ELSEVIER

Chemometrics and Intelligent Laboratory Systems 45 (1999) 335–338

Chemometrics and intelligent laboratory systems

# On the equivalence of window factor analysis and orthogonal projection resolution

Qingsong Xu [a], Yizeng Liang [b,*]

[a] *Department of Mathematics, Institute of Chemometrics and Chemical Sensing Technology, Hunan University, Changsha, 410082, China*
[b] *Institute of Chemometrics and Chemical Sensing Technology, College of Chemistry and Chemical Engineering, Hunan University, Changsha, 410082, China*

**Abstract**

Window factor analysis (WFA) and orthogonal projection resolution (OPR) are two factor analytical techniques for extracting component concentration profiles from two-way evolutionary data. Both methods take advantage of the fact that each component lies in a special region along the evolutionary axis. Theoretical equations are derived to prove that WFA is equivalent to OPR, if errors are ignored. 

*Keywords:* Window factor analysis; Orthogonal projection resolution; Equivalence

**Contents**



## 1. Introduction

A variety of factor analytical techniques has been developed for the purpose of extracting component concentration profiles from evolutionary processes. These methods take advantage of the fact that each component lies in a special region along the evolutionary axis. Window factor analysis (WFA) [1] and orthogonal projection resolution (OPR) [2,3] are two of them. Both are based upon principal component analysis. By locating the region of existence of a component, the concentration profile of that compo-

* Corresponding author.

0169-7439/99/$ - see front matter 
PII: S0169-7439(98)00140-3

nent can be determined. The two procedures perform principal component analysis on the submatrix comprising the zero-concentration region of each component. They differ in the way the concentration profiles are obtained. WFA projects all of the mixture spectra to a vector that has the information of only one component and get the concentration profile of this component. On the other hand, OPR projects all of the mixture spectra to a subspace that is perpendicular to the subspace generated by all the components except one and obtains the concentration profile of that component.

Theoretical expressions are derived to prove that the results based on WFA are the same as those based on OPR, if errors are ignored. That is to say, in resolution of two-way data from hyphenated evolutionary processes, such as chromatography, flow injection, reaction kinetics, etc. WFA and OPR are equivalent to each other.

## 2. Data from hyphenated evolutionary processes

The $m \times n$ matrix $\mathbf{X}$ discussed here contains data in the time direction and the wavelength direction. The $m$ rows correspond to spectra taken at regular time intervals and $n$ columns represent concentrations profiles measured at successive wavelengths. For a mixture of $d$ components (compounds), the matrix $\mathbf{X}$ can be expressed as the sum of $d$ bilinear matrices, one for each component:

$$\mathbf{X} = \sum_{i=1}^{d} \boldsymbol{c}_i \boldsymbol{s}_i^T = \mathbf{C}\mathbf{S}^T \quad (1)$$

Where $\{\boldsymbol{c}_i = (c_{i1}, c_{i2}, \ldots, c_{im})^T, i = 1, 2, \ldots, d\}$, $\{\boldsymbol{s}_i = (s_{i1}, s_{i2}, \ldots, s_{in})^T, i = 1, 2, \ldots, d\}$ are the unknown concentration and spectral vectors, respectively, of the $d$ components.

By principal component analysis [4], this matrix is separated into a product of an abstract spectral matrix $\mathbf{P}$ and an abstract concentration matrix $\mathbf{T}$.

$$\mathbf{X} = \sum_{i=1}^{d} \boldsymbol{t}_i \boldsymbol{p}_i^T = \mathbf{T}\mathbf{P}^T \quad (2)$$

The columns of matrix $\mathbf{P}$ are mutually orthonormal. The columns of matrices $\mathbf{P}$ and $\mathbf{S}$ span the same linear space. The vectors $\boldsymbol{p}_i$ $(i = 1, 2, \ldots, d)$ forming the columns of matrix $\mathbf{P}$ are linear combinations of the true spectral vectors (the columns of matrix $\mathbf{S}$). The true spectral vectors can be expressed, conversely, as a linear combinations of the abstract spectra,

$$\boldsymbol{s}_i = \sum_{j=1}^{d} \alpha_{ij} \boldsymbol{p}_j^T \quad (3)$$

where $\alpha_{ij}$ is the linear coefficients.

## 3. Window factor analysis (WFA)

Let $\mathbf{X}^k$ be a submatrix of $\mathbf{X}$ that is constructed by deleting the elution 'window' (the region of existence) of component $k$ (Fig. 1). Principal component analysis of submatrix $\mathbf{X}^k$ yields a matrix $\mathbf{P}^k$ containing $d-1$ orthonormal abstract spectral vectors $\boldsymbol{p}_i^k$ $(i = 1, 2, \ldots, d-1)$ and a matrix $\mathbf{T}^k$ containing $d-1$ abstract concentration profile vectors $\boldsymbol{t}_i^k$ $(i = 1, 2, \ldots, d-1)$.

$$\mathbf{X}^k = \sum_{i=1}^{d-1} \boldsymbol{t}_i^k \left(\boldsymbol{p}_i^k\right)^T = \mathbf{T}^k \left(\mathbf{P}^k\right)^T \quad (4)$$

Because $\boldsymbol{s}_i$ $(i = 1, 2, \ldots, d, i \neq k)$, the columns of matrix $\mathbf{S}$, belong to the linear space spanned by the columns of matrix $\mathbf{P}^k$, $\boldsymbol{s}_i$ satisfies Eq. (5).

$$\boldsymbol{s}_i = \sum_{j=1}^{d-1} \gamma_{ij} \boldsymbol{p}_j^k \quad (5)$$

In fact, the space spanned by the columns of the matrix $\mathbf{P}^k$ is $(d-1)$-dimensional subspace of the full $d$-dimensional factor space. The vectors $\boldsymbol{p}_j^k$ $(j = 1, 2, \ldots, d-1)$ form a basis of this subspace. We can expand this subspace to $d$-dimensional space by finding a vector $\boldsymbol{p}_d^k$ which is orthonormal to these $d-1$ basis vectors. And this $d$-dimensional space is

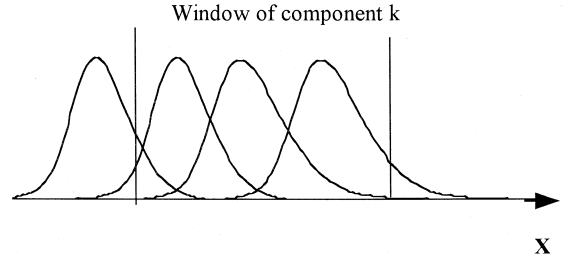


Fig. 1. Window of component $k$ in retention time ($\mathbf{X}$) direction.

congruent with the space spanned by the columns of the matrix **S**. Hence,

$$\boldsymbol{p}_i = \sum_{j=1}^{d} \kappa_{ij} \boldsymbol{p}_j^k \quad (i = 1,2,\ldots,d) \tag{6}$$

Because of orthonormality, $\kappa_{ij}$ is easily calculated.

$$\kappa_{ij} = \boldsymbol{p}_i^T \boldsymbol{p}_j^k \quad (j = 1,2,\ldots,d) \tag{7}$$

In order to determine $\boldsymbol{p}_d^k$, consider the sum of $\boldsymbol{p}_i$ in Eq. (6),

$$\sum_{i=1}^{d} \boldsymbol{p}_i = \sum_{i=1}^{d} \left( \sum_{j=1}^{d} \kappa_{ij} \boldsymbol{p}_j^k \right) \tag{8}$$

Rearranging Eq. (8) gives

$$\left( \sum_{j=1}^{d} \kappa_{id} \right) \boldsymbol{p}_j^k = \sum_{i-1}^{d} \boldsymbol{p}_i - \sum_{i=1}^{d} \left( \sum_{j=1}^{d-1} \kappa_{ij} \right) \boldsymbol{p}_j^k \tag{9}$$

Eq. (9) can give normalized $\boldsymbol{p}_d^k$.

Eq. (5) can now be expanded to include the $d$-th component.

$$\boldsymbol{s}_i = \sum_{j=1}^{d} \gamma_{ij} \boldsymbol{p}_j^k \tag{10}$$

Inserting Eq. (10) into Eq. (1) gives

$$\mathbf{X} = \sum_{i=1}^{d} \boldsymbol{c}_1 \left( \sum_{j=1}^{d} \gamma_{ij} \left( \boldsymbol{p}_j^k \right)^T \right) = \sum_{i=1}^{d} \sum_{j=1}^{d} \gamma_{ij} \boldsymbol{c}_i \left( \boldsymbol{p}_j^k \right)^T \tag{11}$$

Multiplying Eq. (11) by $\boldsymbol{p}_d^k$ gives

$$\mathbf{X} \boldsymbol{p}_d^k = \sum_{i=1}^{d} \gamma_{id} \boldsymbol{c}_i \tag{12}$$

If $i \neq k$, then $\boldsymbol{s}_i$ is linear combination of $\boldsymbol{p}_j$ ($j = 1, 2, \ldots, d-1$) and it is easy to know from Eq. (10) that $\gamma_{ik} = 0$ for all $i$ except $i = k$. Then Eq. (11) becomes

$$\mathbf{X} \boldsymbol{p}_d^k = \gamma_{kk} \boldsymbol{c}_k \tag{13}$$

Let

$$\boldsymbol{c}_k * = \mathbf{X} \boldsymbol{p}_d^k / \| \mathbf{X} \boldsymbol{p}_d^k \| \tag{14}$$

$\boldsymbol{c}_k^*$ is the normalized concentration profile vector of component $k$.

## 4. Orthogonal projection resolution

In the same way as for WFA, let $\mathbf{X}^k$ be a submatrix of **X** that is constructed by deleting the elution 'window' of component $k$. Principal component analysis of submatrix $\mathbf{X}^k$ yields a matrix $\mathbf{P}^k$ containing $d-1$ orthonormal abstract spectral vectors $\boldsymbol{p}_i^k$ ($i = 1, 2, \ldots, d-1$) and a matrix $\mathbf{T}^k$ containing $d-1$ abstract concentration profile vectors $\boldsymbol{t}_i^k$ ($i = 1, 2, \ldots, d-1$). Let be $\boldsymbol{p}_d^k$ the vector produced by Eq. (9). Including $\boldsymbol{p}_d^k$ into $d-1$ orthonormal abstract spectral vectors $\boldsymbol{p}_i^k$ ($i = 1, 2, \ldots, d-1$) described above, then the space spanned by $\{\boldsymbol{p}_j^k,\ j = 1, 2, \ldots, d\}$ is the same space as that spanned by the spectral vectors of $d$ components. Since the space spanned by the spectral vectors of $d$ components is a subspace of $n$-dimensional space (wavelength space), thus $\{\boldsymbol{p}_j^k,\ j = 1, 2, \ldots, d\}$ is the basis of $d$-dimensional subspace of $n$-dimension space.

We can expand $\{\boldsymbol{p}_j^k,\ j = 1, 2, \ldots, d\}$ into the basis of $n$-dimensional space by finding $n$–$d$ vectors $\boldsymbol{p}_{d+1}^k, \boldsymbol{p}_{d+2}^k, \ldots, \boldsymbol{p}_n^k$, which are orthonormal to these $d$ vectors and themselves are mutually orthogonal. Let

$$\begin{aligned} \mathbf{P}^n &= \left( \boldsymbol{p}_1^k \boldsymbol{p}_2^k \cdots \boldsymbol{p}_{d-1}^k \boldsymbol{p}_d^k \boldsymbol{p}_{d+1}^k \cdots \boldsymbol{p}_n^k \right) \\ &= \left( \mathbf{P}^k \vdots \boldsymbol{p}_d^k \vdots \mathbf{P}^\wedge \right) \end{aligned} \tag{15}$$

Where $\mathbf{P}^k = (\boldsymbol{p}_1^k, \boldsymbol{p}_2^k, \ldots, \boldsymbol{p}_{d-1}^k)$, $\mathbf{P}^\wedge = (\boldsymbol{p}_{d+1}^k, \boldsymbol{p}_{d+2}^k, \ldots, \boldsymbol{p}_n^k)$. The columns of the matrix $\boldsymbol{p}^\wedge$ span the ($n$–$d$)-dimensional subspace that is perpendicular to the subspace spanned by the spectral vectors of $d$ components (or by the vectors $\boldsymbol{p}_j^k,\ j = 1, 2, \ldots, d$). Obviously, $\mathbf{P}^n$ is an $n \times n$ orthonormal matrix. That is, $\mathbf{I} = (\mathbf{P}^n)^T \mathbf{P}^n = \mathbf{P}^n (\mathbf{P}^n)^T$. So

$$\begin{aligned} \mathbf{I} &= \mathbf{P}^n (\mathbf{P}^n)^T = \left( \mathbf{P}^k \vdots \boldsymbol{p}_d^k \vdots \mathbf{P}^\wedge \right) \left( \mathbf{P}^k \vdots \boldsymbol{p}_d^k \vdots \mathbf{P}^\wedge \right)^T \\ &= \mathbf{P}^k (\mathbf{P}^k)^T + \boldsymbol{p}_d^k \left( \boldsymbol{p}_d^k \right) + \mathbf{P}^\wedge (\mathbf{P}^\wedge)^T \end{aligned} \tag{16}$$

The projection matrix relating $\mathbf{P}^k$ is

$$\mathbf{M}_k = \mathbf{I} - \mathbf{P}^k (\mathbf{P}^k)^T \tag{17}$$

Inserting Eq. (16) into Eq. (17) gives

$$\mathbf{M}_k = \boldsymbol{p}_d^k \left( \boldsymbol{p}_d^k \right)^T + \mathbf{P}^\wedge (\mathbf{P}^\wedge)^T \tag{18}$$

Let $\boldsymbol{x}_j^T$ be the $j$-th row of matrix **X**. Projecting $\boldsymbol{x}_j^T$ to the subspace which is perpendicular to the subspace spanned by the vectors $\{\boldsymbol{p}_j^k,\ j = 1, 2, \ldots, d-1\}$ (multiplying $\boldsymbol{x}_j^T$ by $\mathbf{M}_k$) gives

$$\boldsymbol{x}_j^T \mathbf{M}_k = \left( \sum_{i=1}^{d} c_{ji} \boldsymbol{s}_i^T \right) \left( \boldsymbol{p}_d^k \left( \boldsymbol{p}_d^k \right)^T + \mathbf{P}^\wedge (\mathbf{P}^\wedge)^T \right) \tag{19}$$

Since $\{\boldsymbol{s}_i,\ i=1,\ 2,\ \ldots,\ d\}$ are the vectors in the subspace spanned by $\{\boldsymbol{p}_j^k,\ j=1,\ 2,\ \ldots,\ d\}$, the vectors $\{\boldsymbol{s}_i,\ \mathrm{i}=1,\ 2,\ \ldots,\ d\}$ are perpendicular to $\{\boldsymbol{p}_j^k,\ j=d+1,\ d+2,\ \ldots,\ n\}$. Thus, Eq. (19) can be written as

$$\begin{aligned}\boldsymbol{x}_j^T\mathbf{M}_k &= \left(\sum_{i=1}^{d} c_{ji}\boldsymbol{s}_i^T\right)\boldsymbol{p}_d^k\left(\boldsymbol{p}_d^k\right)^T \\ &= \boldsymbol{x}_j^T\boldsymbol{p}_d^k\left(\boldsymbol{p}_d^k\right)^T\end{aligned} \tag{20}$$

The norm of $\boldsymbol{x}_j^T\mathbf{M}_k$ is

$$\begin{aligned}\|\boldsymbol{x}_j^T\mathbf{M}_k\| &= \left(\boldsymbol{x}_j^T\boldsymbol{p}_d^k\left(\boldsymbol{p}_d^k\right)^T\boldsymbol{p}_d^k\left(\boldsymbol{p}_d^k\right)^T x_j\right)^{1/2} \\ &= \left(\boldsymbol{x}_j^T\boldsymbol{p}_d^k\left(\boldsymbol{p}_d^k\right)^T\boldsymbol{x}_j\right)^{1/2} \\ &= \boldsymbol{x}_j^T\boldsymbol{p}_d^k\end{aligned} \tag{21}$$

Rearranging Eq. (20) gives

$$\boldsymbol{x}_j^T\mathbf{M}_k = \left(\sum_{i=1,i\neq k}^{d} c_{ji}\boldsymbol{s}_i^T\right)\boldsymbol{p}_d^k\left(\boldsymbol{p}_d^k\right)^T + c_{jk}\boldsymbol{s}_k^T\boldsymbol{p}_d^k\left(\boldsymbol{p}_d^k\right)^T \tag{22}$$

Because vectors $\{\boldsymbol{s}_i,\ i=1,\ 2,\ \ldots,\ d-1\}$ are in the subspace spanned by $\{\boldsymbol{p}_j^k,\ j=1,\ 2,\ \ldots,\ d-1\}$, they are perpendicular to vector $\boldsymbol{p}_d^k$. Therefore, Eq. (22) becomes

$$\boldsymbol{x}_j^T\mathbf{M}_k = c_{jk}\boldsymbol{s}_k^T\boldsymbol{p}_d^k\left(\boldsymbol{p}_d^k\right)^T \tag{23}$$

The norm of $\boldsymbol{x}_j^T\mathbf{M}_k$ can also be written as

$$\begin{aligned}\|\boldsymbol{x}_j^T\mathbf{M}_k\| &= \left(c_{jk}\boldsymbol{s}_k^T\boldsymbol{p}_d^k\left(\boldsymbol{p}_d^k\right)^T\boldsymbol{p}_d^k\left(\boldsymbol{p}_d^k\right)^T\boldsymbol{s}_k c_{jk}\right)^{1/2} \\ &= c_{jk}\boldsymbol{s}_k^T\boldsymbol{p}_d^k\end{aligned} \tag{24}$$

The procedure can be repeated for all the mixture spectra vectors $\{\boldsymbol{x}_j^T,\ j=1,2,\ \ldots,\ m\}$. In this way, we obtain a concentration profile which is proportional to the profile of pure component $k$. This can be done simply by combining Eq. (23) with Eq. (24) in the vector form:

$$\begin{aligned}\boldsymbol{h} &= \left(\|\boldsymbol{x}_1^T\mathbf{M}_k\|,\|\boldsymbol{x}_2^T\mathbf{M}_k\|,\ldots,\|\boldsymbol{x}_m^T\mathbf{M}_k\|\right)^T \\ &= \left(\boldsymbol{x}_1^T\boldsymbol{p}_d^k,\boldsymbol{x}_2^T\boldsymbol{p}_d^k,\ldots,\boldsymbol{x}_m^T\boldsymbol{p}_d^k\right)^T \\ &= \left(c_{1k}\boldsymbol{s}_k^T\boldsymbol{p}_d^k,c_{2k}\boldsymbol{s}_k^T\boldsymbol{p}_d^k,\ldots,c_{mk}\boldsymbol{s}_k^T\boldsymbol{p}_d^k\right) \\ &= \left(c_{1k},c_{2k},\ldots,c_{mk}\right)^T\boldsymbol{s}_k^T\boldsymbol{p}_d^k \\ &= \boldsymbol{c}_k\left(\boldsymbol{s}_k^T\boldsymbol{p}_d^k\right)\end{aligned} \tag{25}$$

Normalizing $\boldsymbol{h}$ gives

$$\boldsymbol{c}_k* = \boldsymbol{h}/\|\boldsymbol{h}\| = \mathbf{X}\boldsymbol{p}_d^k/\|\mathbf{X}\boldsymbol{p}_d^k\| = \boldsymbol{c}_k/\|\boldsymbol{c}_k\| \tag{26}$$

$\boldsymbol{c}_k^*$ is the normalized concentration profile vector of component $k$. Eq. (26) is the same as Eq. (14).

## 5. Conclusion

Although WFA and OPR get the resolution by different ways, from the equations derived in Section 2, if the elution window of a component is same, the normalized concentration profile vectors of this component gained by the two procedures are the same (Eq. (22) is the same as Eq. (14)). Therefore window factor analysis and orthogonal projection resolution are equivalent methods for resolution of two-way data from hyphenated evolutionary processes.

## Acknowledgements

This work was financially supported by the National Natural Science Foundation of People's Republic of China and Fok Ying Tung Foundation of Education Commission of People's Republic of China. The authors also thank Prof. Rolf Manne of Bergen University in Norway for his valuable comments.

## References

[1] E.R. Malinowski, Window factor analysis: theoretical derivation and application to flow injection analysis data, J. Chemom. 6 (1992) 29–40.

[2] A. Lorber, Error propagation and figures of merit for quantification by solving matrix equations, Anal. Chem. 58 (1986) 1167–1172.

[3] Y. Liang, O.M. Kvalheim, Diagnosis and resolution of multiwavelength chromatograms by rank map, orthogonal projections and sequential rank analysis, Anal. Chim. Acta 66 (1994) 43–51.

[4] E.R. Malinowski, Factor analysis in Chemistry, 2nd edn., Wiley, New York, 1989.