Anal. Chem. **1999,** *71,* 4254−4262

Coupled Vectors Resolution Method for Chemometric Calibration with Three-Way Data

Jian-Hui Jiang, Hai-Long Wu, Zeng-Ping Chen, and Ru-Qin Yu*

College of Chemistry and Chemical Engineering, Hunan University, Changsha, 410082, P. R. China

A new second-order calibration procedure, the coupled vectors resolution (COVER) method, has been developed. The objective of the method is to seek a couple of vectors that minimize a least-squares criterion. With the knowledge indispensable for quantitation, the method yields direct solutions to various cases of second-order calibration. Moreover, it allows a statistically plausible way to make use of multisample information. In the case of multiple calibration samples, the method uses the calibration samples to resolve the profiles of the analytes in each order, and then calculates the concentrations of the analytes. This offers the advantage that unknown mixtures newly collected can be predicted in a direct manner. In the case of one calibration sample, the method provides an effective way to utilize the information of spectral profiles of the analytes. Results of simulated experiments and a real analytical example show that the proposed method produces acceptable performance in profile resolution and concentration estimation.

The advancement of second-order instrumentation, such as excitation−emission fluorescence, liquid chromatography with ultraviolet/visible detection (LC−UV/vis), and gas chromatography hyphenated with mass spectrometry (GC/MS), which produces two-way data for a single sample, has brought growing interest in the development of second-order calibration methodologies.[1−3] Tremendous potential can be derived from bilinear data in that the data enable one to quantify the components of analytical interest in the presence of unknown interferences not included in the calibration samples. This is known as the second-order advantage.[4] Mathematically, the second-order advantage makes the final goal of analytical chemistry achievable even without the aid of complicated preseparation procedures. Primarily, there are two types of methodologies approaching second-order calibration. Methods of the first type are built upon eigenanalysis or generalized eigenanalysis.[5−8] Prominent examples are the generalized rank annihilation method (GRAM) [5,9−11] as well as its extension, the trilinear decomposition (TLD) method.[12−19] Unfortunately, these methods need to construct two pseudosamples to formulate a eigenproblem, which unavoidably incurs a loss of information for multiple samples. Moreover, the algorithm for the methods is subjected to the danger of yielding imaginary solutions. Methods of the second type are established on an iterative trilinear decomposition of the data cube integrated by the calibration matrices and the data measured on the unknown samples.[20−29] The methods do provide a sensible way to make use of the whole data. However, it is reported that iterative algorithms are plagued by degenerate solutions and computational swamps, yielding chemically meaningless solutions.[25,27] Additionally, these methods by themselves are incapable of generalizing to new unknown samples in an immediate way. One has to redo an entire decomposition of the data cube augmented by the data matrices measured on the new samples.

In this paper, a *co*upled *vec*tors *r*esolution (COVER) method is developed. The method aims at seeking a couple of vectors, called coupled vectors, which minimize a least-squares criterion proposed by the present authors. With the prior knowledge of partial spectral or concentration profiles of the analytes, the COVER method can produce direct solutions to second-order calibration and provides a statistically plausible manner to make use of multisample information. A salient characteristic of the

* Corresponding author: (e-mail) rqyu@mail.hunu.edu.cn; (fax) 86-0731-8824525.

(1) Ho, C.-N.; Christian, G. D.; Davidson, E. R. *Anal. Chem*. **1978**, *50*, 1108−1113.
(2) Lorber, A. *Anal. Chim. Acta* **1984**, *164*, 293−297.
(3) Ohman, J.; Geladi, P.; Wold, S. *J. Chemom*. **1990**, *4*, 79−90.
(4) Booksh, K. S.; Kowalski, B. R. *Anal. Chem*. **1994**, *66*, 782A−791A.
(5) Sanchez, E.; Kowalski, B. R. *Anal. Chem*. **1986**, *58*, 499−501.
(6) Millican, D. W.; McGown, L. B. *Anal. Chem*. **1989**, *61*, 580−583.
(7) Millican, D. W.; McGown, L. B. *Anal. Chem*. **1990**, *62*, 2242−2247.
(8) Scarminio, I.; Kubista, M. *Anal. Chem*. **1993**, *65*, 409.
(9) Wilson, B.; Sanchez, E.; Kowalski, B. R. *J. Chemom*. **1989**, *3*, 493.
(10) Li, S,; Hamilton, C.; Gemperline, P. *Anal. Chem*. **1992**, *64*, 599−607.
(11) Faber, K.; Lorber, A.; Kowalski, B. R. *J. Chemom*. **1997**, *11*, 95−109.
(12) Sanchez, E.; Kowalski, B. R. *J. Chemom*. **1990**, *4*, 29−45.
(13) Li, S.; Gemperline P. *J. Chemom*. **1993**, *7*, 77.
(14) Lin, Z.; Booksh, K. S.; Burgess, L. W.; Kowalski, B. R. *Anal. Chem*. **1994**, *66*, 2552−2560.
(15) Booksh, K. S.; Lin, Z.; Wang, Z.; Kowalski, B. R. *Anal. Chem*. **1994**, *66*, 2561−2569.
(16) Henshaw, J. M.; Burgess, L. W.; Booksh, K. S.; Kowalski, B. R. *Anal. Chem*. **1994**, *66*, 3328−3336.
(17) Gui, M.; Rutan, S. C.; Agbodian, A. *Anal. Chem*. **1995**, *67*, 3293−3299.
(18) Saurina, J.; Hemandez-Cassou, S.; Tauler, R. *Anal. Chem*. **1997**, *69*, 2329−2336.
(19) Juan, A.; Rutan, S. C.; Tauler, R.; Massart, D. L. *Chemom. Intell. Lab. Syst*. **1998**, *40*, 19−32.
(20) Appellof, C. J.; Davidson, E. R. *Anal. Chem*. **1981**, *53*, 2053−2056.
(21) Geladi, P. *Chemom. Intell. Lab. Syst*. **1989**, *7*, 11−30.
(22) Smilde, A. K.; Doornbos, D. A. *J. Chemom*. **1991**, *5*, 345−360.
(23) Smilde, A. K.; Doornbos, D. A. *J. Chemom*. **1992**, *6*, 11−28.
(24) Paetero, P. *Chemom. Intell. Lab. Syst*. **1997**, *38*, 223−242.
(25) Mitchell, B. C.; Burdick, D. S. *J. Chemom*. **1994**, *8*, 155−168.
(26) Bro, R. *Chemom. Intell. Lab. Syst*. **1997**, *38*, 149−171.
(27) Rayens, W. S.; Mitchell, B. C. *Chemom. Intell. Lab. Syst*. **1997**, *38*, 173−181.
(28) Wu, H.-L.; Shibukawa, M.; Oguma, K. *J. Chemom*. **1998**, *12*, 1−26.
(29) Beltran, J. L.; Guiteras, J.; Ferrer, R. *Anal. Chem*. **1998**, *70*, 1949−1955.

10.1021/ac990009n CCC: $18.00
Published on Web 08/26/1999

method is that it deals with the calibration and prediction stages separately in the case where two or more calibration samples are available. This offers the advantage that unknown samples newly collected can be predicted in a direct manner. In the case of only one calibration sample available, the method provides an effective way to utilize the information of pure spectral profile.

THEORY

1. Least-Squares Criterion for Coupled Vectors Resolution. Trilinear instrumental responses, say the two-way matrices obtained using LC−UV/vis, for K mixture samples can be expressed as

$$\mathbf{R}_k = \sum_{n=1}^{N} c_{kn}\mathbf{x}_n\mathbf{y}_n^{\mathrm{T}} + \mathbf{E}_k \qquad k = 1, ..., K \tag{1}$$

where $\mathbf{R}_k$ is the response matrix for the kth sample (I wavelengths by J retention times). c_{kn} is the concentration of the nth component in the kth sample. $\mathbf{x}_n$ and $\mathbf{y}_n$ are the spectral and the chromatographic profiles, respectively, for the nth component. It is important to note that it is implied by eq 1 that $\mathbf{x}_n\,\mathbf{y}_n{}^{\mathrm{T}}$ is the two-way response of the pure nth component of unit concentration. To guarantee unique representations of $\mathbf{x}_n$ and $\mathbf{y}_n$, one can assume without loss of generality that $||\,\mathbf{x}_n\,|| = 1$ with $||\cdot||$ denoting the Euclidean norm of a vector. $\mathbf{E}_k$ is the measurement error matrix for the kth sample, and N is the number of components in the samples. It is known that the mth contravariant vector [12] $\mathbf{q}_m$ of the base vectors $\mathbf{y}_1, ..., \mathbf{y}_N$ satisfies

$$\mathbf{y}_n{}^{\mathrm{T}}\mathbf{q}_m = \delta_{mn} \qquad m, n = 1, ..., N \tag{2}$$

where $\delta_{mn} = 1$ if $m = n$, and $\delta_{mn} = 0$ if $m \neq n$. Therefore, one obtains

$$\mathbf{R}_k\mathbf{q}_n = c_{kn}\mathbf{x}_n + \mathbf{e}_{kn} \qquad k = 1, ..., K \tag{3}$$

where $\mathbf{e}_{kn}$ is a vector of errors. Because eq 3 holds for all the K samples, a statistically plausible way to estimate $\mathbf{q}_n$, $\mathbf{x}_n$ and unknown c_{kn}'s is the least-squares method. That is, $\mathbf{q}_n$, $\mathbf{x}_n$ and unknown c_{kn}'s can be estimated by the minimizers of the following least-squares criterion

$$L_1 = \sum_{k=1}^{K} ||c_{kn}\mathbf{x}_n - \mathbf{R}_k\mathbf{q}_n||^2 \tag{4}$$

where the vector pair, $\mathbf{x}_n$ and $\mathbf{q}_n$, are called coupled vectors, because they are dependent upon each other in the method. Analogously, one can derive another least-squares criterion

$$L_2 = \sum_{k=1}^{K} ||c_{kn}\mathbf{y}_n - \mathbf{R}_k^{\mathrm{T}}\mathbf{p}_n||^2 \tag{5}$$

where $\mathbf{p}_n$ is the nth contravariant vector of the base vectors, $\mathbf{x}_1$, ..., $\mathbf{x}_N$, and $\mathbf{y}_n$ and $\mathbf{p}_n$ are called coupled vectors, too. The minimizers of the criterion give the least-squares estimates of $\mathbf{y}_n$, $\mathbf{p}_n$ and unknown c_{kn}'s. Because the resolution of coupled vectors constitutes the core of the proposed method, it is called the coupled vectors resolution method.

In subsequent sections, it will be shown that the proposed method provides closed-form solutions to various cases of second-order calibration. For simplicity and without loss of generality, it is supposed that only one component, say the nth one, is of analytical interest.

2. Multisample Prediction with $\mathbf{x}_n$ and $\mathbf{y}_n$ Known: COVER1. It is known that with bilinear two-way measurements on unknown samples, the spectral and the chromatographic profiles, $\mathbf{x}_n$ and $\mathbf{y}_n$, or equivalently $\mathbf{x}_n\,\mathbf{y}_n{}^{\mathrm{T}}$, are sufficient for the quantitation of the component in unknown mixtures even in the presence of new interferences. It will be shown that with known $\mathbf{x}_n$ and $\mathbf{y}_n$ direct estimates of the concentrations of the analyte in K unknown samples can be derived using the COVER method.

In the case where $\mathbf{x}_n$ and $\mathbf{y}_n$ are known, the COVER method aims at seeking the estimates of c_{kn}'s and $\mathbf{q}_n$ which minimize the least-squares criterion, eq 4. A necessary condition for the estimates is

$$\partial L_1/\partial c_{kn} = 2\mathbf{x}_n^{\mathrm{T}}(\mathbf{x}_n c_{kn} - \mathbf{R}_k\mathbf{q}_n) = 0 \qquad k = 1, ..., K \tag{6}$$

$$\partial L_1/\partial \mathbf{q}_n = -2\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}(\mathbf{x}_n c_{kn} - \mathbf{R}_k\mathbf{q}_n) = 0 \tag{7}$$

One can derive from eqs 6 and 7 that

$$c_{kn} = \mathbf{x}_n{}^{\mathrm{T}}\mathbf{R}_k\mathbf{q}_n \qquad k = 1, ..., K \tag{8}$$

$$\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{x}_n\mathbf{x}_n^{\mathrm{T}}\mathbf{R}_k\mathbf{q}_n = \sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_k\mathbf{q}_n \tag{9}$$

It is important to note that, in the case of second-order calibration, the rank of $\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_k$ is intrinsically N, the number of components present in the K samples. This implies ill-conditioning of the eigenproblem, eq 9. Therefore, rank-reduced approximations to $\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_k$ and $\mathbf{R}_k$ ($k = 1, ..., K$) should be used for combating ill-conditioned solutions. Let the singular value decomposition (SVD) of $\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_k$ be

$$\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_k = \mathbf{U}\mathbf{S}^2\mathbf{U}^{\mathrm{T}} \tag{10}$$

where $\mathbf{U}$ and $\mathbf{S}$ are both $J \times J$ matrices. These matrices are truncated by removing the right-hand columns and the bottom rows to give $\bar{\mathbf{U}}$ ($J \times N$) and $\bar{\mathbf{S}}$ ($N \times N$). The rank-reduced approximation of $\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_k$ is thus given by the truncated SVD

$$\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_k \approx \bar{\mathbf{U}}\bar{\mathbf{S}}^2\bar{\mathbf{U}}^{\mathrm{T}} \tag{11}$$

Because $\bar{\mathbf{U}}$ spans the common subspace of the rows of $\mathbf{R}_k$ ($k = 1$, ..., K), then $\mathbf{R}_k$ ($k = 1$, ..., K) can be approximated by

$$\mathbf{R}_k \approx \mathbf{R}_k\,\bar{\mathbf{U}}\bar{\mathbf{U}}^{\mathrm{T}} \qquad k = 1, ..., K \tag{12}$$

In practice, N is generally unknown but can be estimated by the number of significant singular values or cross-validation. Moreover, provided N selected is not smaller than the actual number of components, the choice of N has little effect on the final results.

Substitution of eqs 11 and 12 into eq 9 yields

$$\bar{\mathbf{U}}^{\mathrm{T}}\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{x}_n\mathbf{x}_n^{\mathrm{T}}\mathbf{R}_k\bar{\mathbf{U}}\bar{\mathbf{U}}^{\mathrm{T}}\mathbf{q}_n = \bar{\mathbf{S}}^2\bar{\mathbf{U}}^{\mathrm{T}}\mathbf{q}_n \tag{13}$$

Letting

$$\alpha = \bar{\mathbf{S}}\bar{\mathbf{U}}^{\mathrm{T}}\mathbf{q}_n \tag{14}$$

one has

$$\bar{\mathbf{U}}^{\mathrm{T}}\,\mathbf{q}_n = \bar{\mathbf{S}}^{-1}\alpha \tag{15}$$

Substituting eqs 14 and 15 into eq 13, one can obtain

$$\bar{\mathbf{S}}^{-1}\bar{\mathbf{U}}^{\mathrm{T}}\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{x}_n\mathbf{x}_n^{\mathrm{T}}\mathbf{R}_k\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\alpha = \alpha \tag{16}$$

Now it can be concluded that α is the eigenvector of the symmetric matrix $\bar{\mathbf{S}}^{-1}\bar{\mathbf{U}}^{\mathrm{T}}\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{x}_n\mathbf{x}_n^{\mathrm{T}}\mathbf{R}_k\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}$ corresponding to the eigenvalue 1. It can be shown that the eigenvalues of the symmetric matrix are not greater than 1. Accordingly, α is the eigenvector associated with the largest eigenvalue, and one can easily solve the eigenproblem to obtain the solution of α. Then $\mathbf{q}_n$ can be calculated by

$$\mathbf{q}_n = a\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\alpha \tag{17}$$

where a is a constant, which can be determined using $\mathbf{y}_n{}^{\mathrm{T}}\mathbf{q}_n = 1$, an equality implied by eq 2. Therefore, one has

$$\mathbf{q}_n = \bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\alpha/\mathbf{y}_n^{\mathrm{T}}\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\alpha \tag{18}$$

With eq 18, the concentrations of the nth component in K unknown samples can be predicted immediately using eq 8.

The method above proposed, called COVER1, is an extension of Lorber's noniterative rank annihilation factor analysis (RAFA) procedure,[2] which can only be used for the case involving one unknown sample. In the case of one unknown sample, an analytical solution of eq 16 can be derived

$$\alpha = \bar{\mathbf{S}}^{-1}\bar{\mathbf{U}}^{\mathrm{T}}\mathbf{R}_1{}^{\mathrm{T}}\mathbf{x}_n \tag{19}$$

where $\bar{\mathbf{S}}$ and $\bar{\mathbf{U}}$ are given by the truncated SVD of $\mathbf{R}_1$.

$$\mathbf{R}_1 = \bar{\mathbf{V}}\bar{\mathbf{S}}\bar{\mathbf{U}}^{\mathrm{T}} \tag{20}$$

Therefore, one obtains

$$\mathbf{q}_n = \bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\bar{\mathbf{V}}^{\mathrm{T}}\mathbf{x}_n/\mathbf{y}_n^{\mathrm{T}}\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\bar{\mathbf{V}}^{\mathrm{T}}\mathbf{x}_n \tag{21}$$

$$c_{1n} = \mathbf{x}_n^{\mathrm{T}}\bar{\mathbf{V}}\bar{\mathbf{V}}^{\mathrm{T}}\mathbf{x}_n/\mathbf{y}_n^{\mathrm{T}}\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\bar{\mathbf{V}}^{\mathrm{T}}\mathbf{x}_n \tag{22}$$

The solution differs slightly from that given by Lorber's method.

There are two important cases for second-order calibration in which direct solutions of the spectral and the chromatographic profiles, $\mathbf{x}_n$ and $\mathbf{y}_n$, for the component under determination are available. One is the case where one has the two-way response $\mathbf{R}_0$ measured on the pure analyte of known concentration c_{0n}. That is, the measurement matrix is available

$$\mathbf{R}_0 = c_{0n}\mathbf{x}_n\mathbf{y}_n{}^{\mathrm{T}} + \mathbf{E}_0 \tag{23}$$

In this case, one can estimate $\mathbf{x}_n$ and $\mathbf{y}_n$ using SVD. Let the SVD of $\mathbf{R}_0/c_{0n}$ be

$$\mathbf{R}_0/c_{0n} = \mathbf{USV}^{\mathrm{T}} \tag{24}$$

then $\mathbf{x}_n$ and $\mathbf{y}_n$ can be estimated by $\mathbf{x}_n = \mathbf{u}_1$ and $\mathbf{y}_n = s_1\mathbf{v}_1$. Here $\mathbf{u}_1$ and $\mathbf{v}_1$ are the first columns of $\mathbf{U}$ and $\mathbf{V}$, respectively. s_1 is the first diagonal entry of $\mathbf{S}$. The other important case is where two or more calibration samples are available. This will be shown in the next section.

3. Resolution of $\mathbf{x}_n$ and $\mathbf{y}_n$ in the Case of Two or More Calibration Samples: COVER2. Suppose one has K ($K \geq 2$) calibration samples, in which the concentrations, c_{kn} ($k = 1, \ldots, K$), of the component under quantitation are all known. The estimates of the coupled vectors, $\mathbf{x}_n$ and $\mathbf{q}_n$, are given by the minimizers of the least-squares criterion, eq 4. A necessary condition for the estimates is that

$$\partial L_1/\partial\mathbf{q}_n = -2\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}(\mathbf{x}_n c_{kn} - \mathbf{R}_k\mathbf{q}_n) = 0 \tag{25}$$

$$\partial L_1/\partial\mathbf{x}_n = 2\sum_{k=1}^{K}c_{kn}(\mathbf{x}_n c_{kn} - \mathbf{R}_k\mathbf{q}_n) = 0 \tag{26}$$

One can derive from eqs 25 and 26 that

$$\mathbf{x}_n = \sum_{k=1}^{K}c_{kn}\mathbf{R}_k\mathbf{q}_n/\sum_{k=1}^{K}c_{kn}^2 \tag{27}$$

$$\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_k\mathbf{q}_n = \Big(\sum_{k,h=1}^{K}c_{kn}c_{hn}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_h/\sum_{k=1}^{K}c_{kn}^2\Big)\mathbf{q}_n \tag{28}$$

Analogously to the derivation of eq 16, one obtains

$$\Big(\bar{\mathbf{S}}^{-1}\bar{\mathbf{U}}^{\mathrm{T}}\sum_{k,h=1}^{K}c_{kn}c_{hn}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_h\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}/\sum_{k=1}^{K}c_{kn}^2\Big)\alpha = \alpha \tag{29}$$

where $\bar{\mathbf{S}}$ and $\bar{\mathbf{U}}$ are given by the truncated SVD of $\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_k$, eq 11, and α is analogously defined by eq 14. Notice that α is the eigenvector of the symmetric matrix $(\bar{\mathbf{S}}^{-1}\mathbf{U}^{\mathrm{T}}\sum_{k,h=1}^{K}c_{kn}c_{hn}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_h\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}/\sum_{k=1}^{K}c_{kn}^2)$ corresponding to the eigenvalue of 1. It can be proved that the eigenvalues of the symmetric matrix

are not larger than 1. Therefore, α is the eigenvector associated with the largest eigenvalue, and it can be easily approached. Then one can obtain the closed-form solution of $\mathbf{q}_n$ analogously in terms of eq 17. By keeping consistency with the preliminary assumption that $||\ \mathbf{x}_n\ || = 1$, the constant a in eq 17 can be determined. Therefore,

$$\mathbf{q}_n = (\sum_{k=1}^{K} c_{kn}^2 / ||\sum_{k=1}^{K} c_{kn}\mathbf{R}_k\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\alpha||)\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\alpha \qquad (30)$$

and the spectral profile, $\mathbf{x}_n$, can be calculated using eq 27.

Analogously, the coupled vectors, $\mathbf{y}_n$ and $\mathbf{p}_n$, can be estimated by the minimizers of the least-squares criterion, eq 5. It can be shown that

$$(\bar{\mathbf{S}}^{-1}\bar{\mathbf{V}}^{\mathrm{T}}\sum_{k,h=1}^{K} c_{kn}c_{hn}\mathbf{R}_k\mathbf{R}_h^{\mathrm{T}}\bar{\mathbf{V}}\bar{\mathbf{S}}^{-1} / \sum_{k=1}^{K} c_{kn}^2)\boldsymbol{\beta} = \boldsymbol{\beta} \qquad (31)$$

where $\bar{\mathbf{S}}$ ($N \times N$) and $\bar{\mathbf{V}}$ ($I \times N$) are given by the truncated SVD of $\sum_{k=1}^{K}\mathbf{R}_k\mathbf{R}_k^{\mathrm{T}}$, that is,

$$\sum_{k=1}^{K}\mathbf{R}_k\mathbf{R}_k^{\mathrm{T}} \approx \bar{\mathbf{V}}\bar{\mathbf{S}}\bar{\mathbf{V}}^{\mathrm{T}} \qquad (32)$$

The coupled vectors can be calculated as follows

$$\mathbf{p}_n = b\bar{\mathbf{V}}\bar{\mathbf{S}}^{-1}\boldsymbol{\beta} \qquad (33)$$

$$\mathbf{y}_n = \sum_{k=1}^{K} c_{kn}\mathbf{R}_k^{\mathrm{T}}\mathbf{p}_n / \sum_{k=1}^{K} c_{kn}^2 \qquad (34)$$

where b is a constant, which can be uniquely determined using eq 2.

So far it has been demonstrated that in the case where two or more calibration samples are available, closed-form solutions to the spectral profiles for the component of interest in each order can be approached only with the calibration set using the proposed COVER method, called COVER2. With the spectral profiles in each order thus estimated, concentrations of the component in multiple unknown mixtures can be predicted directly by using the method developed in the preceding section.

4. Quantitation in the Case of One Calibration Sample: COVER3. It is acknowledged that only one calibration mixture, in which the concentration of the component under determination is known, is insufficient for quantifying the concentrations of the analyte in unknown mixtures, unless extra information is available.[5] One can at most calculate the relative concentrations and the spectral profiles in each order for the components. Identification of the sought-for analyte is practically unfeasible. Conventionally, to identify the component of interest, the actual spectral profile of the analyte in a certain order need to be used for matching to the calculated one. The concentrations of the component are then computed from its relative concentrations and its concentration in the calibration sample. It will be shown that, by utilizing the information of the spectral profile, computationally easily achievable closed-form solutions of the concentrations in multiple unknown samples can be obtained using the method developed here. The method, called COVER3, is merely slightly different from that proposed in section 2.

In fact, suppose one has a calibration sample and the spectral profile in a certain order, say $\mathbf{x}_n$, then the concentrations, c_{kn} ($k = 2, ..., K$), of the analyte in $K - 1$ unknown mixtures and the nth contravariant vector, $\mathbf{q}_n$, can be estimated by the minimizers of the least-squares criterion, eq 4. Following the derivation in section 2, one obtains that $\mathbf{q}_n$ can be estimated using eqs 16 and 17. The constant a in eq 17 can be determined from eq 8 and the concentration of the component in the calibration sample. One thus has,

$$\mathbf{q}_n = (c_{1n} / \mathbf{x}^{\mathrm{T}}{}_n\mathbf{R}_1\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\alpha)\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\alpha \qquad (35)$$

where c_{1n} and $\mathbf{R}_1$ are the concentration of the component and the two-way response for the calibration sample, respectively. Analogously, α is given by eq 16. $\bar{\mathbf{U}}$ and $\bar{\mathbf{S}}$ are calculated by eq 11. With $\mathbf{q}_n$ achieved, the concentrations of the analyte can be estimated using eq 8 for the $K - 1$ unknown mixtures ($k = 2, ..., K$).

5. Algorithms and Implementation. In this section, the implementation of the COVER methods in various cases of second-order calibration is discussed.

The first case of second-order calibration is that where the profiles of the component under quantitation in x and y orders are both known a priori. As mentioned above, a typical example of this case is that the two-way response of the pure analyte of known concentration is available. In this case, the COVER1 method can be used to directly determine the concentrations of the analyte in unknown mixtures. The computing procedure is given as follows.

a1. Compute the eigenvector α of the symmetric matrix $\bar{\mathbf{S}}^{-1}\bar{\mathbf{U}}^{\mathrm{T}}\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{x}_n\mathbf{x}_n^{\mathrm{T}}\mathbf{R}_k\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}$ corresponding to the largest eigenvalue.

a2. Calculate $\mathbf{q}_n$ by

$$\mathbf{q}_n = \bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\alpha / \mathbf{y}_n^{\mathrm{T}}\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\alpha$$

a3. Calculate c_{kn} by

$$c_{kn} = \mathbf{x}_n{}^{\mathrm{T}}\mathbf{R}_k\mathbf{q}_n \qquad k = 1, ..., K$$

The second case of second-order calibration is that where the profiles $\mathbf{x}_n$ and $\mathbf{y}_n$ are unavailable, but one has two or more calibration samples in which the concentrations of the component under determination are known. In this case, the COVER2 method can be used to resolve directly the profiles $\mathbf{x}_n$ and $\mathbf{y}_n$ of the analyte. The procedure for resolving $\mathbf{x}_n$ is given by

b1. Compute the eigenvector α of the symmetric matrix $\bar{\mathbf{S}}^{-1}\bar{\mathbf{U}}^{\mathrm{T}}\sum_{k,h=1}^{K} c_{kn}c_{hn}\mathbf{R}_k^{\mathrm{T}}\mathbf{R}_h\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1} / \sum_{k=1}^{K} c_{kn}^2$ corresponding to the largest eigenvalue.

b2. Calculate $\mathbf{q}_n$ by

$$\mathbf{q}_n = (\sum_{k=1}^{K} c_{kn}^2 / ||\sum_{k=1}^{K} c_{kn}\mathbf{R}_k\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\alpha||)\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\alpha$$

b3. Calculate $\mathbf{x}_n$ by

$$\mathbf{x}_n = \sum_{k=1}^{K} c_{kn}\mathbf{R}_k\mathbf{q}_n / \sum_{k=1}^{K} c_{kn}^2$$

Then the procedure for estimating $\mathbf{y}_n$ is given by

c1. Compute the eigenvector $\boldsymbol{\beta}$ of the symmetric matrix $\bar{\mathbf{S}}^{-1}\bar{\mathbf{V}}^{\mathrm{T}}\sum_{k,h=1}^{K} c_{kn}c_{hn}\mathbf{R}_k\mathbf{R}_h^{\mathrm{T}}\bar{\mathbf{V}}\bar{\mathbf{S}}^{-1}/\sum_{k=1}^{K} c_{kn}^2$ corresponding to the largest eigenvalue.

c2. Calculate $\mathbf{p}_n$ by

$$\mathbf{p}_n = \bar{\mathbf{V}}\bar{\mathbf{S}}^{-1}\boldsymbol{\beta} / \mathbf{x}_n^{\mathrm{T}}\bar{\mathbf{V}}\bar{\mathbf{S}}^{-1}\boldsymbol{\beta}$$

c3. Calculate $\mathbf{y}_n$ by

$$\mathbf{y}_n = \sum_{k=1}^{K} c_{kn}\mathbf{R}_k^{\mathrm{T}}\mathbf{p}_n / \sum_{k=1}^{K} c_{kn}^2$$

With $\mathbf{x}_n$ and $\mathbf{y}_n$ thus identified, the second case is transformed into the first one, and one can use COVER1 for the quantification of the component in unknown mixtures.

The third case of second-order calibration is that where one has the profile of the analyte in a certain order, say $\mathbf{x}_n$, and one calibration sample in which the concentration of the analyte, c_{kn}, is known. In this case, the COVER3 method can be used to directly estimate the concentrations of the component in unknown mixtures. The computing procedure is implemented as follows.

d1. Compute the eigenvector α of the symmetric matrix $\bar{\mathbf{S}}^{-1}\bar{\mathbf{U}}^{\mathrm{T}}\sum_{k=1}^{K}\mathbf{R}_k^{\mathrm{T}}\mathbf{x}_n\mathbf{x}_n^{\mathrm{T}}\mathbf{R}_k\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}$ corresponding to the largest eigenvalue.

d2. Calculate $\mathbf{q}_n$ by

$$\mathbf{q}_n = (c_{1n}/\mathbf{x}_n^{\mathrm{T}}\mathbf{R}_1\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\alpha)\bar{\mathbf{U}}\bar{\mathbf{S}}^{-1}\alpha$$

d3. Calculate c_{kn} by

$$c_{kn} = \mathbf{x}_n{}^{\mathrm{T}}\mathbf{R}_k\mathbf{q}_n \qquad k = 2, ..., K$$

Note that in the aforementioned three cases the prior knowledge is indispensable for quantification of the analyte. In analytical practice one sometimes encounters the cases where excessive knowledge is available for second-order calibration. An example is the case where one has the profile $\mathbf{x}_n$ and more than one calibration samples. In this case, one can simply use the COVER3 method for estimating the relative concentrations of the analytes in all samples and then utilize the calibration samples to determine the absolute values of these concentrations.

EXPERIMENTAL SECTION

1. Simulated Example. The spectral profiles of four components, $\mathbf{s}_1$, $\mathbf{s}_2$, $\mathbf{s}_3$, and $\mathbf{s}_4$, were generated by

$$s_{1,i} = 0.2\ \mathrm{gs}(2i - 1, 30, 30) + 0.5\ \mathrm{gs}(2i - 1, 70, 10) \qquad i = 1, 2, ..., 50$$

$$s_{2,i} = 0.6\ \mathrm{gs}(2i - 1, 20, 10) + 0.3\ \mathrm{gs}(2i - 1, 80, 30) \qquad i = 1, 2, ..., 50$$

$$s_{3,i} = 0.7\ \mathrm{gs}(2i - 1, 40, 10) + 0.2\ \mathrm{gs}(2i - 1, 90, 20) \qquad i = 1, 2, ..., 50$$

$$s_{4,i} = 0.7\ \mathrm{gs}(2i - 1, 50, 25) \qquad i = 1, 2, ..., 50$$

where $\mathrm{gs}(x, a, b)$ is the value at x of Gaussian function with center a and standard deviation b; i.e., $\mathrm{gs}(x, a, b) = \exp\{-(x - a)^2/2b^2\}$. The chromatographic profiles of the components, $\mathbf{c}_1$, $\mathbf{c}_2$, $\mathbf{c}_3$, and $\mathbf{c}_4$, were simulated by

$$c_{1,i} = 0.5\ \mathrm{gs}(4i - 3, 40, 5) \qquad i = 1, 2, ..., 20$$

$$c_{2,i} = 0.5\ \mathrm{gs}(4i - 3, 30, 10) \qquad i = 1, 2, ..., 20$$

$$c_{3,i} = 0.5\ \mathrm{gs}(4i - 3, 50, 10) \qquad i = 1, 2, ..., 20$$

$$c_{4,i} = 0.5\ \mathrm{gs}(4i - 3, 40, 9) \qquad i = 1, 2, ..., 20$$

Ten samples were generated, in which the first five samples only contained the components of analytical interest, i.e., the first three components, with the analyte concentrations uniformly distributed in the range of 0−1. The remaining five samples contained the three analytes as well as an interference, i.e., the fourth component, with the concentrations of each component uniformly distributed in the range of 0−1. The two-way response of each sample were generated exactly in terms of eq 1, in which the random errors were normally distributed with a mean of 0 and a standard deviation of 0.002. To investigate the performance of the proposed method, these 10 samples were divided into two sets. The calibration set was composed of the first five samples, and the prediction set consisted of the remaining five samples. With the calibration samples, COVER2 was used for the resolution of the spectral and the chromatographic profiles for each analyte. COVER1 was then used to estimate the concentrations of the analytes in the samples. To examine the behavior of COVER3, the first of these 10 samples was used as a calibration sample. With the spectral profiles assumed known, the concentrations of the samples were estimated using COVER3. For comparison, these 10 samples were also treated using the TLD method and the PARAFAC algorithm.

2. HPLC-DAD Data. Nine mixtures of *o*-dichlorobenzene (*o*-DCB), *p*-chlorotoluene (*p*-CT), and *o*-chlorotoluene (*o*-CT) as well as an internal retention time standard, chlorobenzene (CB), were analyzed using a high-performance liquid chromatography (HPLC) system with diode array detection. The concentrations for each component are shown in Table 1. Details of the experimental procedures are as given previously.[28] With the samples 3−6 as the calibration set, COVER2 was used to resolve the spectral and the chromatographic profiles of the analytes, *p*-CT and *o*-CT, as well as the interference *o*-DCB. COVER1 was then applied to estimating the concentrations of the two analytes in the samples. With the spectral profiles experimentally obtained and sample 1 as a calibration sample, the concentrations of *p*-CT in the samples were calculated by COVER3. With the spectral profiles experimentally obtained and sample 2 as a calibration sample, the concentrations of *o*-CT in the samples were also estimated using COVER3. For comparison, these nine samples were also treated using the TLD method and the PARAFAC algorithm.

Table 1. Compositions of Nine Mixtures in HPLC-DAD Data

	concentration (μg mL^{-1})								
sample	1	2	3	4	5	6	7	8	9
o-DCB	0.0	0.0	0.0	0.0	152.2	15.2	60.8	91.2	30.4
p-CT	75.6	0.0	50.4	25.2	12.6	12.6	25.2	50.4	75.6
o-CT	0.0	91.2	30.4	60.8	15.2	152.0	91.2	30.4	60.8
CB	62.4	62.4	62.4	62.4	62.4	62.4	62.4	62.4	62.4

All computer programs were written in Matlab and run on a personal computer (Pentium processor). The algorithm for TLD used in the investigation was the improved version developed by Li et al.[10] and Booksh and co-workers.[15] The PARAFAC algorithm used in the study is the version given by Krijnen.[30] The stopping criterion for the PARAFAC algorithm is that the improvement of the PARAFAC error between consecutive iterations is less than 10^{-5}, or the total computational epochs are greater than a predefined maximum, set to 10 000 in the investigation.

RESULTS AND DISCUSSION

1. Simulated Example. The goal of the simulated experiments was two-fold. First, the performance of the proposed method was investigated in comparison with the TLD method as well as the PARAFAC algorithm. Throughout the investigation, the number of components in the samples was chosen to be four, which was determined by the number of significant singular values given by eqs 10 and 32. The spectral and the chromatographic profiles of the three analytes resolved using COVER2 with the five-sample calibration set are depicted against the actually simulated ones in Figure 1a and d.

It took 338 cycles for PARAFAC to achieve the resolution of the 10-sample data. The resulting spectral and chromatographic profiles of the three components are shown in Figure 1c and f. One can observe that the profiles estimated by COVER2 fit the actual ones as well as those given by PARAFAC. The spectral and chromatographic profiles of the three analytes calculated using TLD with all the 10 samples are plotted in Figure 1b and e. It can be seen that the deviations between the estimated and the actual profiles are much larger than those given by COVER2 and PARAFAC. With the 5-sample calibration set, the concentrations of the analytes in 10 mixtures were estimated by TLD as well as PARAFAC. For TLD, the mean squared errors (MSEs) between the estimated and the actual concentrations were 1.4179×10^{-3}, 7.2776×10^{-5}, and 1.1752×10^{-4}, respectively, for the three analytes, while for PARAFAC the MSEs were 3.8450×10^{-6}, 1.6114×10^{-6}, and 2.9667×10^{-6}, respectively, for the three components. With the profiles estimated by COVER2, the concentrations of the analytes in the 10 samples were calculated using COVER1. The MSEs of the calculated concentrations from the actual ones were 1.0699×10^{-4}, 9.1394×10^{-6}, and 2.0322×10^{-5}, respectively, for the three analytes. Using the first sample as calibration set and the simulated spectral profile of each analyte, the concentrations of the analytes in the 10 samples were predicted separately by COVER3. The MSEs were 1.3263×10^{-4}, 2.5999×10^{-5}, and 7.2945×10^{-5}, respectively, for the three analytes. For comparison with the first sample as calibration set and the spectral profiles simulated, the concentrations of the three components in 10 mixtures were estimated separately using the restricted PARAFAC algorithm in which one of the spectral profiles was restricted to the known spectral profile of the analyte. The MSEs were 4.0420×10^{-6}, 3.3726×10^{-5}, and 3.3823×10^{-5}, respectively, for the three analytes. The concentrations of the first component in 10 samples calculated by different methods are shown in Table 2. It can be seen that for the simulated data the estimating errors given by COVER1 and COVER3 are a little larger than those given by PARAFAC and its restricted version. However, it was found that for 10 times (or runs) one started PARAFAC from randomly selected values, this algorithm converged to degenerate solutions twice. This indicates the PARAFAC algorithm is subjected to the danger of getting trapped into degenerate solutions. Moreover, for the simulated data when the component number was not chosen to be four, the performance of PARAFAC was undesirably bad, hinting that the behavior of PARAFAC was very sensitive to the choice of the component number. Because in practical problem solving determination of the component number in unknown mixtures is usually a difficulty that is hard to handle, the aforementioned results do suggest that in practical problem solving the robustness of PARAFAC to errors is frequently offset by the degenerate solutions and its sensitivity to the component number. This can be shown from the results of the real data presented below.

Second, the effect of the estimated number of components present in the samples, N, on the behavior of the proposed method was examined. The results are shown in Figure 2. It was found that, when N selected was smaller than the number of components actually present in the samples, the profiles in each order resolved by COVER2 deviated severely from the true ones. The performance of COVER3 in concentration estimation of unknown mixtures was also very poor. However, as N increased and became equal to or greater than the actual number of components in the samples, the performance of COVER2 and COVER3 tended to be very stable and was little affected by the estimates of N. Similar observations were obtained for COVER1, since it differs merely slightly from COVER3. The results really indicate an attractive characteristic of the proposed method in that its performance is very stable with respect to the overestimates of the component number. Because an upper bound of the component number can always be easily estimated, this characteristic does imply an advantage of the COVER method over the PARAFAC algorithm in that the COVER method is free of the difficulty of determining the number of components present in the samples, while for PARAFAC one cannot circumvent this difficulty, since the performance of PARAFAC is very sensitive to the choice of the component number. It seems that this advantage derives from the fact that one uses SVD in eqs 9 and 28 to obtain well-defined eigenproblems.

2. HPLC-DAD Data. With the four-sample calibration set, COVER2 was applied to the resolution of the spectral and the chromatographic profiles of *o*-DCB, *p*-CT, and *o*-CT. The spectral profiles of the three components computed by COVER2 are plotted against those experimentally measured from pure components in Figure 3a. One can observe a very good fit of the estimated profiles to the actual ones. This further confirmed the performance of

(30) Krijnen, W. P. *The Analysis of Three-Way Arrays by Constrained PARAFAC Methods*; DSWO Press: Leiden, 1993.

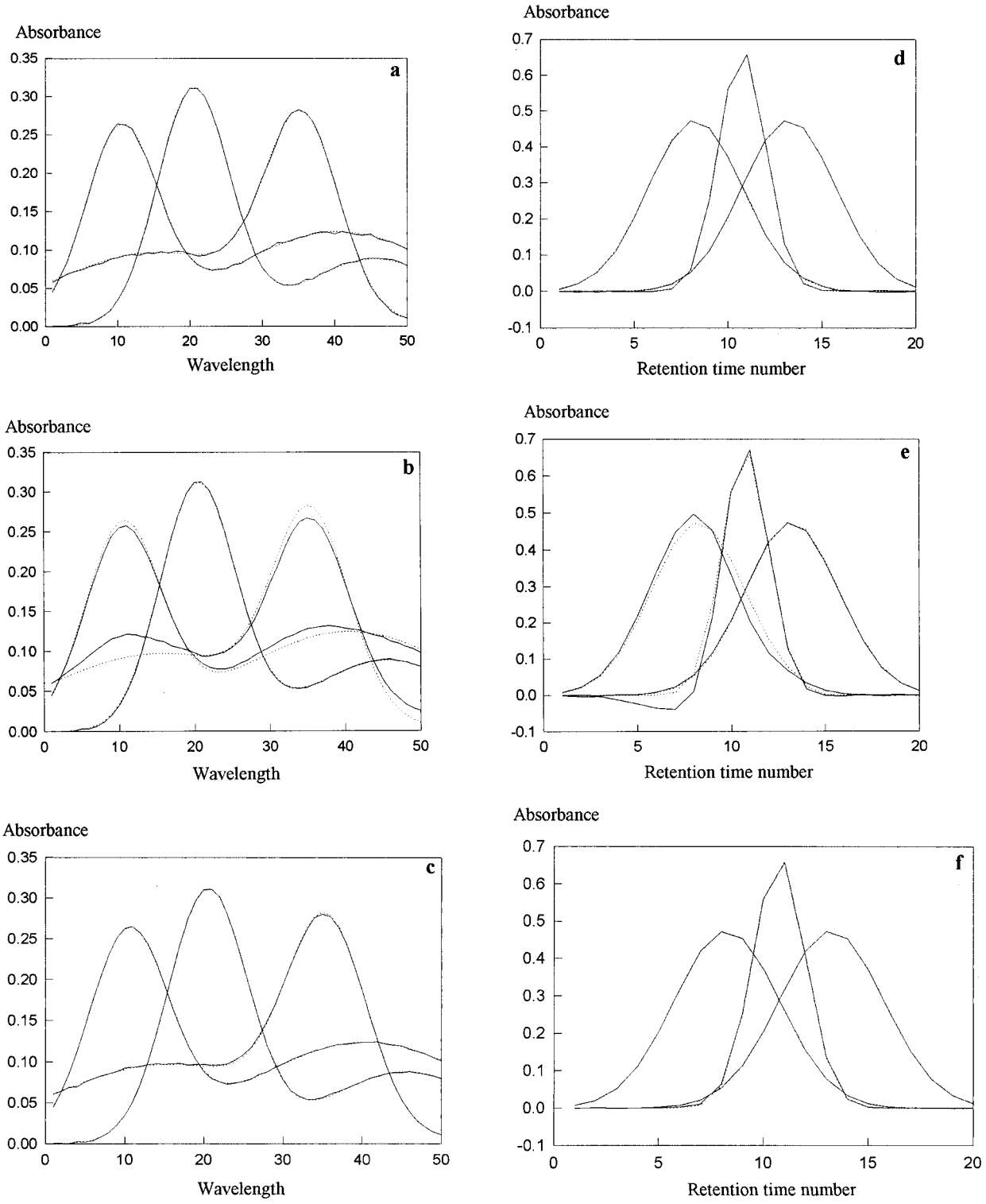

Figure 1. Spectral and chromatographic profiles of simulated data. (a) Spectral profiles simulated (dotted line) and calculated by COVER2 (solid line). (b) Spectral profiles simulated (dotted line) and calculated by TLD (solid line). (c) Spectral profiles simulated (dotted line) and calculated by PARAFAC (solid line). (d) Chromatographic profiles simulated (dotted line) and calculated by COVER2 (solid line). (e) Chromatographic profiles simulated (dotted line) and calculated by TLD (solid line). (f) Chromatographic profiles simulated (dotted line) and calculated by PARAFAC (solid line).

COVER2. Using all the nine samples, TLD and PARAFAC were used for calculating the spectral profiles, too. It was found that with different numbers of components chosen for the model, the performance of TLD and PARAFAC varied, and the best resolutions for PARAFAC and TLD were both achieved when the component number was set to three. It took 774 iterations for PARAFAC to locate the best solution. The spectral profiles of the three components obtained using TLD and PARAFAC are depicted against the experimentally measured ones in Figure 3b and c. The discrepancies between the calculated and the measured

Table 2. Concentrations of the First Component in Ten Samples in Simulated Data

sample no.	actual values	values calculated by different methods: TLD	COVER1	PARAFAC	COVER3	restricted PARAFAC[a]
1	0.8130	0.7715	0.8118	0.8145	0.8130	0.8130
2	0.8979	0.9234	0.8877	0.8959	0.8899	0.8953
3	0.2446	0.2304	0.2496	0.2473	0.2499	0.2468
4	0.7606	0.7307	0.7594	0.7603	0.7607	0.7590
5	0.6949	0.7402	0.6954	0.6951	0.6980	0.6949
6	0.4519	0.4643	0.4541	0.4484	0.4558	0.4481
7	0.2636	0.2304	0.2773	0.2643	0.2776	0.2635
8	0.6655	0.6878	0.6787	0.6631	0.6821	0.6626
9	0.1166	0.1360	0.1347	0.1178	0.1363	0.1178
10	0.2096	0.2904	0.2252	0.2074	0.2283	0.2082
MSE[b]		1.4179[c]	1.0699[d]	3.8450[e]	1.3263[d]	4.0420[e]

[a] The restricted PARAFAC algorithm is the one where the knowledge of known spectra of the component under determination is employed. [b] MSE is the mean squared error. [c] The number is multiplied by 10^3. [d] The number is multiplied by 10^4. [e] The number is multiplied by 10^6.

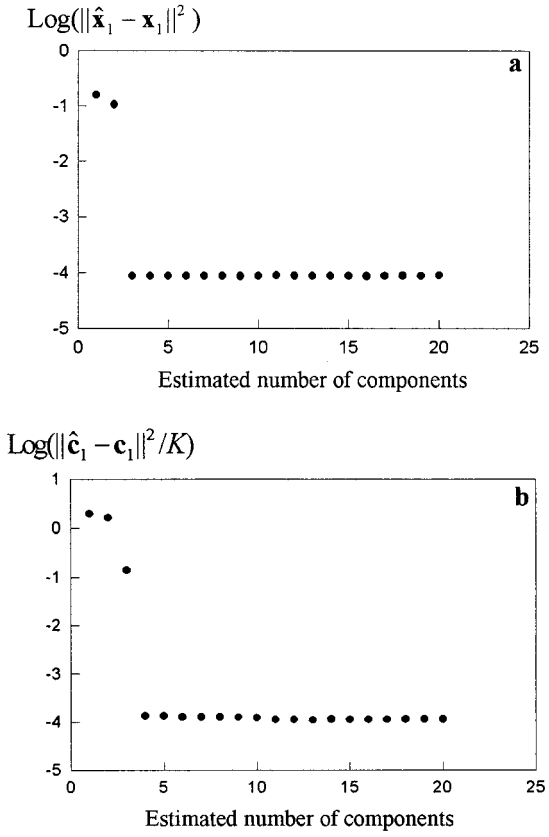

Figure 2. Relationship of the errors of spectral resolution and concentration estimation for the first component to the estimated number of components present in the samples. (a) Relationship between the logarithmic Euclidean distance of the spectral profiles estimated by COVER2 from the actual ones and the estimated number of components. The actual number of components in the calibration set is 3. (b) Relationship between the logarithmic Euclidean distance of the concentrations estimated by COVER3 from the actual ones and the estimated number of components. The actual number of components in the prediction set is 4.

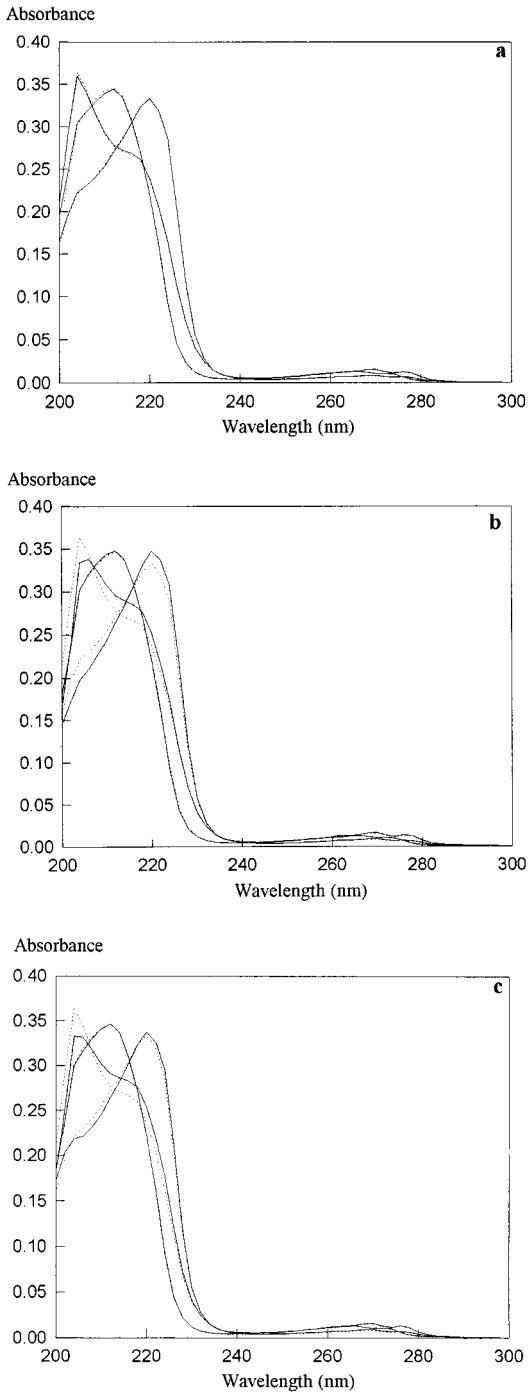

Figure 3. Spectral and chromatographic profiles of HPLC-DAD data. (a) Spectral profiles experimentally measured from pure compounds (dotted line) and calculated by COVER2 from mixture samples (solid line). (b) Spectral profiles experimentally measured from pure compounds (dotted line) and calculated by TLD from mixture samples (solid line). (c) Spectral profiles experimentally measured from pure compounds (dotted line) and calculated by PARAFAC from mixture samples (solid line).

profiles are much larger than that given by COVER2. With the four calibration samples, the concentrations were computed by TLD as well as PARAFAC. The results are shown in Tables 3 and 4. For TLD the MSEs were 7.9 and 56.9, respectively, for *p*-CT and *o*-CT, while for PARAFAC the MSEs were 11.9 and 31.8, respectively, for *p*-CT and *o*-CT. Using the profiles resolved by COVER2, COVER1 was applied to estimating the concentrations of the two analytes in nine samples. The calculated concentrations are shown in Tables 3 and 4. The MSEs were 5.8 and 9.5,

Table 3. Concentrations of *p*-CT in Nine Mixtures in HPLC-DAD Data

sample no.	actual values	values calculated by different methods				
		TLD	COVER1	PARAFAC	COVER3	restricted PARAFAC[a]
1	75.6	77.2	76.0	77.3	75.6	75.6
2	0	−0.2	2.1	−0.4	2.1	0.1
3	50.4	51.4	51.4	51.4	51.1	50.4
4	25.2	25.6	26.8	25.5	26.6	25.2
5	12.6	6.6	16.3	6.2	16.3	7.0
6	12.6	10.6	12.1	10.4	12.4	10.9
7	25.2	19.5	26.6	17.4	26.1	17.7
8	50.4	48.8	55.8	47.8	55.3	47.4
9	75.6	75.1	77.8	74.3	77.3	73.1
MSE[b]		7.9	5.8	11.9	4.8	10.5

[a] The restricted PARAFAC algorithm is the one where the knowledge of known spectra of the component under determination is employed. [b] MSE is the mean squared error.

Table 4. Concentrations of *o*-CT in Nine Mixtures in HPLC-DAD Data

sample no.	actual values	values calculated by diferent methods				
		TLD	COVER1	PARAFAC	COVER3	restricted PARAFAC[a]
1	0	−6.4	2.1	1.8	2.0	1.1
2	91.2	99.8	89.8	100.7	91.2	91.2
3	30.4	29.0	31.3	34.7	31.8	31.2
4	60.8	64.4	60.6	67.7	61.5	61.2
5	15.2	29.1	16.1	11.3	17.1	10.0
6	152.0	148.0	148.2	148.1	148.7	134.1
7	91.2	104.3	83.4	97.7	84.6	88.2
8	30.4	36.1	28.1	29.2	28.3	26.0
9	60.8	65.9	63.3	70.0	64.2	62.9
MSE[b]		56.9	9.5	31.8	8.0	38.2

[a] The restricted PARAFAC algorithm is the one where the knowledge of known spectra of the component under determination is employed. [b] MSE is the mean squared error.

respectively, for *p*-CT and *o*-CT. Using sample 1 as the calibration set and the experimentally measured spectral profile of *p*-CT, the concentrations of *p*-CT in the samples were computed using COVER3 and the restricted PARAFAC algorithm. The results obtained are shown in Table 3. The MSEs were 4.8 and 10.5, respectively, for COVER3 and the restricted PARAFAC algorithm. With the second sample as the calibration set and the experimentally obtained spectral profile of *o*-CT, COVER3 and the restricted PARAFAC algorithm were used to estimate the concentrations of *o*-CT in the samples. The concentrations computed are shown in Table 4. The MSEs were 8.0 and 38.2, respectively, for COVER3 and the restricted PARAFAC algorithm. The results verified the conclusion obtained in previous simulated experiments that the proposed method could provide acceptable performance in second-order calibration.

CONCLUSIONS

The coupled vectors resolution method has been developed. The method can produce directly solutions to various cases of second-order calibration and can generalize in a direct manner to new unknown samples. The presented results demonstrate that the proposed method is capable of yielding accurate estimates for the profiles of the analytes in each order and for the concentrations of the analytes, and the performance of the proposed method is very stable when the number of components is chosen to be equal to or greater than the actual number present in the samples. This offers the advantage that in second-order calibration one can circumvent the difficulty of determining a proper number of components present in the samples.

ACKNOWLEDGMENT

The work is supported by the Natural Science Foundation of China (Grant 29735150).

SUPPORTING INFORMATION AVAILABLE

Appendix A, showing that the eigenvalues of the symmetric matrix in eq 16 are not larger than 1, Appendix B, proving that the eigenvalues of the symmetric matrix in eq 29 are not larger than 1, and Appendix C, giving the MATLAB coded programs for COVER1, COVER2 and COVER3. This material is available free of charge via the Internet at http://pubs.acs.org.

Received for review January 7, 1999. Accepted April 29, 1999.

AC990009N