

Comments on “Near-Infrared Hyperspectral Unmixing Based on a Minimum Volume Criterion for Fast and Accurate Chemometric Characterization of Counterfeit Tablets”

Róbert Rajkó*

Department of Process Engineering, Faculty of Engineering, University of Szeged, P.O. Box 433, H-6701 Szeged, Hungary

Very recently Lopes et al.¹ have published a paper on chemometric characterization of counterfeit tablets unmixing near-infrared hyperspectral data by using simplex identification via split augmented Lagrangian (SISAL), minimum volume simplex analysis (MVSA), and minimum-volume enclosing simplex (MVES) methods. They concluded that the three minimum volume criterion based methods outperform the popular and widely used multivariate curve resolution-alternating least-squares (MCR-ALS).

First of all, for data sets without pure pixels, they cited only Craig² as the researcher who suggested finding the vertices of the minimum volume simplex containing the data. However, 3 years before Craig's paper, Perczel et al.³ published a paper introducing convex constraint analysis (CCA) as a natural deconvolution of circular dichroism (CD) curves of proteins. It would have been more natural to cite papers of Perczel et al.,^{3,4} instead of one outside of chemistry. CCA can be downloaded from the Internet (CD spectrum analyzer program CCA+, <http://www.chem.elte.hu/departments/protnmr/cca/>).

Correspondingly, Henry's paper⁵ (and some references therein) on multivariate receptor modeling by n -dimensional edge detection should have been treated and cited as well. It is welcome if one scientific area fertilizes other scientific areas with its own algorithms and methods (that is the engine of the multidisciplinary); however, during introducing such external algorithms and methods, the technical terms and background history should have been smoothly transformed to the technical terms and background history used in the fertilized scientific area. Note that Lopes et al.¹ demonstrated their methods to identify counterfeit tablets using near infrared (NIR) spectroscopy (i.e., analytical chemical problem published in analytical chemistry journal which is not devoted at all to the remote sensing (geosciences) area).

Second, and most harmfully, Lopes et al. stated the following¹ “The main disadvantage of MCR-ALS is the so-called rotational ambiguity problem, i.e., a set of simplices with different orienta-

tions, all enclosing the data points, are minimizers of the least-squares criterion.”

The rotational ambiguity is not the belongings of MCR-ALS (or any other/related algorithms)! It is an embedded property of the bilinear data matrix \mathbf{Y} , which can be decomposed in the usual way as follows:

$$\mathbf{Y} = \mathbf{M} \mathbf{A} = \sum_{i=1}^p \mathbf{m}_i \mathbf{a}_i \quad (1)$$

where \mathbf{M} is an $l \times p$ matrix of the estimated component spectra; \mathbf{A} is a $p \times n$ matrix of the estimated composition profiles; and \mathbf{m}_i and \mathbf{a}_i are the i th column and row (i.e., profiles for the i th component) of the \mathbf{M} and \mathbf{A} , respectively. The data matrix \mathbf{Y} can be decomposed to orthogonal product matrices by singular value decomposition (SVD) or principal component analysis (PCA):⁶

$$\mathbf{Y} = (\mathbf{U} \mathbf{D}) \mathbf{V}^T = \mathbf{X} \mathbf{V}^T \quad (2)$$

where \mathbf{U} and \mathbf{V} are the left and right eigenvectors of \mathbf{Y} , respectively, \mathbf{D} is the diagonal matrix of the singular values. In terms of the PCA, $\mathbf{U}\mathbf{D} = \mathbf{X}$ is the score matrix and \mathbf{V} is the loading matrix. However, these decompositions are ambiguous because production of the decomposition with any transformation matrix \mathbf{T} and its inverse can yield a new solution with the same optimal properties:⁶

$$\mathbf{Y} = [\mathbf{X} \mathbf{T}] [\mathbf{T}^{-1} \mathbf{V}^T] = \tilde{\mathbf{X}} \tilde{\mathbf{V}}^T \quad (3)$$

Thus the rotational ambiguity depends on the possible amount of different transformation matrix \mathbf{T} . The feasible set of \mathbf{T} can be reduced using constraints.^{7,8} Constraints usually applied are non-negativity, unimodality, closure, and some type of equality constraints based on local rank and on selectivity from previously known information. Without any constraints, the points in the whole abstract subspace composing a simplex containing the data points according to the columns or rows of \mathbf{Y} are feasible. The

* To whom correspondence should be addressed. Phone: (+36 62) 546030. Fax: (+36 62) 546549. E-mail: rajko@sol.cc.u-szeged.hu.

(1) Lopes, M. B.; Wolff, J.-C.; Bioucas-Dias, J. M.; Figueiredo, M. A. T. *Anal. Chem.* **2010**, *82*, 1462–1469.

(2) Craig, M. D. *IEEE Trans. Geosci. Remote Sens.* **1994**, *32*, 542–552.

(3) Perczel, A.; Hollósi, M.; Tusnády, G.; Fasman, G. D. *Protein Eng.* **1991**, *4*, 669–679.

(4) Perczel, A.; Park, K.; Fasman, G. D. *Anal. Biochem.* **1992**, *203*, 83–93.

(5) Henry, R. C. *Chemom. Intell. Lab. Syst.* **2003**, *65*, 179–189.

(6) Malinowski, E. R. *Factor Analysis in Chemistry*, 3rd ed.; Wiley: New York, 2002.

(7) Tauler, R.; Smilde, A.; Kowalski, B. R. *J. Chemom.* **1995**, *9*, 31–58.

(8) de Juan, A.; Tauler, R. *Crit. Rev. Anal. Chem.* **2006**, *36*, 163–176.

natural constraint is the non-negativity because the concentration, composition, abundance, etc. are always non-negative. Most of the spectra are also non-negative (but, e.g., CD spectra can have negative parts). For two- and three-component systems, Lawton and Sylvestre,⁹ Borgen et al.,^{10,11} Maeder et al.,^{12,13} and Rajkó et al.^{14–19} published seminal papers with clear algebraic and visual (geometric) interpretations using this minimal constraint of non-negativity.

Lopes et al.¹ applied the non-negativity constraints and the necessary condition: “each facet of the simplex contains at least $p - 1$ spectral vectors”. The chemometric meaning of this condition is, there should exist at least $p - 1$ spectral bands at which only one component has zero absorbance (signal) for all components (endmembers). This strict condition is rather arbitrary and can be unattainable. On the other hand MCR-ALS is so flexible that it can be extended with the constraint getting the minimal volume simplex and with assuming the condition.

For illustration that the concepts used by Lopes et al. are not always proper, i.e., the methods used by them are not always optimal, let us use the data published by Henry and Kim (Table 1 in their paper)²⁰ and the details can be found in the Supporting Information. On the basis of the results, the main conclusion can be drawn, i.e., there exists no general algorithm which is the best (optimal) for all data sets, in other words we have to be careful to use the proper constraints and conditions.

To be *unique* or not to be *unique*, that is the *following* question! There is a deep confusion in the term of unique solution. Programmers and mathematicians have been and are developing algorithms providing “unique solution”. However, most of these unique solutions are only mathematical artifacts. Most of the algorithms can provide only one solution at the end of the iterations; however, a large set of equally optimal solutions (not revealed by the used algorithms) can exist. PCA is the most adequate example for chemometrics. Mathematically (algorithmically/numerically) PCA can provide a “unique solution” because of the used restrictions of orthogonality, normalization, sign (e.g., only non-negative components of the first eigenvector), and maximum variance. However, there exist infinite equally optimal solutions based on eq 3. Of course, in general, the spectra are not orthogonal and the concentration profiles are not orthogonal to each other as well, e.g., because of the non-negativity properties (for the orthogonality between two non-negative vectors, some suitable sparseness should be fulfilled), thus the abstract solution of PCA is just a mathematical artifact in general. Chemically/physically interpretable solutions should/can be obtained using further constraints (i.e.,

constructing a proper transformation matrix **T** in eq 3) taking into consideration of the investigated chemical systems.^{7,8} It should be noted that Manne²¹ postulated three theorems on unique solution based on local rank information for chromatographic data. It is worth mentioning that mathematically and chemically/physically unique solutions can be obtained by using some multilinear methods/restrictions.^{7,22}

It should be mentioned here that both SISAL and MVSA suffer from the nonuniqueness because of the used lambda parameter: for different lambdas, different results can be obtained even if the strict condition mentioned above is fulfilled. Every different result is “unique” for a special lambda; however, the result is not always equal to the true one of course and it even may not lie in the feasible region. Moreover it can be easily constructed such that a data set for which the condition is fulfilled but not uniquely, i.e., there exist several minimal volume simplices. For example, let us consider any point-pairs at the sides of a regular hexagon for a three-component system, there will be two equal and minimal area triangles for using the points of every second side.

Another critical problem is the interpretation of the normalization. In nature, the data are not normalized at all in general. The normalization may be needed for the used algorithms or for easier interpretation. Thus, the proper use of eqs 1 and 2 in the paper of Lopes et al.¹ should be as follows. Matrix **M** can be considered as the linear napierian absorption coefficient matrix.²³ Thus the abundance matrix **A** contains the information of penetration depth which is related to the composition. Matrix **A** is not naturally normalized as eq 2 in the paper of Lopes et al.¹ indicates. The penetration depths are not directly related to the abundance fractions valued between 0 and 1. Thus the assumption of 1-norm normalization is arbitrary in this case. Because of the scale ambiguity of the decomposition in eq 1, any concentration profile can be multiplied with a constant if we divide the corresponding absorbance profile with the same constant simultaneously. In this sense, any normalization is possible and adequate; however, we lose the ease of comparison of different measurements and the chemically/physically interpretable solution.

Of course we can normalize the data matrix **Y** using 1-norm for the columns:¹⁶

$$\bar{\mathbf{Y}}_C = \mathbf{Y}_{b \times n} / (\mathbf{1}_{b \times 1} \mathbf{1}_{1 \times l}^T \mathbf{Y}) = \bar{\mathbf{M}}_{b \times p} \bar{\mathbf{W}}_{p \times n} \quad (4)$$

where

$$\bar{\mathbf{M}}_{b \times p} = \mathbf{M}_{b \times p} / (\mathbf{1}_{b \times 1} \mathbf{1}_{1 \times l}^T \mathbf{M}), \quad \bar{\mathbf{W}}_{p \times n} = \mathbf{A}^*_{p \times n} / (\mathbf{1}_{p \times 1} \mathbf{1}_{1 \times n}^T \mathbf{A}^*),$$

$$\mathbf{A}^*_{p \times n} = \mathbf{A}_{p \times n} \cdot (\mathbf{M}_{p \times l}^T \mathbf{1}_{l \times 1} \mathbf{1}_{1 \times n}^T)$$

(9) Lawton, W. H.; Sylvestre, E. A. *Technometrics* **1971**, *13*, 617–633.

(10) Borgen, O. S.; Kowalski, B. R. *Anal. Chim. Acta* **1985**, *174*, 1–26.

(11) Borgen, O. S.; Davidsen, N.; Mingyang, Z.; Øyen, Ø. *Microchim. Acta* **1986**, *II*, 63–73.

(12) Vosough, M.; Mason, C.; Tauler, R.; Jalali-Heravi, M.; Maeder, M. *J. Chemom.* **2006**, *20*, 302–310.

(13) Abdollahi, A.; Maeder, M.; Tauler, R. *Anal. Chem.* **2009**, *81*, 2115–2122.

(14) Rajkó, R.; István, K. *J. Chemom.* **2005**, *19*, 1–16.

(15) Rajkó, R. *J. Chemom.* **2006**, *20*, 164–169.

(16) Rajkó, R. *J. Chemom.* **2009**, *23*, 265–274.

(17) Rajkó, R. *Anal. Chim. Acta* **2009**, *645*, 18–24.

(18) Rajkó, R.; Nassab, P. R.; Szabó-Révész, P. *Talanta* **2009**, *79*, 268–274.

(19) Rajkó, R. *Anal. Chim. Acta* **2010**, *661*, 129–132.

(20) Henry, R. C.; Kim, B. M. *Chemom. Intell. Lab. Syst.* **1990**, *8*, 205–216.

(21) Manne, R. *Chemom. Intell. Lab. Syst.* **1995**, *27*, 89–94.

(22) Smilde, A.; Bro, R.; Geladi, P. *Multi-Way Analysis: Applications in Chemical Sciences*; Wiley: Chichester, U.K., 2004.

(23) Chalmers, J.; Griffiths, P., Eds.; *Handbook of Vibrational Spectroscopy*; Wiley: Chichester, U.K., 2002.

and ./ and .* means the element-wise matrix division and element-by-element product of the matrices, respectively, according to Matlab²⁴ concepts. Now $\bar{\mathbf{W}}_{p \times p}^T \mathbf{1}_{p \times 1} = \mathbf{1}_{n \times 1}$ however, this normalized abundance matrix depends on the spectral matrix as well.

And in the end, most of the results presented in the paper of Lopes et al.¹ have already been published earlier elsewhere,²⁵ a fact that was not indicated.

(24) The Mathworks Inc. *MATLAB, version 7.2 (R2006a), User's Guide*; The Mathworks Inc.: Natick, MA, 2006.

SUPPORTING INFORMATION AVAILABLE

Additional information as noted in text. This material is available free of charge via the Internet at <http://pubs.acs.org>.

Received for review July 24, 2010. Accepted August 23, 2010.

AC101962X

(25) Lopes, M. B.; Bioucas-Dias, J. M.; Figueiredo, M. A. T.; Wolff, J.-C. First Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, WHISPERS '09, 2009; DOI: 10.1109/WHISPERS.2009.5289081.