Additional knowledge for determining and interpreting feasible band boundaries in self-modeling/multivariate curve resolution of two-component systems

Róbert Rajkó*
Department of Mechanical and Process Engineering, Faculty of Engineering, University of Szeged, P.O. Box 433, H-6701 Szeged, Hungary

**A R T I C L E   I N F O**

Article history:
Received 29 July 2009
Received in revised form 16 December 2009
Accepted 17 December 2009
Available online 24 December 2009

Keywords:
Self-modeling curve resolution (SMCR)
Band boundaries of feasible solutions
Signal contribution function
Grid method

**A B S T R A C T**

Recently Tauler’s mcrbands Matlab script and Maeder’s grid method were used by Abdollahi et al. to calculate the elements of transformation matrix for obtaining feasible band boundaries in multivariate curve resolution of a two-component system. Neither method is analytical, instead they are iterative. For long time it is well-known that Lawton and Sylvester’s approach can provide the feasible band boundaries analytically and non-iteratively. In this paper, firstly in the literature, the clear relationship is given between Lawton and Sylvester’s approach and Tauler’s approach (as well as Maeder’s approach). It was found that all approaches are identical for noiseless or moderately noisy two-component systems and, it was illustrated by figures and tables provided in Supplementary Material.

© 2009 Elsevier B.V. All rights reserved.

**1. Introduction**

Very recently Abdollahi et al. [1] has published a paper on calculation and interpretation of feasible band boundaries for multivariate curve resolution (MCR) estimated profiles of a system containing two components. The paper really contains several excellent novel and relevant results relating to two-component systems. The authors have used grid method and the optimization of the so-called “signal contribution function” (SCF):

\[ f_k(T) = \frac{||c_k s_k^T||}{||c_k s_k^T||}, \]  

(1)

where \( T \) is the rotation (transformation) matrix; \( C \) is an \( I \times K \) matrix of the estimated composition profiles; \( S \) is a \( J \times K \) matrix of the estimated component spectra; and \( c_k \) and \( s_k \) are the \( k \)th columns (i.e., profiles for the \( k \)th component) of the \( C \) and \( S \), respectively. The method based on the numerator of Eq. (1) was introduced first by Gemperline [2], and later it was generalized in the present form by Tauler [3].

In Ref. [1], the authors have introduced a key equation (namely Eq. (13)) which determines the exact relationship between the numerator of the SCF and the elements of the transformation matrix for two-component systems; the generalized form of the equation is

\[ ||c_k s_k^T|| = ||c_k|| ||s_k^T|| = \frac{\sqrt{\sigma_{11}^2 t_{13}^2 + \sigma_{22}^2 t_{23}^2 - k_1 t_{12} - t_{12}^2}}{t_{11} t_{22} - t_{12} t_{21}}, \]  

(2)

where \( \sigma_{11} \) and \( \sigma_{22} \) are the first and second singular values given by the singular value decomposition with \( \sigma_{11} > \sigma_{22}, k_1 \) and \( t_{12} \) are the elements of the transformation matrix \( T \), and \( k = 1 \) or 2 (see more details below).

For finding the minimum and maximum of Eq. (13) of Ref. [1] (which can be calculated by making pre-transformation using Eq. (13) shown later), the results can be obtained analytically by forming the partial derivatives of the equation with respect to \( t_{12} \) and \( t_{21} \); and then finding the roots of the derivatives. After some algebraic manipulations, surprisingly the solution is rather simple: for both components \( t_{12} = 0 \) and \( t_{21} = 0 \), the transformation matrix is the unit one which means no transformation. For components 1 and 2 the values of the SCF are \( \sigma_{11} \) and \( \sigma_{22} \), i.e. the first and second singular values, respectively. Since the determinants of the matrices of the second-order partial derivatives of \( f_k(T) \) are negative and positive (\( \sigma_{22} - \sigma_{11} \) and \( \sigma_{11} - \sigma_{22} \), respectively), thus there is no extreme value for component 1, but for component 2 the extremum is minimum because (\( \partial^2 f_2(T)/\partial t_{12}^2 \)|\( t_{12}=0,t_{21}=0 = \sigma_{11}^2/\sigma_{22} \) > 0. It means that \( f_2(T) \) must have greater and greater values if the singular value decomposition is more and more transformed. It should be mentioned that the denominator of Eq. (13) of Ref. [1] must be greater than 0 which is naturally fulfilled if \( t_{12} = 0 \) and \( t_{21} = 0 \),

0003-2670/$ – see front matter © 2009 Elsevier B.V. All rights reserved.
natural one. Thus the above trivial reasoning cannot be used for practical calculation of $t_{12}$ and $t_{21}$. Instead we utilize the original Lawton and Sylvestre (LS) [5] approach as an alternative way to calculate the elements of the transformation matrices $T$ and $T$ beside grid method and optimization of Eq. (1) used both in Ref. [1]. Recently, for example, Booksh [6], Rajkó and István [7], and Rutan et al. [8] dealt with LS method.

In this paper, using and complementing the derivations of Abdollahi et al. [1], some additional results will be presented for two-component systems. For the same numerical example we reproduced the synthetic data used by Abdollahi et al. [1] with the Matlab [9] script Code 1 presented in Supplementary Material (SM).

2. Theoretical foundation

The original data matrix $R = CS^T$ can be decomposed using principal component analysis (PCA) or singular value decomposition (SVD):

$$R = XV^T = UDV^T,$$

where $X$ and $V$ are the orthogonal score and orthonormal load-

matrices, respectively, $U$ and $V$ are both orthonormal matrices containing left and right eigenvectors, respectively, and $D$ is a diagonal matrix containing singular values $\sigma_k$ in its diagonal and zero elsewhere.

The abstract solution $X$ and $V$ can be transformed to the physically interpretable $C$ and $S$ with some appropriate transformation matrix $T$:

$$R = (XT^{-1})(TV^T) = CS^T,$$

It is known that the norm of $s_k$ and $c_k$ cannot be uniquely determined without calibration samples with known concentrations, thus one can prescribe a certain scale constraint for the pure spectral profiles exclusive or the pure concentration profiles. In addition the response matrix $R$ can be normalized to have unit norm using so-called Borgen norms [10] (|$x$|$_{\beta_k} = z^T|x|$ for $z > 0$) in two ways according to the rows and columns, respectively:

$$\bar{R}_C = R / \sqrt{\left(\frac{R}{I_{lx_j}}\right)_{lx_j}} = \left(\frac{R}{I_{lx_j}}\frac{z}{I_{lj}}\frac{1}{I_{lj}}\right),$$

$$\bar{R}_s = R / \sqrt{\left(\frac{1}{I_{lx_z}}\frac{z}{I_{lj}}\frac{1}{I_{lj}}\right)}.$$

where $j$ means the element-wise matrix division according to the Matlab [9] concept and $I$ is a suitable sized vector of ones.

The normalized response matrix $\bar{R}$ can be also decomposed using PCA and because of the non-negativity:

$$\bar{R} = \bar{X}\bar{V}^T \geq 0.$$

Applying a suitable transformation matrix $\bar{T}$, the physically interpretable solution can be obtained here as well:

$$R = (XT^{-1})(TV^T) = CS^T.$$

All points represented by the rows of the matrix $\bar{X}$ in the abstract space spanned by the columns of the matrix $\bar{V}$ will be constrained to a line (see the black line defined by points 1–5 in Fig. 1) because of the chosen unit 1-norm (one of the Borgen norms [10] when $z=1$) normalization of rows of $\bar{R}$:

$$\bar{R} = \bar{X}\bar{V}^T,$$

$$\bar{R} = \bar{X}\bar{V}^T,$$

$$\bar{R} = \bar{X}\bar{V}^T,$$

$$\bar{R} = \bar{X}\bar{V}^T,$$

$$\bar{R} = \bar{X}\bar{V}^T,$$

$$\bar{R} = \bar{X}\bar{V}^T,$$

$$\bar{R} = \bar{X}\bar{V}^T,$$
The elements of the first eigenvector \( \tilde{v}_1 \), corresponding to the largest eigenvalue, will be always positive (or, if negative, the signs of all elements may be arbitrarily changed). In addition because
\[
\tilde{X} \tilde{v}_1^T \geq 0 \Rightarrow \tilde{X}^T \tilde{v}_1 \geq 0 \Rightarrow \tilde{X}^T \tilde{v}_1^T \geq 0 \Rightarrow \tilde{v}_1^T \tilde{X}^T \tilde{v}_1 \geq 0,
\]
there are \( J \) half-planes to restrict the largest possible positive valued region in the plane spanned by \([v_1, v_2]^T\), see the area between the continuous lines crossing zero (outer bounds) in Fig. 1. On the inside, the retransformed points (i.e., the profiles) have no negative values.

Because of the non-negativity constraint, convex linear combination of the two outermost points (points 3 and 5 in Fig. 1) can make all points of \( \tilde{X} \). The broken lines crossing zero and the outermost points of Fig. 1 marks the inner bounds. The areas between the outer and the inner bounds will be the feasible regions according to the two components.

Intersections of the straight line defined by Eq. (10) and the straight lines of the outer and the inner bounds produce 4–4 points for components 1 and 2, respectively, whose coordinates can be used to create a suitable transformation matrix \( T \). In Fig. 1 for component 1 \( t_{11}, t_{12} \) are the \( \tilde{x}_1 \) and \( t_{12}, t_{22} \) are the \( \tilde{x}_2 \) coordinates:
\[
T = \begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix}.
\]

In Fig. 1 the indicated transformation matrix elements are for getting the minimum of the SCF of component 1 and at the same time the maximum of the SCF of component 2 in spectral space.

Using the simplification suggested by Vosough et al. [4], i.e. transforming the value of the main diagonal elements of \( T \) to one:
\[
\tilde{T} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 \tilde{t}_{12} \\ \tilde{t}_{21} \end{pmatrix} = \begin{pmatrix} 1 \tilde{t}_{12} \\ 1 \tilde{t}_{21} \end{pmatrix}.
\]

Code 2 in SM presents how to calculate the transformation matrices based on LS method in the form used in Tauler’s mcrbands Matlab script [11]. Some more numerical details on LS calculations can be found in the literature [5,12,13].

Tauler’s Matlab script provides \( \delta \) as one of the outputs which is the collection of the \( 2 \times 2 \) transformation matrices in the following manner: in the first two rows the \( 2 \times 2 \) matrix is for getting the maximum of the SCF of component 1, in the second two rows the \( 2 \times 2 \) matrix is for getting the minimum of the SCF of component 1 in the third two rows the \( 2 \times 2 \) matrix is for getting the maximum of the SCF of component 2, and in the final two rows the \( 2 \times 2 \) matrix is for getting the minimum of the SCF of component 2.

Fig. 1 (hypothetically) and Fig. 2 (depicted using data reproduced by Code 1 in SM) help to comprehend graphically the relationship between Tauler’s approach and Lawton and Sylvester’s approach. In Fig. 2 \( xAl \), \( xBl \), \( xBu \) and \( xBd \) are the intersection points of the straight line defined by Eq. (10) and the straight lines of the outer and the inner bounds. For getting the maximum (minimum) of the SCF of component 1, the elements of the transformation matrix \( \tilde{T} \) will be the coordinates of the points \( xBl \) and \( xAl \) (\( xBu \) and \( xAa \)):
\[
\tilde{t}_{12, \text{MAX}, 1} = \frac{\tilde{x}_{3,xBl}}{\tilde{x}_{1,xBl}}, \quad \tilde{t}_{21, \text{MAX}, 1} = \frac{\tilde{x}_{1,xBu}}{\tilde{x}_{2,xBu}};
\]
\[
\tilde{t}_{12, \text{MIN}, 1} = \frac{\tilde{x}_{2,xBu}}{\tilde{x}_{1,xBu}}, \quad \tilde{t}_{21, \text{MIN}, 1} = \frac{\tilde{x}_{1,xAl}}{\tilde{x}_{2,xAl}}.
\]

For getting the maximum (minimum) of the SCF of component 2, the elements of the transformation matrix \( \tilde{T} \) will be the coordinates of the points \( xBu \) and \( xAl \) (\( xBl \) and \( xAa \)):
\[
\tilde{t}_{12, \text{MAX}, 2} = \frac{\tilde{x}_{2,xBu}}{\tilde{x}_{1,xBu}}, \quad \tilde{t}_{21, \text{MAX}, 2} = \frac{\tilde{x}_{1,xAl}}{\tilde{x}_{2,xAl}};
\]
\[
\tilde{t}_{12, \text{MIN}, 2} = \frac{\tilde{x}_{2,xBl}}{\tilde{x}_{1,xBl}}, \quad \tilde{t}_{21, \text{MIN}, 2} = \frac{\tilde{x}_{1,xBu}}{\tilde{x}_{2,xBu}}.
\]

3. Results and discussion

The left and right panels of Fig. 3 show the mesh plots of \( f_0(\tilde{T}) \) for components 1 and 2. This figure helps for explaining why Eq. (14) and (15) must be applied. For component 2 the explanation of the situation is easier, because \( f_2(\tilde{T}) \) has minimum at \( \tilde{t}_{12} = 0 \) and \( \tilde{t}_{31} = 0 \). \( f_2(\tilde{T}) \) has the smallest value nearest the point \((0,0)\), and the largest value farthest from point \((0,0)\). In Eq. (15) the values of every \( \tilde{t} \) can be written as tangent of suitable angles: \( \tan(-\beta_2), \tan(\alpha_N), \tan(-\beta_N) \), respectively (see the angles in Fig. 2). Considering Fig. 2 and right panel of Fig. 3, \( f_2(\tilde{T}) \) has the largest value at largest angles, and smallest value at smallest angles. It is not true for component 1, because \( f_1(\tilde{T}) \) has not any extremum. The function \( f_1(\tilde{T}) \) is, however, strictly monotonic above the feasible region because of the convex linearity property (which means that convex linear combination of the two endpoints of the feasible regions, i.e. \( xBl \) and \( xBu \), or \( xAl \) and \( xAa \) for component 1 or 2, respectively, can make all points of the feasible regions). \( f_1(\tilde{T}) \) should have unique maximum and minimum at the domain boundary (indicated as rectangle in Fig. 3) because of this monotonic property. Considering Fig. 2 and left panel of Fig. 3, \( f_1(\tilde{T}) \) has the largest value at smallest angles and smallest value at largest angles. Now in Eq. (14) the values of every \( \tilde{t} \) can be written as tangent

![Fig. 3](image-url)
of suitable angles: \( \tan(-\beta_N) \), \( \tan(\alpha_N) \), \( \tan(-\beta_F) \), \( \tan(\alpha_F) \), respectively (see Fig. 2).

After the theoretical discussion we give a brief practical discussion as well just for the sake of visual interpretation. Using the synthetic data reproduced by Code 1, Table SM-1–4 (see in Supplementary Material) were generated and the tables confirm that Tauler’s \textit{mcrbands} and LS method calculated transformation matrix elements coincide for noiseless or moderately noisy (standard deviation of the zero mean Gaussian noise used were 0 and 0.01, respectively) cases. There is some noticeable repetition of the values in the transformation matrices, it is because of the natural duality between the concentration and spectral spaces of the minimal constrained curve resolution task, see Fig. 3 in Ref. [14].

Figure SM-4–11 show that in these cases the outer bounds (solid red and blue lines) are really external, and inner bounds (dashed red and blue lines) are really internal, and of course the true profiles marked with green crossed circle or dotted white line are inside the area defined by outer and inner bounds. Using extremely large noise (standard deviation was 0.4) the situation is dramatically changed. Table SM-5 and -6 corroborate that Tauler’s \textit{mcrbands} and LS method calculated transformation matrix elements are now different. Figure SM-12–15 show surprisingly that outer bounds are internal, and inner bounds are external. The true profiles are not always inside the feasible ranges. Even the inner bounds have some negative values (see Figure SM-15) despite of the used non-negativity constraint. The reason is that the two-dimensional subspace selected by PCA contained large amount of the subspace spanned by the noise. Figure SM-16 and -17 depict that \textit{mcrbands} provided worse \( T_{12} \) and \( T_{21} \) values than LS-based method in case of large amount of noise.

4. Conclusions

To sum it up it can be concluded that all three approaches work well and their results are equivalent in the cases of noiseless or moderately noisy (on LS plot the outer and inner bounds must be external and internal, respectively) systems. Although this conclusion was illustrated with a single set of synthetic data, the Matlab code based on the theoretical reasoning above is provided in Supplementary Material: thus anybody can check it out using as many data sets as possible.

As an advantage Tauler’s \textit{mcrbands} and Maeder’s grid method can give bounds for the original initial estimations, while LS-based method can provide only normalized bounds, but this latter approach is analytical and non-iterative. It should be mentioned that for three- or more component systems the situation is much more complex and the outcomes obtained for two-component systems cannot be straightforwardly generalized [15].

Acknowledgements

Prof. Attila Felinger is acknowledged for his suggestion (made at the conference ACE-CC 2003, Budapest) to investigate Lawton and Sylvestre’s method, which turned the author’s interest on S/MCR methods.

Appendix A. Supplementary data


References