JOURNAL OF CHEMOMETRICS
*J. Chemometrics* 2001; **15**: 149–167

# Pseudo alternating least squares algorithm for trilinear decomposition

Zeng-Ping Chen, Yang Li and Ru-Qin Yu*

*College of Chemistry and Chemical Engineering, Hunan University, Changsha 410082, People's Republic of China*

SUMMARY

In chemistry, PARAFAC is one of the most widely used algorithms for trilinear decomposition. However, the problem of PARAFAC requiring an accurate estimation of the number of factors in the system under study limits its applications to some extent. This troublesome problem has been tackled by the pseudo alternating least squares (PALS) algorithm designed in this paper. PALS is a unique algorithm which tries to alternately optimize three different objective functions to obtain the solutions for the trilinear decomposition model. It has the outstanding feature of being resistant to the influence of $N$ (the number of factors chosen in calculation), which has been proved mathematically under some mild conditions. Although the optimization procedure of PALS is different from that of PARAFAC, an alternating least squares scheme, and hinders a straightforward analysis of its convergence properties, studies on simulated as well as real data arrays reveal that PALS can often converge to satisfactory results within a reasonable computation time, even if excess factors are used in calculation. 

KEY WORDS: trilinear decomposition model; underlying factors; loading matrixes; PARAFAC; alternating least squares algorithm; PALS

## 1. INTRODUCTION

With the increasing popularity of advanced instruments capable of generating multidimensional arrays, trilinear decomposition models with the so-called second-order advantage have become an area of much current activity in chemometrics. All the methods developed for trilinear decomposition may be classified into two main categories, i.e. non-iterative [1–9] and iterative [10–15] algorithms. Within the aiterative trilinear decomposition category, PARAFAC [10] is probably the most widely used method in chemistry. However, it is often criticized for its low convergence rate and requirement of an accurate estimation of the number of underlying factors [13]. Concerning the problem of slow convergence, some remedies have been proposed, [16–18] but to date, as far as we know, there has been little work on the latter problem. *Alternative trilinear decomposition* (ATLD), recently proposed by Wu *et al* [15], may be the first iterative trilinear decomposition algorithm to possess the capacity of being insensitive to the number of factors chosen in calculation. ATLD has the other prominent

* Correspondence to: R.-Q. Yu, College of Chemistry and Chemical Engineering, Hunan University, Changsha 410082, People's Republic of China.
E-mail: rqyu@mail.hunu.edu.cn
Contract/grant sponsor: National Natural Science Foundation of China; Contract/grant number: 29735150


*Received 14 January 2000*
*Accepted 30 March 2000*

advantage of fast convergence. It can often converge to satisfactory solutions within a few iterations. Unfortunately, ATLD also has its deficiencies, such as a symmetry constraint on the data array (i.e. the three modes of the data array should have equal rank) and the introduction of possible optima other than the solutions desired. Therefore, although ATLD has the above prominent advantages over other methods, some improvements or remedies may be needed for wider applications.

In this paper, efforts are made to develop a new algorithm, called *pseudo alternating least squares* (PALS), which completely or partially overcomes the deficiencies of ATLD. Some interesting properties of PALS under mild conditions are proved mathematically in Appendix II, while other features such as convergence are scrutinized by simulations in Section 5. An application to a real data array is described in Section 6.

## 2. NOMENCLATURE

Throughout this paper, scalars are represented by lower-case italics, vectors are denoted by bold lower-case characters, bold capitals designate two-way matrices and underlined bold capitals symbolize three-way arrays. Before reading the main text of this paper, readers are recommended to refer to Appendix I for detailed nomenclature information.

## 3. THE MODEL

All the discussion in this paper is based on the famous trilinear decomposition model proposed by Harshman [10] and Carroll and Chang [11]:

$$x_{ijk} = \sum_{f=1}^{F} a_{if} b_{jf} c_{kf} + e_{ijk}, \quad i = 1, \ldots, I, \quad j = 1, \ldots, J, \quad k = 1, \ldots, K \tag{1}$$

For chemometricians the following matrix form may be more convenient and clearer to understand:

$$\mathbf{X}_{..k} = \mathbf{A}\mathrm{diag}(\mathbf{c}_k^{\mathrm{T}})\mathbf{B}^{\mathrm{T}} + \mathbf{E}_{..k}, \quad k = 1, \ldots, K \tag{2}$$

As proved by Kruskal [19,20], the decomposition of the above model is unique regardless of some scaling and permutation indeterminacy under some mild conditions, i.e. correct estimation of the number of factors and $k_1 + k_2 + k_3 \geq 2F + 2$ ($k_1$, $k_2$ and $k_3$ are the $k$-ranks of the underlying loading matrices **A, B** and **C** respectively).

## 4. THE ALGORITHM

The most straightforward way to obtain the underlying loading matrices **A, B** and **C** is by the alternating least squares approach proposed by Harshman along with the model (which is often called PARAFAC). PARAFAC is popular in psychology as well as in chemometrics. However, it is often criticized for its slow convergence and requirement of the correct choice of the number of factors. Although many alternative ways have been proposed to accelerate the optimization procedure of PARAFAC [16–18], most of them take advantage of the compression technique. As far as we know, little attention has been paid to how to relieve experimenters from the troublesome factor estimation task. Both problems of PARAFAC have been touched on by ATLD [15], and encouraging results such as fast convergence and insensitivity to excess factors used have been observed. Unfortunately, a symmetry constraint (rank(**A**) = rank(**B**) = rank(**C**)) on the data array limits its applications. Our recent study revealed that the symmetry constraint was introduced by the objective functions employed in ATLD. Moreover, it was also found that the utilization of

 

$$\sum_{k=1}^{K}\left\|\mathbf{X}_{..k}((\hat{\mathbf{B}})^{\mathrm{T}})^{+}-\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})\right\|_{\mathrm{F}}^{2} \quad \text{or} \quad \sum_{k=1}^{K}\left\|(\hat{\mathbf{A}})^{+}\mathbf{X}_{..k}-\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2}$$

as objective function will endow the algorithm established on it with the property of insensitivity to excess factors used in calculation [21]. Therefore it is theoretically sound to eliminate the influence of excess factors on the final results when one of the above functions is minimized to obtain the solutions. Nevertheless, the multi-optimum feature of

$$\min\sum_{k=1}^{K}\left\|\mathbf{X}_{..k}((\hat{\mathbf{B}})^{\mathrm{T}})^{+}-\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})\right\|_{\mathrm{F}}^{2} \quad \text{or} \quad \min\sum_{k=1}^{K}\left\|(\hat{\mathbf{A}})^{+}\mathbf{X}_{..k}-\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2}$$

under the assumption of absence of errors will sometimes cause it to converge to undesired solutions such as $\hat{\mathbf{A}}=0$, $\hat{\mathbf{B}}=0$ and $\hat{\mathbf{C}}=0$. Fortunately, a simple combination of the PARAFAC loss function with one of the above objective functions can circumvent the multi-optimum problem:

$$\min\sum_{k=1}^{K}\left(\left\|\mathbf{X}_{..k}-\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2}+\lambda\left\|\mathbf{X}_{..k}((\hat{\mathbf{B}})^{\mathrm{T}})^{+}-\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})\right\|_{\mathrm{F}}^{2}\right) \tag{3}$$

or

$$\min\sum_{k=1}^{K}\left(\left\|\mathbf{X}_{..k}-\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2}+\lambda\left\|(\hat{\mathbf{A}})^{+}\mathbf{X}_{..k}-\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2}\right) \tag{4}$$

When optimizing one of the above objective functions (3) or (4) to obtain the solutions, one will no longer be troubled by the multi-optimum problem and symmetry constraint. However, there is not a straightforward optimization procedure to minimize the objective function (3) or (4). Inspired by the success of ATLD which alternately optimizes three different objective functions with intrinsic relationships, in the present paper the following three objective functions $S(\hat{\mathbf{C}})$, $S(\hat{\mathbf{A}})$ and $S(\hat{\mathbf{B}})$ are alternately minimized to find the solutions:

$$\begin{aligned}S(\hat{\mathbf{C}})=\sum_{k=1}^{K}\Bigg(2\left\|\mathbf{X}_{..k}-\mathbf{A}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2}\\ +\lambda\left(\left\|\mathbf{X}_{..k}((\hat{\mathbf{B}})^{\mathrm{T}})^{+}-\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})\right\|_{\mathrm{F}}^{2}+\left\|(\hat{\mathbf{A}})^{+}\mathbf{X}_{..k}-\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2}\right)\Bigg)\end{aligned} \tag{5}$$

(minimizing $S(\hat{\mathbf{C}})$ with fixed $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ to obtain loading matrix $\hat{\mathbf{C}}$)

$$S(\hat{\mathbf{A}})=\sum_{k=1}^{K}\left(\left\|\mathbf{X}_{..k}-\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2}+\lambda\left\|\mathbf{X}_{..k}((\hat{\mathbf{B}})^{\mathrm{T}})^{+}-\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})\right\|_{\mathrm{F}}^{2}\right) \tag{6}$$

(minimizing $S(\hat{\mathbf{A}})$ with fixed $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$ to obtain loading matrix $\hat{\mathbf{A}}$)

 

$$S(\hat{\mathbf{B}}) = \sum_{k=1}^{K}\left(\left\|\mathbf{X}_{..k} - \hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2} + \lambda\left\|(\hat{\mathbf{A}})^{+}\mathbf{X}_{..k} - \mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2}\right) \tag{7}$$

(minimizing $S(\hat{\mathbf{B}})$ with fixed $\hat{\mathbf{A}}$ and $\hat{\mathbf{C}}$ to obtain loading matrix $\hat{\mathbf{B}}$).

Based on the above three objective functions, an algorithm called *pseudo alternating least squares* (PALS) is designed. The main steps of PALS are as follows.

1. Randomly initialize loading matrixes $\hat{\mathbf{A}}_{I\times N}$ and $\hat{\mathbf{B}}_{J\times N}$ and choose an appropriate $\lambda$.

2. $$SC_1 = \sum_{i}^{I}(\mathbf{X}_{i..}^{\mathrm{T}}(2\hat{\mathbf{B}} + \lambda((\hat{\mathbf{B}})^{\mathrm{T}})^{+})\mathrm{diag}((\hat{\mathbf{a}}_i)^{\mathrm{T}})) + \lambda\left(\sum_{j}^{J}(\mathbf{X}_{.j.}((\hat{\mathbf{A}})^{\mathrm{T}})^{+}\mathrm{diag}((\hat{\mathbf{b}}_j)^{\mathrm{T}}))\right)$$

$$SC_2 = \sum_{i}^{I}(\mathrm{diag}((\hat{\mathbf{a}}_i)^{\mathrm{T}})(2(\hat{\mathbf{B}})^{\mathrm{T}}\hat{\mathbf{B}} + \lambda\mathbf{I})\mathrm{diag}((\hat{\mathbf{a}}_i)^{\mathrm{T}})) - \lambda\left(\sum_{j}^{J}(\mathrm{diag}((\hat{\mathbf{b}}_j)^{\mathrm{T}})\mathrm{diag}((\hat{\mathbf{b}}_j)^{\mathrm{T}}))\right)$$

$$\hat{\mathbf{C}} = SC_1 \times (SC_2)^{+}$$

3. $$\hat{\mathbf{A}} = \left(\sum_{k}^{K}(\mathbf{X}_{..k}(\hat{\mathbf{B}} + \lambda((\hat{\mathbf{B}})^{\mathrm{T}})^{+})\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}}))\right)\left(\sum_{k}^{K}(\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})((\hat{\mathbf{B}})^{\mathrm{T}}\hat{\mathbf{B}} + \lambda\mathbf{I})^{+})\right)$$

4. $$\hat{\mathbf{B}} = \left(\sum_{k}^{K}(\mathbf{X}_{..k}(\hat{\mathbf{A}} + \lambda((\hat{\mathbf{A}})^{\mathrm{T}})^{+})\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}}))\right)\left(\sum_{k}^{K}(\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})((\hat{\mathbf{A}})^{\mathrm{T}}\hat{\mathbf{A}} + \lambda\mathbf{I})^{+})\right)$$

5. Update $\hat{\mathbf{C}}$, $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ according to steps 2–4 until a certain stopping criterion has been reached.

Under some mild conditions, some interesting properties of PALS are explained in a more or less rigorous mathematical way in Appendix II. The optimization scheme of PALS is not as straightforward as that of PARAFAC. It seems impossible to provide a rigorous mathematical analysis of its convergence behaviour, which will be investigated by simulations in the following section.

## 5. SIMULATION STUDIES OF THE PROPERTIES OF PALS

### *5.1. Simulated data arrays*

It is obvious that the optimization procedure of PALS is not well defined when $\lambda \neq 0$. However, the three objective functions (5)–(7) have an intrinsic relationship (when $\lambda = 0$, PALS is equivalent to PARAFAC). Therefore, although the optimization procedure of PALS may not be a consistently

 

decreasing one as in PARFAC, it is reasonable to expect that PALS can converge to satisfactory solutions. Because of the lack of a straightforward description of the optimization procedure of PALS, its performance has to be demonstrated by simulation studies. Data arrays were simulated according to the following scheme:

$$\mathbf{A} = \mathbf{rand}(I, F)$$

$$\mathbf{B} = \mathbf{rand}(J, F)$$

$$\mathbf{C} = \mathbf{rand}(K, F)$$

$$\mathbf{X}^p_{..k} = \mathbf{A}\mathrm{diag}(\mathbf{c}_k^{\mathrm{T}})\mathbf{B}^{\mathrm{T}}$$

$$\mathbf{E}_{..k} = \mathrm{Max}(\mathbf{X}^p_{..k}) \times \mathbf{randn}(I, J) \times a_{\mathrm{noise}}, \qquad k = 1, 2, \ldots, K$$

$$\mathbf{X}_{..k} = \mathbf{X}^p_{..k} + \mathbf{E}_{..k}, \qquad k = 1, 2, \ldots, K$$

$$\underline{\mathbf{X}} = \{\mathbf{X}_{..k}\}$$

where $a_{\mathrm{noise}}$ is a scalar controlling the noise level added, **rand**(*I*,*F*) is an $I \times F$ matrix with random elements taking values in the range (0,1), **randn**(*I*,*J*) is an $I \times J$ matrix with random entries chosen from a normal distribution with mean zero and variance one, and $\mathrm{Max}(\mathbf{X}^p_{..k})$ is the maximal entry of matrix $\mathbf{X}^p_{..k}$.

### 5.2. *Implementation of PALS*

For all the data arrays, random initialization was carried out to start the iterative optimization procedure of PALS. The optimization procedure is terminated when the following criterion reaches a certain threshold $\varepsilon$ ($\varepsilon = 1 \times 10^{-6}$ in the present paper)—a maximal iteration number of 3000 is adopted to avoid possible unduly slow convergence (a maximal iteration number of 5000 is set for PARAFAC):

$$SSR^{(m)} = \sum_{k=1}^{K} \left\| \mathbf{X}_{..k} - (\hat{\mathbf{A}})^{(m)} \mathrm{diag}(((\hat{\mathbf{c}}_k)^{\mathrm{T}})^{(m)})((\hat{\mathbf{B}})^{\mathrm{T}})^{(m)} \right\|_{\mathrm{F}}^2$$

$$\left| \frac{SSR^{(m)} - SSR^{(m-1)}}{SSR^{(m-1)}} \right| \leq \varepsilon$$

where *SSR* is the residual sum of squares and *m* is the current iteration number.

### 5.3. *Convergence properties of PALS*

Ten random data arrays of size $20 \times 20 \times 5$ were simulated. All the data arrays are of three-component systems. Into each data array, noise with $a_{\mathrm{noise}}$ ranging from 0·004 to 0·02 at intervals of 0·004 was added. Therefore a total of 50 data arrays were simulated to investigate the convergence properties of PALS. Each data array was decomposed 20 times by PALS with $N = 3$, $\lambda = 1$ and random initialization.

Out of the 1000 trials, 98·4% of runs have reached the desired global optima (or near global optima), and the obtained loading matrices $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$ are in perfect agreement with the actual ones (Figure 1 shows one example). Only 16 runs have been trapped in local optima, which can be easily

 

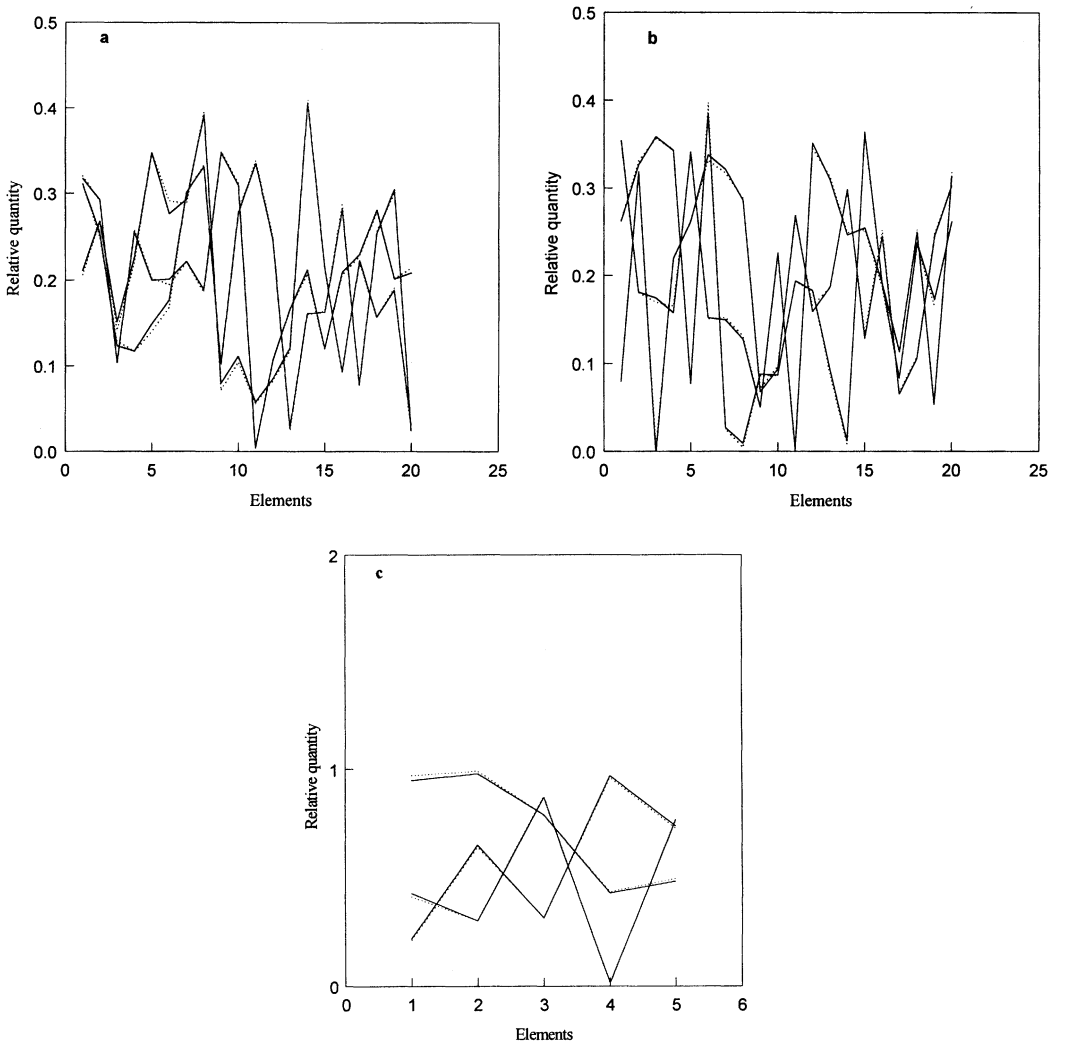


Figure 1. True (dotted line) and resolved (full line) loading matrices (a, **A**; b, **B**; c, **C**) in three modes by PALS ($N = 3$ and $\lambda = 1$) for a randomly simulated three-component data array of size $20 \times 20 \times 5$ with $a_{\text{noise}} = 0{\cdot}02$.

perceived from the obvious difference between their residual sums of squares (*SSR*) and those when the global optima are attained (for instance, $SSR = 0{\cdot}0935$ for a certain local optimum, while $SSR = 0{\cdot}0230$ for the global optimum for a data array with $a_{\text{noise}} = 0{\cdot}02$). The optimization pattern of PALS is different from that of PARAFAC. In our studies, *SSR* in 52·1% of trials decreases consistently (Figure 2a), and most of the other runs show a minor increase in *SSR* after a consistent decrease (Figure 2b). The increase is so small (for instance, from $SSR = 9{\cdot}21318329 \times 10^{-4}$ to $SSR = 9{\cdot}21318516 \times 10^{-4}$ during 10 iterations) that it should not influence the final results (actually, it does not). It should be mentioned that out of the 1000 trials, there exist several trials with oscillating optimization behaviour. To our surprise, these 'abnormal individuals' have also attained the desired global optima. In our opinion the ability of PALS to converge to the desired optima may be due to the

 

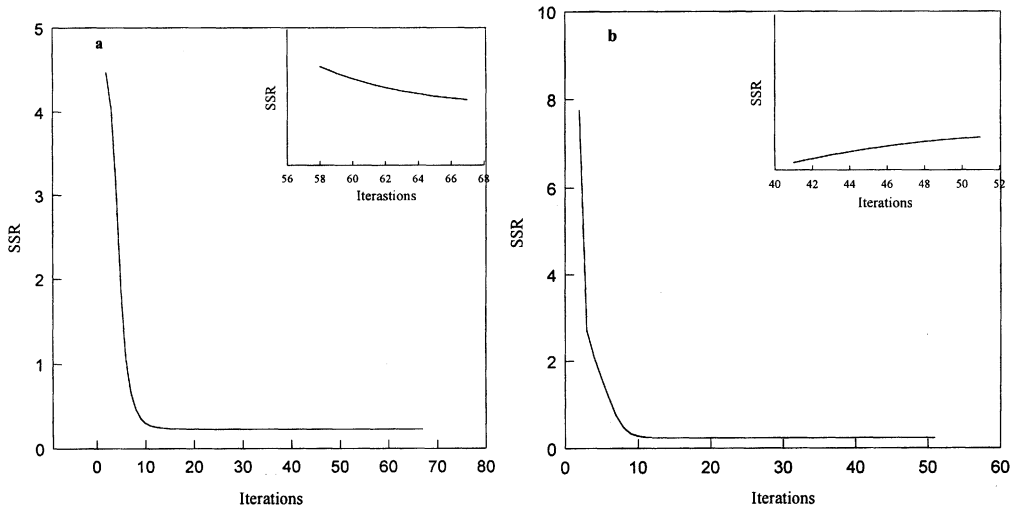


Figure 2. Two typical convergence patterns of PALS with $N = 3$ and $\lambda = 1$ for randomly simulated three-component data arrays of size $20 \times 20 \times 5$.

intrinsic relationship among the three objective functions employed by PALS. For a deeper understanding of this phenomenon, further work is needed.

Concerning the convergence efficiency, PALS can generally converge within 50–300 iterations, with a high frequency around 100 iterations. Iteration numbers larger than 500 are rarely encountered in our simulation. Varying the noise level $a_{\text{noise}}$ within a certain range (from 0·004 to 0·02) has little influence on the optimization efficiency of PALS. Therefore, although the optimization pattern of PALS may not be a consistently decreasing one as in PARAFAC, its ability to converge to satisfactory results within a reasonable number of iterations is without doubt.

### *5.4. Influence of N (number of factors used in calculation) on PALS*

Although it has been proved in Appendix II that the excess factors used in PALS will not affect the final results, the above conclusion is drawn conditionally on $\underline{\mathbf{E}} = 0$. Therefore, when $\underline{\mathbf{E}} \neq 0$, simulation studies on the possible influence of excess factors on the performance of PALS are necessary.

Table I lists the results of PALS ($\lambda = 1$) for a randomly simulated three-component data array of size $20 \times 20 \times 5$ with $a_{\text{noise}} = 0{\cdot}02$. As expected, an increase in the number of excess factors has little effect on the quality of the results obtained by PALS. The choice of an $N$ that is two times the number of actual factors only slightly deteriorated the quality of the results. Even when $N = 10$, the resolved profiles of the underlying factors in the three modes are in good agreement with the true ones (Figure 3). It is worth pointing out that when excess factors are used in calculation, the columns of the loading matrices resolved by PALS will either equal the corresponding columns of the underlying loading matrices or represent noise. Moreover, for every column of the underlying loading matrices there is an equal column in the corresponding loading matrices, regardless of some scale indeterminacy. The noise profiles can be distinguished from the profiles of the underlying factors, since they make a very small contribution to the variances of the data arrays compared to that of the underlying factors. In chemical applications the task is further simplified by the fact that the profiles of the underlying

 

Table I. Influence of $N$ on final results of PALS ($\lambda = 1$) for a randomly simulated data array of three components with $a_{noise} = 0{\cdot}02$ (for each $N$ value, five randomly initialized runs were performed)

| | | Mode 1 | | | Mode 2 | | | Mode 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | IND[a] | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |
| 3 | Max | 0·9997[b] | 0·9999 | 0·9994 | 0·9997 | 0·9997 | 0·9998 | 0·9998 | 0·9998 | 0·9996 |
| | Min | 0·9997 | 0·9999 | 0·9994 | 0·9997 | 0·9997 | 0·9998 | 0·9998 | 0·9998 | 0·9996 |
| | Ave | 0·9997 | 0·9999 | 0·9994 | 0·9997 | 0·9997 | 0·9998 | 0·9998 | 0·9998 | 0·9996 |
| | Var | 0[c] | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | Max | 0·9998 | 0·9999 | 0·9994 | 0·9997 | 0·9997 | 0·9998 | 0·9998 | 0·9998 | 0·9996 |
| | Min | 0·9997 | 0·9999 | 0·9994 | 0·9997 | 0·9997 | 0·9998 | 0·9998 | 0·9998 | 0·9996 |
| | Ave | 0·9998 | 0·9999 | 0·9994 | 0·9997 | 0·9997 | 0·9998 | 0·9998 | 0·9998 | 0·9996 |
| | Var | $2{\cdot}0\times10^{-9}$ | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | Max | 0·9998 | 0·9999 | 0·9995 | 0·9997 | 0·9997 | 0·9998 | 0·9999 | 0·9998 | 0·9996 |
| | Min | 0·9997 | 0·9970 | 0·9993 | 0·9997 | 0·9981 | 0·9998 | 0·9998 | 0·9997 | 0·9996 |
| | Ave | 0·9998 | 0·9993 | 0·9994 | 0·9997 | 0·9994 | 0·9998 | 0·9998 | 0·9998 | 0·9996 |
| | Var | $3{\cdot}0\times10^{-9}$ | $1{\cdot}7\times10^{-6}$ | $7{\cdot}0\times10^{-9}$ | 0 | $5{\cdot}1\times10^{-7}$ | 0 | $3{\cdot}0\times10^{-9}$ | $2{\cdot}0\times10^{-9}$ | 0 |
| 6 | Max | 0·9997 | 0·9999 | 0·9994 | 0·9997 | 0·9997 | 0·9998 | 0·9999 | 0·9998 | 0·9997 |
| | Min | 0·9997 | 0·9984 | 0·9993 | 0·9997 | 0·9968 | 0·9998 | 0·9998 | 0·9996 | 0·9996 |
| | Ave | 0·9997 | 0·9996 | 0·9994 | 0·9997 | 0·9991 | 0·9998 | 0·9999 | 0·9998 | 0·9997 |
| | Var | 0 | $4{\cdot}5\times10^{-7}$ | $2{\cdot}0\times10^{-9}$ | 0 | $1{\cdot}7\times10^{-4}$ | 0 | $2{\cdot}0\times10^{-9}$ | $8{\cdot}0\times10^{-9}$ | $3{\cdot}0\times10^{-9}$ |
| 10 | Max | 0·9998 | 0·9999 | 0·9994 | 0·9997 | 0·9996 | 0·9998 | 0·9998 | 0·9997 | 0·9996 |
| | Min | 0·9997 | 0·9846 | 0·9992 | 0·9996 | 0·9881 | 0·9993 | 0·9996 | 0·9987 | 0·9993 |
| | Ave | 0·9997 | 0·9968 | 0·9993 | 0·9996 | 0·9971 | 0·9996 | 0·9997 | 0·9995 | 0·9995 |
| | Var | $3{\cdot}0\times10^{-9}$ | $4{\cdot}6\times10^{-5}$ | $5{\cdot}0\times10^{-9}$ | $2{\cdot}0\times10^{-9}$ | $2{\cdot}5\times10^{-5}$ | $5{\cdot}3\times10^{-8}$ | $1{\cdot}2\times10^{-8}$ | $2{\cdot}0\times10^{-7}$ | $1{\cdot}2\times10^{-8}$ |

[a] IND, Max, Min, Ave and Var denote 'index', 'maximum', 'minimum', 'average' and 'variance' respectively.
[b] 0·9997 is the relative coefficient between the resolved and the true profile. For convenience of presentation, all relative coefficients in this paper have only four significant digits after the decimal point.
[c] Since the numbers following the fourth decimal digit after the decimal point are cut away, PALS and PARFAC will sometimes have replications with no 'apparent variance', though actually the algorithms converge to different answers at machine precision.

chemical factors generally have some distinguishable characteristics, while those of the noise do not. In contrast, a good quality of the performance of PARAFAC depends heavily on the correct choice of the number of factors (Table II). Just one excess factor can cause the final results to be misleading (the relative coefficient between resolved and true loading in the second mode may be as small as 0·2035). The more excess factors that are used, the lower is the confidence in obtaining satisfactory results and the worse is the performance of PARAFAC. Owing to the difficulty in estimating the actual number of factors, compared with PARAFAC, PALS may be more user-friendly in practice. Since a corresponding increase in iterations following an increment of $N$ was observed for PARAFAC as well as PALS (Table III), the gross estimation of the number of underlying factors in the data arrays may be compensated by the reduction in computation time, and hence recommended in practice.

### 5.5. *Choice of* $\lambda$

The behaviour of PALS with respect to the parameter $\lambda$ was also scrutinized. Table IV indicates that a very small $\lambda$ (such as 0·001) is not enough to ensure that the final results of PALS are immune from excess factors. Different runs may converge to different final results which have almost the same SSR. On the other hand, a larger $\lambda$ (such as $\lambda = 1$) will endow PALS with the capability of being insensitive to excess factors. All 10 runs with $\lambda = 1$ have converged to satisfactory results. A further increase in $\lambda$ will make PALS perform even better in terms of variance among different trials and

 

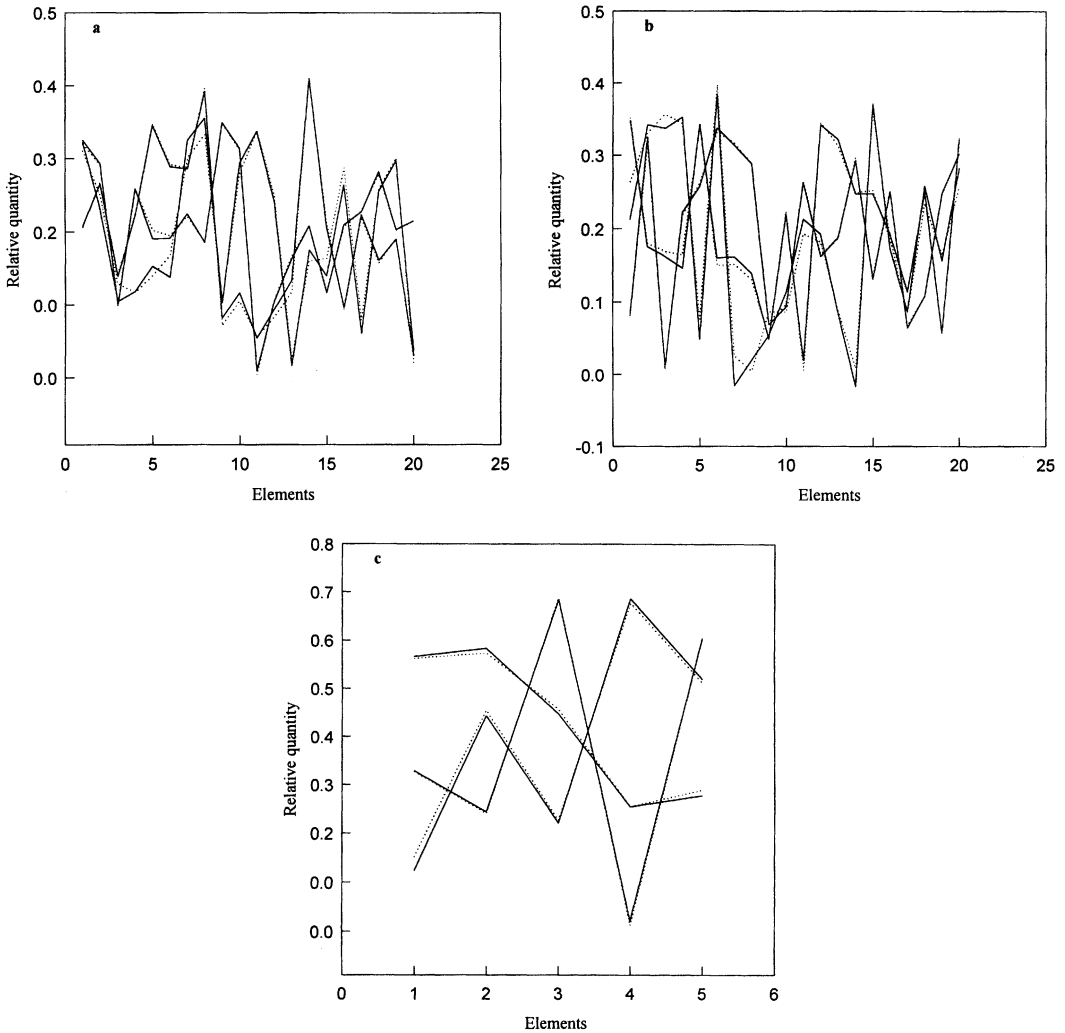


Figure 3. True (dotted line) and resolved (full line) loading matrices (a, **A**; b, **B**; c, **C**) in three modes by PALS ($N = 10$ and $\lambda = 1$) for a randomly simulated three-component data array of size $20 \times 20 \times 5$ with $a_{\text{noise}} = 0{\cdot}02$.

computational burden, although in our simulations no obvious deterioration of the quality of results has been observed with $\lambda$ varying from 0·01 to $10^6$. A large $\lambda$ such as $10^6$ should be used with special care because of the large bias it might introduce. In practice, therefore, a moderate value such as $\lambda = 1$ is preferred.

## 6. REAL DATA ARRAY

The real data array used in this section is a collection of fluorescent excitation–emission response matrices of 11 mixtures of tyrosine, tryptophan and phenylalanine with different concentration ratios. All the response matrices were recorded by a Hitachi 850 fluorescence spectrophotometer with

 

Table II. Influence of $N$ on final results of PARAFAC for a randomly simulated data array of three components with $a_{noise} = 0{\cdot}02$ (for each $N$ value, five randomly initialized runs were performed)

| | | Mode 1 | | | Mode 2 | | | Mode 3 | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | IND | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |
| 3 | Max | 0·9999 | 0·9999 | 0·9995 | 0·9998 | 0·9998 | 0·9999 | 0·9999 | 0·9998 | 0·9999 |
| | Min | 0·9999 | 0·9999 | 0·9995 | 0·9998 | 0·9998 | 0·9998 | 0·9999 | 0·9998 | 0·9999 |
| | Ave | 0·9999 | 0·9999 | 0·9995 | 0·9998 | 0·9998 | 0·9998 | 0·9999 | 0·9998 | 0·9999 |
| | Var | 0 | 0 | 0 | 0 | 0 | $2{\cdot}0\times10^{-9}$ | 0 | 0 | 0 |
| 4 | Max | 0·9999 | 0·9999 | 0·9996 | 0·9998 | 0·9998 | 0·9998 | 0·9999 | 0·9998 | 0·9999 |
| | Min | 0·9988 | 0·9979 | 0·8828 | 0·7155 | 0·2035 | 0·9923 | 0·8816 | 0·9996 | 0·9770 |
| | Ave | 0·9995 | 0·9995 | 0·9726 | 0·9426 | 0·8405 | 0·9983 | 0·9648 | 0·9998 | 0·9950 |
| | Var | $2{\cdot}5\times10^{-7}$ | $8{\cdot}0\times10^{-7}$ | $2{\cdot}6\times10^{-3}$ | $1{\cdot}6\times10^{-2}$ | $1{\cdot}3\times10^{-1}$ | $1{\cdot}1\times10^{-5}$ | $2{\cdot}8\times10^{-3}$ | $8{\cdot}0\times10^{-9}$ | $1{\cdot}0\times10^{-4}$ |
| 5 | Max | 0·9999 | 0·9999 | 0·9992 | 0·9998 | 0·9998 | 0·9998 | 0·9999 | 1·0000 | 1·0000 |
| | Min | 0·9995 | 0·6701 | 0·4240 | 0·5854 | 0·9979 | 0·5326 | 0·8877 | 0·9995 | 0·9938 |
| | Ave | 0·9998 | 0·9150 | 0·7473 | 0·9085 | 0·9993 | 0·8870 | 0·9395 | 0·9998 | 0·9986 |
| | Var | $2{\cdot}8\times10^{-7}$ | $2{\cdot}0\times10^{-2}$ | $6{\cdot}1\times10^{-2}$ | $3{\cdot}3\times10^{-2}$ | $6{\cdot}3\times10^{-7}$ | $4{\cdot}1\times10^{-2}$ | $2{\cdot}7\times10^{-3}$ | $3{\cdot}2\times10^{-8}$ | $7{\cdot}3\times10^{-6}$ |
| 6 | Max | 0·9999 | 0·9994 | 0·9994 | 0·9999 | 0·9994 | 0·9998 | 0·9999 | 0·9999 | 0·9999 |
| | Min | 0·9998 | 0·0428 | 0·8605 | 0·2274 | 0·4475 | 0·9895 | 0·7878 | 0·9993 | 0·6139 |
| | Ave | 0·9999 | 0·6206 | 0·9631 | 0·8453 | 0·8598 | 0·9976 | 0·9241 | 0·9997 | 0·9227 |
| | Var | $3{\cdot}0\times10^{-9}$ | $2{\cdot}7\times10^{-1}$ | $3{\cdot}6\times10^{-3}$ | $1{\cdot}2\times10^{-1}$ | $5{\cdot}5\times10^{-2}$ | $2{\cdot}1\times10^{-5}$ | $1{\cdot}0\times10^{-2}$ | $5{\cdot}5\times10^{-8}$ | $3{\cdot}0\times10^{-2}$ |

excitation and emission wavelengths ranging from 205 to 290 nm and from 270 to 385 nm respectively in intervals of 5 nm. Therefore the data array is of size $18 \times 24 \times 11$.

The results of PARAFAC and PALS for the real data array are listed in Tables V and VI respectively. Since the system under study is a three-component system, it is natural for PARAFAC to attain the best results when $N = 3$. The quality of the results obtained by PARAFAC shows a rapid deterioration as $N$ increases from 3 to 6. The multi-optimum feature of PARAFAC when $N > 3$ causes

Table III. Influence of $N$ on iterations of PALS and PARAFAC for a randomly simulated data array of three components with $a_{noise} = 0{\cdot}02$ (for each $N$ value, 20 randomly initialized runs were performed)

| | | Iterations | | Time per iteration (s) | |
|---|---|---|---|---|---|
| $N$ | IND | PALS | PARAFAC | PALS | PARAFAC |
| 3 | Max | 176 | 168 | | |
| | Min | 50 | 86 | 4·27 | 1·39 |
| | Ave | 69 | 127 | | |
| | Var | $7{\cdot}1\times10^{2}$ | $6{\cdot}8\times10^{2}$ | | |
| 4 | Max | 675 | 2019 | | |
| | Min | 129 | 120 | 4·75 | 1·54 |
| | Ave | 244 | 694 | | |
| | Var | $1{\cdot}7\times10^{4}$ | $3{\cdot}2\times10^{5}$ | | |
| 5 | Max | 737 | 2955 | 5·23 | 1·66 |
| | Min | 190 | 62 | | |
| | Ave | 411 | 1091 | | |
| | Var | $3{\cdot}0\times10^{4}$ | $7{\cdot}1\times10^{5}$ | | |
| 6 | Max | 1012 | 4954 | | |
| | Min | 127 | 702 | 5·61 | 1·70 |
| | Ave | 519 | 2107 | | |
| | Var | $4{\cdot}9\times10^{4}$ | $1{\cdot}4\times10^{6}$ | | |

 

Table IV. Influence of $\lambda$ on final results and iterations of PALS ($N = 5$) for a randomly simulated data array of three components with $a_{noise} = 0{\cdot}02$ (for each $\lambda$ value, 10 randomly initialized runs were performed)

| | | Mode 1 | | | Mode 2 | | | Mode 3 | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $\lambda$ | IND | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 | IT[a] |
| 0·001 | Max | 0·9999 | 0·9999 | 0·9999 | 0·9999 | 0·9999 | 0·9999 | 0·9999 | 1·0000 | 1·0000 | 3000 |
| | Min | 0·8480 | 0·6039 | 0·9108 | 0·2925 | 0·5705 | 0·3517 | 0·8043 | 0·8102 | 0·8750 | 162 |
| | Ave | 0·9773 | 0·9600 | 0·9854 | 0·8938 | 0·9502 | 0·8948 | 0·9768 | 0·9756 | 0·9784 | 1549 |
| | Var | $2{\cdot}3 \times 10^{-3}$ | $1{\cdot}6 \times 10^{-2}$ | $9{\cdot}8 \times 10^{-4}$ | $5{\cdot}0 \times 10^{-2}$ | $1{\cdot}8 \times 10^{-2}$ | $4{\cdot}8 \times 10^{-2}$ | $3{\cdot}7 \times 10^{-3}$ | $3{\cdot}6 \times 10^{-3}$ | $1{\cdot}7 \times 10^{-3}$ | $1{\cdot}6 \times 10^{6}$ |
| 0·1 | Max | 0·9998 | 0·9999 | 0·9999 | 0·9999 | 0·9998 | 0·9999 | 0·9999 | 1·0000 | 1·0000 | 3000 |
| | Min | 0·9892 | 0·9995 | 0·9974 | 0·9896 | 0·99986 | 0·9965 | 0·9981 | 0·9999 | 0·9995 | 464 |
| | Ave | 0·9987 | 0·9999 | 0·9995 | 0·9988 | 0·9997 | 0·9995 | 0·9997 | 1·0000 | 0·9999 | 1855 |
| | Var | $1{\cdot}1 \times 10^{-5}$ | $1{\cdot}6 \times 10^{-5}$ | $6{\cdot}1 \times 10^{-7}$ | $1{\cdot}1 \times 10^{-5}$ | $1{\cdot}4 \times 10^{-7}$ | $1{\cdot}1 \times 10^{-4}$ | $3{\cdot}2 \times 10^{-7}$ | $1{\cdot}0 \times 10^{-9}$ | $2{\cdot}5 \times 10^{-8}$ | $1{\cdot}0 \times 10^{6}$ |
| 1 | Max | 0·9997 | 0·9999 | 0·9998 | 0·9999 | 0·9998 | 0·9998 | 0·9999 | 1·0000 | 1·0000 | 1009 |
| | Min | 0·9959 | 0·9996 | 0·9998 | 0·9919 | 0·9993 | 0·9998 | 0·9993 | 0·9999 | 1·0000 | 162 |
| | Ave | 0·9993 | 0·9999 | 0·9998 | 0·9990 | 0·9997 | 0·9998 | 0·9998 | 1·0000 | 1·0000 | 327 |
| | Var | $1{\cdot}4 \times 10^{-6}$ | $9{\cdot}0 \times 10^{-9}$ | $5{\cdot}5 \times 10^{-32}$ | $6{\cdot}3 \times 10^{-4}$ | $2{\cdot}4 \times 10^{-8}$ | $5{\cdot}5 \times 10^{-32}$ | $3{\cdot}6 \times 10^{-8}$ | $1{\cdot}0 \times 10^{-9}$ | 0 | $6{\cdot}3 \times 10^{4}$ |
| 10 | Max | 0·9996 | 0·9999 | 0·9998 | 0·9998 | 0·9998 | 0·9998 | 0·9999 | 0·9999 | 1·0000 | 365 |
| | Min | 0·9995 | 0·9997 | 0·9998 | 0·9996 | 0·9995 | 0·9997 | 0·9998 | 0·9999 | 1·0000 | 86 |
| | Ave | 0·9996 | 0·9998 | 0·9998 | 0·9998 | 0·9997 | 0·9998 | 0·9999 | 0·9999 | 1·0000 | 209 |
| | Var | $1{\cdot}0 \times 10^{-9}$ | $2{\cdot}2 \times 10^{-9}$ | $5{\cdot}5 \times 10^{-32}$ | $5{\cdot}0 \times 10^{-9}$ | $5{\cdot}4 \times 10^{-9}$ | $2{\cdot}3 \times 10^{-9}$ | $1{\cdot}0 \times 10^{-9}$ | 0 | 0 | $7{\cdot}0 \times 10^{3}$ |
| 100 | Max | 0·9996 | 0·9998 | 0·9998 | 0·9997 | 0·9997 | 0·9997 | 0·9999 | 0·9999 | 1·0000 | 533 |
| | Min | 0·9995 | 0·9998 | 0·9998 | 0·9997 | 0·9996 | 0·9997 | 0·9999 | 0·9999 | 1·0000 | 87 |
| | Ave | 0·9996 | 0·9998 | 0·9998 | 0·9997 | 0·9997 | 0·9997 | 0·9999 | 0·9999 | 1·0000 | 190 |
| | Var | $2{\cdot}8 \times 10^{-9}$ | $5{\cdot}5 \times 10^{-32}$ | $5{\cdot}5 \times 10^{-32}$ | 0 | $2{\cdot}8 \times 10^{-9}$ | 0 | 0 | 0 | 0 | $1{\cdot}9 \times 10^{4}$ |
| $1 \times 10^{6}$ | Max | 0·9996 | 0·9998 | 0·9998 | 0·9997 | 0·9997 | 0·9997 | 0·9999 | 0·9999 | 1·0000 | 179 |
| | Min | 0·9995 | 0·9998 | 0·9998 | 0·9996 | 0·9996 | 0·9997 | 0·9998 | 0·9999 | 1·0000 | 73 |
| | Ave | 0·9996 | 0·9998 | 0·9998 | 0·9997 | 0·9997 | 0·9997 | 0·9999 | 0·9999 | 1·0000 | 111 |
| | Var | $2{\cdot}8 \times 10^{-9}$ | $5{\cdot}5 \times 10^{-32}$ | $5{\cdot}5 \times 10^{-32}$ | $1{\cdot}0 \times 10^{-9}$ | $2{\cdot}8 \times 10^{-9}$ | 0 | $1{\cdot}0 \times 10^{-9}$ | 0 | 0 | $9{\cdot}9 \times 10^{4}$ |

[a] IT denotes the iterations required by PALS.

Table V. Influence of $N$ on final results of PARAFAC for real data array (for each $N$ value, five randomly initialized runs were performed)

| | | Excitation profiles | | | Emission profiles | | | Concentration profiles | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | IND | 1[a] | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |
| 3 | Max | 0·9997 | 0·9996 | 0·9845 | 0·9994 | 0·9990 | 0·9936 | 0·9986 | 0·9969 | 0·9953 |
| | Min | 0·9997 | 0·9996 | 0·9844 | 0·9994 | 0·9990 | 0·9936 | 0·9986 | 0·9968 | 0·9953 |
| | Ave | 0·9997 | 0·9996 | 0·9844 | 0·9994 | 0·9990 | 0·9936 | 0·9986 | 0·9968 | 0·9953 |
| | Var | 0 | 0 | $3{\cdot}0\times10^{-9}$ | 0 | 0 | 0 | 0 | $3{\cdot}0\times10^{-9}$ | 0 |
| 4 | Max | 1·0000 | 0·9996 | 0·9896 | 0·9997 | 0·9995 | 0·9977 | 0·9989 | 0·9970 | 0·9982 |
| | Min | 0·8584 | 0·9980 | 0·8631 | 0·6006 | 0·9990 | 0·9897 | 0·8575 | 0·9967 | 0·7072 |
| | Ave | 0·9716 | 0·9988 | 0·9604 | 0·9178 | 0·9992 | 0·9952 | 0·9704 | 0·9968 | 0·9143 |
| | Var | $4{\cdot}0\times10^{-3}$ | $6{\cdot}2\times10^{-7}$ | $3{\cdot}0\times10^{-3}$ | $3{\cdot}2\times10^{-2}$ | $6{\cdot}7\times10^{-3}$ | $1{\cdot}3\times10^{-5}$ | $4{\cdot}0\times10^{-3}$ | $2{\cdot}7\times10^{-8}$ | $1{\cdot}6\times10^{-2}$ |
| 5 | Max | 0·9999 | 0·9980 | 0·9711 | 0·9995 | 0·9995 | 0·9980 | 0·9995 | 0·9980 | 0·9883 |
| | Min | 0·9963 | 0·9948 | 0·6869 | 0·8841 | 0·9990 | 0·9939 | 0·9859 | 0·9913 | 0·9845 |
| | Ave | 0·9981 | 0·9962 | 0·8226 | 0·9515 | 0·9993 | 0·9963 | 0·9925 | 0·9940 | 0·9865 |
| | Var | $3{\cdot}0\times10^{-3}$ | $2{\cdot}6\times10^{-6}$ | $1{\cdot}9\times10^{-2}$ | $2{\cdot}9\times10^{-3}$ | $7{\cdot}5\times10^{-8}$ | $2{\cdot}8\times10^{-6}$ | $4{\cdot}4\times10^{-5}$ | $1{\cdot}1\times10^{-5}$ | $1{\cdot}9\times10^{-6}$ |
| 6 | Max | 0·9997 | 0·9981 | 0·9904 | 0·9995 | 0·9968 | 0·9988 | 0·9988 | 0·9983 | 0·9986 |
| | Min | 0·9239 | 0·9710 | 0·1024 | 0·9013 | 0·1048 | 0·9851 | 0·8572 | 0·9315 | 0·9878 |
| | Ave | 0·9797 | 0·9923 | 0·5134 | 0·9665 | 0·7815 | 0·9921 | 0·9659 | 0·9846 | 0·9941 |
| | Var | $1{\cdot}1\times10^{-3}$ | $1{\cdot}4\times10^{-4}$ | $1{\cdot}9\times10^{-2}$ | $1{\cdot}6\times10^{-3}$ | $1{\cdot}5\times10^{-1}$ | $2{\cdot}9\times10^{-5}$ | $3{\cdot}8\times10^{-3}$ | $8{\cdot}8\times10^{-4}$ | $2{\cdot}3\times10^{-5}$ |

[a] Numbers 1, 2 and 3 represent tyrosine, tryptophan and phenylalanine respectively.

large variations in the final results in different runs, which will prevent experimenters from drawing correct conclusions. In comparison, for each $N$ value, all five runs of PALS have reached the same results, and no variations among runs were observed. It was observed that the quality of the results obtained by both PARAFAC and PALS for phenylalanine is obviously inferior to that for the other two components. Its emission and excitation profiles (Figure 4) resolved by PARAFAC ($N = 3$) show

Table VI. Influence of $N$ on final results of PALS with $\lambda = 1$ for real data array (for each $N$ value, five randomly initialized runs were performed)

| | | Excitation profiles | | | Emission profiles | | | Concentration profiles | | |
|---|---|---|---|---|---|---|---|---|---|---|
| $N$ | IND | 1 | 2 | 3 | 1 | 2 | 3 | 1 | 2 | 3 |
| 3 | Max | 0·9993 | 0·9967 | 0·9839 | 0·9993 | 0·9989 | 0·9938 | 0·9966 | 0·9878 | 0·9505 |
| | Min | 0·9993 | 0·9967 | 0·9839 | 0·9993 | 0·9989 | 0·9938 | 0·9966 | 0·9878 | 0·9505 |
| | Ave | 0·9993 | 0·9967 | 0·9839 | 0·9993 | 0·9989 | 0·9938 | 0·9966 | 0·9878 | 0·9505 |
| | Var | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 4 | Max | 0·9997 | 0·9987 | 0·9861 | 0·9993 | 0·9987 | 0·9985 | 0·9977 | 0·9869 | 0·9509 |
| | Min | 0·9997 | 0·9987 | 0·9861 | 0·9993 | 0·9987 | 0·9985 | 0·9977 | 0·9869 | 0·9509 |
| | Ave | 0·9997 | 0·9987 | 0·9861 | 0·9993 | 0·9987 | 0·9985 | 0·9977 | 0·9869 | 0·9509 |
| | Var | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 5 | Max | 0·9998 | 0·9986 | 0·9839 | 0·9992 | 0·9989 | 0·9992 | 0·9978 | 0·9864 | 0·9543 |
| | Min | 0·9998 | 0·9986 | 0·9839 | 0·9992 | 0·9989 | 0·9992 | 0·9978 | 0·9864 | 0·9543 |
| | Ave | 0·9998 | 0·9986 | 0·9839 | 0·9992 | 0·9989 | 0·9992 | 0·9978 | 0·9864 | 0·9543 |
| | Var | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |
| 6 | Max | 0·9997 | 0·9984 | 0·9379 | 0·9988 | 0·9970 | 0·9976 | 0·9982 | 0·9865 | 0·9606 |
| | Min | 0·9997 | 0·9984 | 0·9379 | 0·9988 | 0·9970 | 0·9976 | 0·9982 | 0·9865 | 0·9606 |
| | Ave | 0·9997 | 0·9984 | 0·9379 | 0·9988 | 0·9970 | 0·9976 | 0·9982 | 0·9865 | 0·9606 |
| | Var | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

 

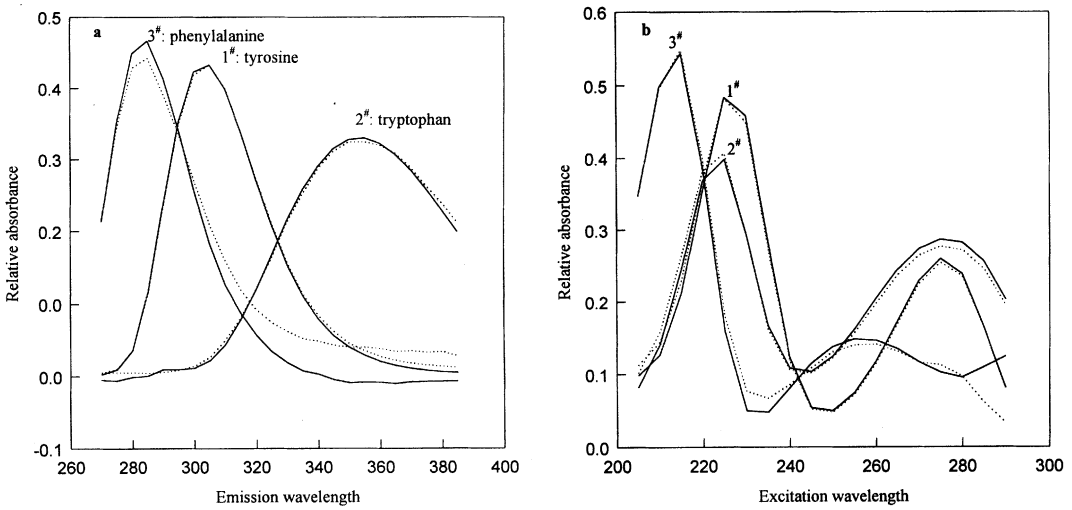


Figure 4. Expected (dotted line) and resolved (full line) emission (a) and excitation (b) profiles by PARAFAC with $N = 3$ for three components in real system.

some aberrations from the expected ones. Particularly in the range 280–290 nm the excitation profile of phenylalanine exhibits an increasing trend opposite to that of the expected profile. Although in the excitation profiles resolved by PALS with $N = 5$ and $\lambda = 1$ (Figure 5b) such an increasing trend is not so prominent, the quality of the corresponding emission profile (Figure 5a) and concentration profile (Table VI) still reveals the existence of some abnormal response of phenylalanine in this system. Therefore it may be more reasonable to evaluate the performance of PALS with respect to $N$ based on the results for tyrosine and tryptophan. As $N$ increases from 3 to 6, the relative coefficients of the loading matrices in the three modes for both tyrosine and tryptophan show no significant

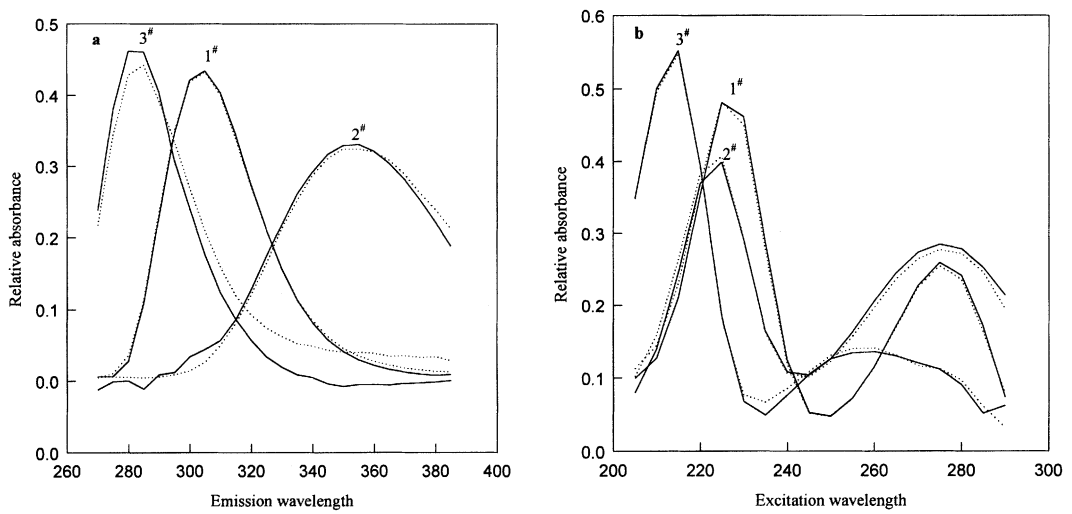


Figure 5. Expected (dotted line) and resolved (full line) emission (a) and excitation (b) profiles by PALS with $N = 5$ and $\lambda = 1$ for three components in real system.

 

deterioration. For instance, the relative coefficient between the real and resolved excitation profiles of tyrosine remains almost unchanged (varying in the range 0·9993–0·9998) as $N$ increases from 3 to 6. The agreement among the results for tyrosine and tryptophan obtained by PALS with different $N$ values further demonstrates its capability of being resistant to the influence of $N$.

All the programs used in this paper are written in Matlab 5·1 and run on a personal computer with a Pentium II processor.

## 7. CONCLUSIONS

The problem of PARAFAC requiring an accurate estimation of the number of factors in the system under study has been solved by the PALS algorithm designed in this paper. PALS is a unique algorithm which tries to alternately optimize three different objective functions to obtain the solutions for the trilinear mode. It has been proved that the result obtained by PALS with $N = F$ is unique regardless of some scaling and permutation indeterminacy. When $N > F$, the loading matrices obtained by PALS should also contain the corresponding actual profiles of the underlying factors in the three modes. Although the above conclusion was attained conditionally on the absence of noise, it was also confirmed by simulated and real data arrays. Owing to the multi-objective function property of PALS, a straightforward analysis of its convergence properties is difficult and hence simulations are resorted to. Simulations have revealed that although the optimization procedure of PALS may not decrease monotonically as in PARAFAC, it can often converge to satisfactory results within a reasonable computation time. The parameter $\lambda$ has some influence on the performance of PALS. For satisfactory results and a reasonable computational burden a moderate $\lambda$ is preferred in practice.

Although PALS has many advantages, it should be regarded as only one of the alternatives for trilinear decomposition. Owing to the possible bias introduced by the parameter $\lambda$ in PALS, it is recommended to use PALS only when the actual number of underlying factors is not available. If one has confidence in the estimation of the underlying factors, PARAFAC may be the first choice because of its statistical merits.

ACKNOWLEDGEMENTS

The authors would like to thank the National Natural Science Foundation of China for financial support (grant 29735150).

## APPENDIX I. NOMENCLATURE

| | |
|---|---|
| $x_{ijk}$ | the $ijk$th element of the three-way array $\underline{\mathbf{X}}$ |
| $e_{ijk}$ | the $ijk$th element of the three-way residue array $\underline{\mathbf{E}}$ |
| $\mathbf{A}_{I \times F}$, $\mathbf{B}_{I \times F}$, $\mathbf{C}_{K \times F}$ | the three underlying loading matrices of $\underline{\mathbf{X}}$ with dimensions $I \times F$, $J \times F$ and $K \times F$ respectively (for simplicity, in this paper, matrices $\mathbf{A}_{I \times F}$, $\mathbf{B}_{J \times F}$ and $\mathbf{C}_{K \times F}$ are represented by matrices **A, B** and **C** respectively) |
| $a_{if}$, $b_{jf}$, $c_{kf}$ | the $if$th, $jf$th and $kf$th elements of the three underlying loading matrices **A, B** and **C** respectively |
| $I$, $J$, $K$ | the dimensions of different modes in three-way arrays |
| $F$ | the number of underlying factors, i.e. the total number of detectable species, including the components of interest and interference(s) as well as background |
| $N$ | the number of factors used in calculation |

 

| | |
|---|---|
| $\hat{\mathbf{A}}_{I\times N}$, $\hat{\mathbf{B}}_{J\times N}$, $\hat{\mathbf{C}}_{K\times N}$ | the three resolved loading matrices of $\underline{\mathbf{X}}$ with dimensions $I\times N$, $J\times N$ and $K\times N$ respectively (in this paper, $\hat{\mathbf{A}}_{I\times N}$, $\hat{\mathbf{B}}_{J\times N}$ and $\hat{\mathbf{C}}_{K\times N}$ are often simply written as $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}$ and $\hat{\mathbf{C}}$ respectively) |
| $(\hat{\mathbf{A}})^{+}$ | the Moore–Penrose generalized inverse of matrix $\hat{\mathbf{A}}$ |
| $\mathbf{X}_{..k}=\mathbf{A}\mathrm{diag}(\mathbf{c}_k^{\mathrm{T}})\mathbf{B}^{\mathrm{T}}+\mathbf{E}_{..k}$ | the $k$th frontal slice of the three-way array $\underline{\mathbf{X}}$ |
| $\mathbf{X}_{.j.}=\mathbf{C}\mathrm{diag}(\mathbf{b}_j^{\mathrm{T}})\mathbf{A}^{\mathrm{T}}+\mathbf{E}_{.j.}$ | the $j$th lateral slice of the three-way array $\underline{\mathbf{X}}$ |
| $\mathbf{X}_{i..}=\mathbf{B}\mathrm{diag}(\mathbf{a}_i^{\mathrm{T}})\mathbf{C}^{\mathrm{T}}+\mathbf{E}_{i..}$ | the $i$th horizontal slice of the three-way array $\underline{\mathbf{X}}$ |
| $\mathbf{E}_{..k}$, $\mathbf{E}_{.j.}$, $\mathbf{E}_{i..}$ | the $k$th frontal, $j$th lateral and $i$th horizontal slices of the three-way residue array $\underline{\mathbf{E}}$ respectively |
| $\mathrm{diag}(\hat{\mathbf{c}}_k{}^{\mathrm{T}})$, $\mathrm{diag}(\hat{\mathbf{a}}_i{}^{\mathrm{T}})$, $\mathrm{diag}\hat{\mathbf{b}}(_j{}^{\mathrm{T}})$ | diagonal matrices with elements equal to the $k$th, $i$th and $j$th rows of **C,** **A** and **B** respectively |
| $\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})$, $\mathrm{diag}((\hat{\mathbf{a}}_i)^{\mathrm{T}})$, $\mathrm{diag}((\hat{\mathbf{b}}_j)^{\mathrm{T}})$ | diagonal matrices with elements equal to the $k$th, $i$th and $j$th rows of $\hat{\mathbf{C}}$, $\hat{\mathbf{A}}$ and $\hat{\mathbf{B}}$ respectively |

## APPENDIX II

*II. 1. The uniqueness of the solutions obtained by PALS when N = F*

*II.1.1. The existence of solutions $\hat{\mathbf{A}}_{I\times F}$, $\hat{\mathbf{B}}_{J\times F}$ and $\hat{\mathbf{C}}_{K\times F}$ satisfying $S(\hat{\mathbf{A}})=0$, $S(\hat{\mathbf{B}})=0$ and $S(\hat{\mathbf{C}})=0$ under the conditions $\underline{\mathbf{E}}=0$ and rank(A) = rank(B) = F.*

Suppose the actual loading matrices of the PARAFAC model are **A, B** and **C**, i.e.

$$\mathbf{X}_{..k}=\mathbf{A}\mathrm{diag}(\mathbf{c}_k^{\mathrm{T}})\mathbf{B}^{\mathrm{T}},\quad k=1,\ldots,K$$

Then the following two equations hold:

$$(\mathbf{A})^{+}\mathbf{X}_{..k}=(\mathbf{A})^{+}\mathbf{A}\mathrm{diag}(\mathbf{c}_k^{\mathrm{T}})\mathbf{B}^{\mathrm{T}}=\mathrm{diag}(\mathbf{c}_k^{\mathrm{T}})\mathbf{B}^{\mathrm{T}},\quad k=1,2,\ldots,K$$

$$\mathbf{X}_{..k}(\mathbf{B}^{\mathrm{T}})^{+}=\mathbf{A}\mathrm{diag}(\mathbf{c}_k^{\mathrm{T}})\mathbf{B}^{\mathrm{T}}(\mathbf{B}^{\mathrm{T}})^{+}=\mathbf{A}\mathrm{diag}(\mathbf{c}_k^{\mathrm{T}}),\quad k=1,2,\ldots,K$$

Therefore there exist at least loading matrices $\hat{\mathbf{A}}_{I\times F}=\mathbf{A}$, $\hat{\mathbf{B}}_{J\times F}=\mathbf{B}$ and $\hat{\mathbf{C}}_{K\times F}=\mathbf{C}$ which satisfy $S(\hat{\mathbf{A}})=0$, $S(\hat{\mathbf{B}})=0$ and $S(\hat{\mathbf{C}})=0$.

*II.1.2. Provided that rank(**A**) = rank(**B**) = F and $\underline{\mathbf{E}}=0$, solutions $\hat{\mathbf{A}}_{I\times F}$, $\hat{\mathbf{B}}_{J\times F}$ and $\hat{\mathbf{C}}_{K\times F}$ which satisfy the equations $S(\hat{\mathbf{A}})=0$, $S(\hat{\mathbf{B}})=0$ and $S(\hat{\mathbf{C}})=0$ should be unique up to some scaling and permutation freedom.*

It has been proved by Kruskal that under some mild conditions [19,20] the trilinear decomposition model $\mathbf{X}_{..k}=\mathbf{A}_{I\times F}\mathrm{diag}(\mathbf{c}_k^{\mathrm{T}})\mathbf{B}_{J\times F}^{\mathrm{T}}$, $k=1,\ldots,K$, in itself is unique up to some scaling and permutation freedom. Therefore it can be easily verified that the solutions satisfying $S(\hat{\mathbf{A}})=0$, $S(\hat{\mathbf{B}})=0$ and $S(\hat{\mathbf{C}})=0$ should also be unique in the same sense as in PARAFAC when rank(**A**) = rank(**B**) = $F$ and $\underline{\mathbf{E}}=0$.

*II.2. The property of PALS of insensitivity to excess factors used in calculation (N > F)*

*II.2.1. The existence of solutions $\hat{\mathbf{A}}_{I\times N}$, $\hat{\mathbf{B}}_{J\times N}$ and $\hat{\mathbf{C}}_{K\times N}$ satisfying $S(\hat{\mathbf{A}})=0$, $S(\hat{\mathbf{B}})=0$ and $S(\hat{\mathbf{C}})=0$ when $\underline{\mathbf{E}}=0$ and rank(**A**) = rank(**B**) = F.*

For simplicity of description it is assumed that $F=4$ and $N=6$ (note that the values of $F$ and $N$ are set randomly; they will not affect the generality of the following conclusions).

 

$$\mathbf{A}_{I\times F}=(\mathbf{a}_1,\mathbf{a}_2,\mathbf{a}_3,\mathbf{a}_4)\qquad \mathbf{B}_{J\times F}=(\mathbf{b}_1,\mathbf{b}_2,\mathbf{b}_3,\mathbf{b}_4)\qquad \mathbf{C}_{K\times F}=(\mathbf{c}_1,\mathbf{c}_2,\mathbf{c}_3,\mathbf{c}_4)$$

are the underlying loading matrices satisfying $\mathbf{X}_{..k}=\mathbf{A}_{I\times F}\mathrm{diag}(\mathbf{c}_k^{\mathrm{T}})\mathbf{B}_{J\times F}^{\mathrm{T}},\ k=1,\ldots,K$.

Define

$$\hat{\mathbf{A}}=(\mathbf{A}_1,\mathbf{A}_2),\qquad \mathbf{A}_1=\mathbf{A}_{I\times F},\qquad \mathbf{A}_2=(\mathbf{a}_1,\mathbf{a}_2)$$

$$\hat{\mathbf{B}}=(\mathbf{B}_1,\mathbf{B}_2),\qquad \mathbf{B}_1=\mathbf{B}_{J\times F},\qquad \mathbf{B}_2=(\mathbf{b}_1,\mathbf{b}_2)$$

$$\hat{\mathbf{C}}=\left(\tfrac{1}{2}\mathbf{c}_1,\tfrac{1}{2}\mathbf{c}_2,\mathbf{c}_3,\mathbf{c}_4,\tfrac{1}{2}\mathbf{c}_1,\tfrac{1}{2}\mathbf{c}_2\right)$$

According to Reference [22], one has

$$(\hat{\mathbf{A}})^{+}=\begin{pmatrix}\mathbf{A}_1^{+}-\mathbf{A}_1^{+}\mathbf{A}_2\mathbf{K}_{\mathbf{A}}^{-1}\mathbf{A}_2^{\mathrm{T}}(\mathbf{A}_1^{+})^{\mathrm{T}}\mathbf{A}_1^{+}\\ \mathbf{K}_{\mathbf{A}}^{-1}\mathbf{A}_2^{\mathrm{T}}(\mathbf{A}_1^{+})^{\mathrm{T}}\mathbf{A}_1^{+}\end{pmatrix},\qquad (\hat{\mathbf{B}})^{+}=\begin{pmatrix}\mathbf{B}_1^{+}-\mathbf{B}_1^{+}\mathbf{B}_2\mathbf{K}_{\mathbf{B}}^{-1}\mathbf{B}_2^{\mathrm{T}}(\mathbf{B}_1^{+})^{\mathrm{T}}\mathbf{B}_1^{+}\\ \mathbf{K}_{\mathbf{B}}^{-1}\mathbf{B}_2^{\mathrm{T}}(\mathbf{B}_1^{+})^{\mathrm{T}}\mathbf{B}_1^{+}\end{pmatrix}$$

$$\mathbf{K}_{\mathbf{A}}=\mathbf{I}+\mathbf{A}_2^{\mathrm{T}}(\mathbf{A}_1^{+})^{\mathrm{T}}\mathbf{A}_1^{+}\mathbf{A}_2,\qquad \mathbf{K}_{\mathbf{B}}=\mathbf{I}+\mathbf{B}_2^{\mathrm{T}}(\mathbf{B}_1^{+})^{\mathrm{T}}\mathbf{B}_1^{+}\mathbf{B}_2$$

Suppose

$$\mathbf{A}_1{}^{+}=\begin{pmatrix}\mathbf{p}\mathbf{a}_1^{\mathrm{T}}\\ \mathbf{p}\mathbf{a}_2^{\mathrm{T}}\\ \mathbf{p}\mathbf{a}_3^{\mathrm{T}}\\ \mathbf{p}\mathbf{a}_4^{\mathrm{T}}\end{pmatrix},\qquad \mathbf{B}_1{}^{+}=\begin{pmatrix}\mathbf{p}\mathbf{b}_1^{\mathrm{T}}\\ \mathbf{p}\mathbf{b}_2^{\mathrm{T}}\\ \mathbf{p}\mathbf{b}_3^{\mathrm{T}}\\ \mathbf{p}\mathbf{b}_4^{\mathrm{T}}\end{pmatrix}$$

Then

$$(\hat{\mathbf{A}})^{+}=\begin{pmatrix}\frac{1}{2}\mathbf{p}\mathbf{a}_1^{\mathrm{T}}\\ \frac{1}{2}\mathbf{p}\mathbf{a}_2^{\mathrm{T}}\\ \mathbf{p}\mathbf{a}_3^{\mathrm{T}}\\ \mathbf{p}\mathbf{a}_4^{\mathrm{T}}\\ \frac{1}{2}\mathbf{p}\mathbf{a}_1^{\mathrm{T}}\\ \frac{1}{2}\mathbf{p}\mathbf{a}_2^{\mathrm{T}}\end{pmatrix},\qquad (\hat{\mathbf{B}})^{+}=\begin{pmatrix}\frac{1}{2}\mathbf{p}\mathbf{b}_1^{\mathrm{T}}\\ \frac{1}{2}\mathbf{p}\mathbf{b}_2^{\mathrm{T}}\\ \mathbf{p}\mathbf{b}_3^{\mathrm{T}}\\ \mathbf{p}\mathbf{b}_4^{\mathrm{T}}\\ \frac{1}{2}\mathbf{p}\mathbf{b}_1^{\mathrm{T}}\\ \frac{1}{2}\mathbf{p}\mathbf{b}_2^{\mathrm{T}}\end{pmatrix}$$

from which it follows that

$$\mathbf{X}_{..k}=\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}},\quad k=1,\ldots,K$$

$$(\hat{\mathbf{A}})^{+}\mathbf{X}_{..k}=\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}},\quad k=1,\ldots K$$

$$\mathbf{X}_{..k}((\hat{\mathbf{B}})^{\mathrm{T}})^{+}=\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}}),\quad k=1,\ldots,K$$

 

Therefore

$$\sum_{k=1}^{K}(2\left\|\mathbf{X}_{..k}-\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2}+\lambda(\left\|\mathbf{X}_{..k}((\hat{\mathbf{B}})^{\mathrm{T}})^{+}-\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})\right\|_{\mathrm{F}}^{2}$$

$$+\left\|(\hat{\mathbf{A}})^{+}\mathbf{X}_{..k}-\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2}))=0$$

$$\sum_{k=1}^{K}(\left\|\mathbf{X}_{..k}-\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2}+\lambda\left\|\mathbf{X}_{..k}((\hat{\mathbf{B}})^{\mathrm{T}})^{+}-\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})\right\|_{\mathrm{F}}^{2})=0$$

$$\sum_{k=1}^{K}(\left\|\mathbf{X}_{..k}-\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2}+\lambda\left\|\hat{\mathbf{A}}^{+}\mathbf{X}_{..k}-\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}}\right\|_{\mathrm{F}}^{2})=0$$

*II.2.2. Conditional on rank(***A***) = rank(***B***) = F and* $\underline{\mathbf{E}}$ *= 0, the columns of the solutions* $\hat{\mathbf{A}}_{I\times N}$, $\hat{\mathbf{B}}_{J\times N}$ *and* $\hat{\mathbf{C}}_{K\times N}$ *which satisfy S(*$\hat{\mathbf{A}}$*) = 0, S(*$\hat{\mathbf{B}}$*) = 0 and S(*$\hat{\mathbf{C}}$*) = 0 should be the actual profiles of underlying factors in corresponding modes, otherwise zero.*

Since $S(\hat{\mathbf{A}})=0$, $S(\hat{\mathbf{B}})=0$ and $S(\hat{\mathbf{C}})=0$, one has the following equations:

$$\mathbf{X}_{..k}=\hat{\mathbf{A}}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}},\quad k=1,\ldots,K$$

$$(\hat{\mathbf{A}})^{+}\mathbf{X}_{..k}=\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}},\quad k=1,\ldots,K$$

Therefore

$$(\hat{\mathbf{A}})^{+}\hat{\mathbf{A}}\mathrm{diag}(\hat{\mathbf{c}}_k^{\mathrm{T}})\mathbf{B}^{\mathrm{T}}=\mathrm{diag}(\hat{\mathbf{c}}_k)^{\mathrm{T}}(\hat{\mathbf{B}})^{\mathrm{T}},\quad k=1,2,\ldots,K$$

Suppose

$$\hat{\mathbf{A}}=(\mathbf{A}_F,\mathbf{A}_{N-F}),\qquad \hat{\mathbf{B}}=(\mathbf{B}_F,\mathbf{B}_{N-F}),\qquad \mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})=\begin{pmatrix}\mathrm{diag}(\mathbf{c}_{k,F}^{\mathrm{T}}) & 0\\ 0 & \mathrm{diag}(\mathbf{c}_{k,N-F}^{\mathrm{T}})\end{pmatrix}$$

where $\mathbf{A}_F$ of size $I\times F$ is a column full-rank matrix. Since one has

$$(\hat{\mathbf{A}})^{+}=\begin{pmatrix}\mathbf{A}_F^{+}-\mathbf{A}_F^{+}\mathbf{A}_{N-F}\mathbf{K}^{-1}\mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}}\mathbf{A}_F^{+}\\ \mathbf{K}^{-1}\mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}}\mathbf{A}_F^{+}\end{pmatrix}$$

$$\mathbf{K}=\mathbf{I}+\mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}}\mathbf{A}_F^{+}\mathbf{A}_{N-F}$$

 

hence

$$(\hat{\mathbf{A}})^{+}\mathbf{A}\mathrm{diag}((\hat{\mathbf{c}}_k)^{\mathrm{T}})(\hat{\mathbf{B}})^{\mathrm{T}} = \begin{pmatrix} \mathbf{A}_F^{+} - \mathbf{A}_F^{+}\mathbf{A}_{N-F}\mathbf{K}_1^{-1}\mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}}\mathbf{A}_F^{+} \\ \mathbf{K}_1^{-1}\mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}}\mathbf{A}_F^{+} \end{pmatrix} (\mathbf{A}_F, \mathbf{A}_{N-F})$$

$$\times \begin{pmatrix} \mathrm{diag}(\mathbf{c}_{k,F}^{\mathrm{T}}) & 0 \\ 0 & \mathrm{diag}(\mathbf{c}_{k,N-F}^{\mathrm{T}}) \end{pmatrix} (\mathbf{B}_F, \mathbf{B}_{N-F})^{\mathrm{T}}$$

$$= \begin{pmatrix} \mathrm{diag}(\mathbf{c}_{k,F}^{\mathrm{T}}) & 0 \\ 0 & \mathrm{diag}(\mathbf{c}_{k,N-F}^{\mathrm{T}}) \end{pmatrix} (\mathbf{B}_F, \mathbf{B}_{N-F})^{\mathrm{T}}$$

which can be rewritten as the following equations:

$$(\mathbf{I} - \mathbf{A}_F^{+}\mathbf{A}_{N-F}\mathbf{K}_1^{-1}\mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}})\mathrm{diag}(\mathbf{c}_{k,F}^{\mathrm{T}})\mathbf{B}_F^{\mathrm{T}}$$

$$+ \mathbf{A}_F^{+}(\mathbf{I} - \mathbf{A}_{N-F}\mathbf{K}_1^{-1}\mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}})\mathbf{A}_{N-F}\mathrm{diag}(\mathbf{c}_{k,N-F}^{\mathrm{T}})\mathbf{B}_{N-F}^{\mathrm{T}} = \mathrm{diag}(\mathbf{c}_{k,F}^{\mathrm{T}})\mathbf{B}_F^{\mathrm{T}}$$

$$\mathbf{K}_1^{-1}\mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}}\mathrm{diag}(\mathbf{c}_{k,F}^{\mathrm{T}})\mathbf{B}_F^{\mathrm{T}} + \mathbf{K}_1^{-1}\mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}}\mathbf{A}_F^{+}\mathbf{A}_{N-F}\mathrm{diag}(\mathbf{c}_{k,N-F}^{\mathrm{T}})\mathbf{B}_{N-F}^{\mathrm{T}}$$

$$= \mathrm{diag}(\mathbf{c}_{k,N-F}^{\mathrm{T}})\mathbf{B}_{N-F}^{\mathrm{T}}$$

Thus we have

$$\mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}}\mathrm{diag}(\mathbf{c}_{k,F}^{\mathrm{T}})\mathbf{B}_F^{\mathrm{T}} + \mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}}\mathbf{A}_F^{+}\mathbf{A}_{N-F}\mathrm{diag}(\mathbf{c}_{k,N-F}^{\mathrm{T}})\mathbf{B}_{N-F}^{\mathrm{T}}$$

$$= \mathbf{K}_1\mathrm{diag}(\mathbf{c}_{k,N-F}^{\mathrm{T}})\mathbf{B}_{N-F}^{\mathrm{T}}$$

which is equivalent to

$$(\mathbf{K}_1 - \mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}}\mathbf{A}_F^{+}\mathbf{A}_{N-F})\mathrm{diag}(\mathbf{c}_{k,N-F}^{\mathrm{T}})\mathbf{B}_{N-F}^{\mathrm{T}} = \mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}}\mathrm{diag}(\mathbf{c}_{k,F}^{\mathrm{T}})\mathbf{B}_F^{\mathrm{T}}$$

Finally we get

$$\mathrm{diag}(\mathbf{c}_{k,N-F}^{\mathrm{T}})\mathbf{B}_{N-F}^{\mathrm{T}} = \mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}}\mathrm{diag}(\mathbf{c}_{k,F}^{\mathrm{T}})\mathbf{B}_F^{\mathrm{T}}$$

If $\mathbf{A}_{N-F} = 0$, it is obvious that

$$\mathbf{X}_{..k} = \mathbf{A}_F\mathrm{diag}(\mathbf{c}_{k,F}^{\mathrm{T}})\mathbf{B}_F^{\mathrm{T}}$$

If $\mathbf{A}_{N-F} \neq 0$, the following equation holds:

$$\mathbf{X}_{..k} = \mathbf{A}_F\mathrm{diag}(\mathbf{c}_{k,F}^{\mathrm{T}})\mathbf{B}_F^{\mathrm{T}} + \mathbf{A}_{N-F}\mathrm{diag}(\mathbf{c}_{k,N-F}^{\mathrm{T}})\mathbf{B}_{N-F}^{\mathrm{T}}$$

$$= (\mathbf{A}_F + \mathbf{A}_{N-F}\mathbf{A}_{N-F}^{\mathrm{T}}(\mathbf{A}_F^{+})^{\mathrm{T}})\mathrm{diag}(\mathbf{c}_{k,F}^{\mathrm{T}})\mathbf{B}_F^{\mathrm{T}}$$

According to the theorem of Kruskal [19,20], the trilinear model is unique up to some scaling and permutation indeterminacy:

$$\mathbf{X}_{..k} = \mathbf{A}_{I\times F}\mathrm{diag}(\mathbf{c}_{k,F}^{\mathrm{T}})\mathbf{B}_{J\times F}^{\mathrm{T}}$$

 

Assuming the columns of $\mathbf{B}_{I \times F}$ and $\mathbf{B}_F$ are arranged according to the same criterion, then

$$\mathbf{B}_F = \mathbf{B}_{I \times F}$$

where $\Lambda$ is a diagonal matrix.

This conclusion implies that loading matrix $\mathbf{B}_F$ is also a column full-rank matrix. Similarly, it can easily be demonstrated that the columns of loading matrices $\mathbf{A}_F$ and $\mathbf{C}_F$ should also be the corresponding profiles of the actual factors with physical meaning in the first and third modes respectively. For any column of $\mathbf{A}$, $\mathbf{a}_i \neq 0$, there exist $F - 1$ other columns $\mathbf{A}_{F-1}$ which satisfy rank($[\mathbf{A}_{F-1}, \mathbf{a}_i]$) = $F$. Thus the corresponding $\mathbf{b}_i$ should be the profile in the second mode of a certain underlying factor, which indicates that $\mathbf{a}_i$ and $\mathbf{c}_i$ are the respective profiles in the first and third modes of the certain underlying factor. In conclusion, the columns of the loading matrices $\hat{\mathbf{A}}_{I \times N}$, $\hat{\mathbf{B}}_{J \times N}$ and $\hat{\mathbf{C}}_{K \times N}$ should be the profiles in the corresponding modes of the underlying factors, otherwise zero.

## REFERENCES


1. Sanchez E, Kowalski BR. *J. Chemometrics* 1990; **4**: 29.
2. Li S, Hamilton JC, Gemperline PJ. *Anal. Chem.* 1992; **64**: 599.
3. Faber NM, Buydens LMC, Kateman G. *J. Chemometrics* 1994; **8**: 141.
4. Faber NM, Buydens LMC, Kateman G. *J. Chemometrics* 1994; **8**: 181.
5. Faber NM, Buydens LMC, Kateman G. *J. Chemometrics* 1994; **8**: 273.
6. Faber K, Lorber A, Kowalski BR. *J. Chemometrics* 1997; **11**: 95.
7. Windig W, Antalek B. *Chemometrics Intell. Lab. Syst.* 1997; **37**: 241.
8. Zeng Y, Hopke PK. *J. Chemometrics* 1992; **6**: 65.
9. Linder M, Sundberg R. *Chemometrics Intell. Lab. Syst.* 1998; **42**: 159.
10. Harshman RA. *UCLA Working Papers Phonet.* 1970; **16**: 1.
11. Carroll JD, Chang J. *Psychometrika* 1970; **35**: 283.
12. Kiers HAL, Krijnen WP. *Psychometrika* 1991; **56**: 147.
13. Leurgans SE, Ross RT, Abel RB. *SIAM J. Matrix Anal. Appl.* 1993; **14**: 1064.
14. Bro R. *Chemometrics Intell. Lab. Syst.* 1997; **38**: 149.
15. Wu HL, Shibukawa M, Oguma K. *J. Chemometrics* 1998; **12**: 1.
16. Alsberg BK, Kvalheim OM. *Chemometrics Intell. Lab. Syst.* 1994; **23**: 29.
17. Kiers HAL, Harshman RA. *Chemometrics Intell. Lab. Syst.* 1997; **36**: 31.
18. Kiers HAL. *J. Chemometrics* 1998; **12**: 155.
19. Kruskal JB. *Linear Algebra Appl.* 1977; **18**: 95.
20. Kruskal JB. *Psychometrika* 1976; **41**: 281.
21. Chen ZP, Wu HL, Yu RQ. *J. Chemometrics* 2001; **15** in press.
22. Cline RE. *J. Soc. Ind. Appl. Math.* 1964; **12**: 588.


 