Analytica Chimica Acta, 146 (1983) 9—14
Elsevier Scientific Publishing Company, Amsterdam — Printed in The Netherlands

THREE-DIMENSIONAL RANK ANNIHILATION FOR MULTI-COMPONENT DETERMINATIONS

C. J. APPELLOF and E. R. DAVIDSON*
Department of Chemistry, University of Washington, Seattle, WA 98195 (U.S.A.)
(Received 26th April 1982)

SUMMARY

The method of rank annihilation for multi-component determinations is extended to a three-dimensional data array. The possibility of improved sensitivity over the two-dimensional method is shown. An illustration using data of the type expected from a liquid chromatograph with a video-fluorimeter as detector is presented. Partial separation of the sample by the liquid chromatograph leads to improved sensitivity in the data analysis of the excitation or emission covariance matrices.

The technique of two-dimensional (2-D) rank annihilation has been used for the determination of fluorescent compounds in mixtures [1, 2]. The major advantage over other data-processing techniques is that a single compound in a mixture can be quantified without knowing what other compounds are in the mixture. Sensitivity criteria have been developed which indicate that the accuracy of determination of a compound concentration using 2-D rank annihilation depends on the spectral overlap of that compound with the other components of the solution [2]. In this paper, we show that by using three-dimensional (3-D) data, better accuracy should be obtained and possible spectral overlap problems avoided. The method is applied to data of the type expected from liquid chromatographic separation with video-fluorimetric detection (v.f./l.c.), but should apply equally well to other types of data.

THEORY

In 2-D rank annihilation [1, 2], the data processed is a two-dimensional emission—excitation matrix (EEM), **M**, the *ij* member of which is the fluorescence emission intensity of a solution at wavelength λ_i when excited by light of wavelength λ_j. For a dilute solution of r components, the EEM is a linear combination of standard EEM matrices.

$$M_{ij} = \sum_{p=1}^{r} \alpha_p x_{ip} y_{jp} \tag{1}$$

where x_{ip} is the normalized, digitized emission spectrum for component p, y_{jp} is the normalized excitation spectrum, and the relative intensity α_p is proportional to the concentration. The rank of the matrix M should be equal to the number of independent components in the sample. Other sources of two-dimensional data, such as gas chromatography/mass spectrometry, also produce data of this form, and are amenable to the same data-processing schemes.

The determination of one compound in a multi-component solution by the method of rank annihilation proceeds by subtraction of a multiple of the standard EEM matrix N, for the pure solution, from the EEM matrix of the mixed solution to form a residual matrix $E = M - \beta N$. Then, the covariance matrix, $D = E^T E$, of E is formed and eigenvectors are established. When the proper multiple (β) of the standard matrix of the analyte has been subtracted from the mixture matrix, the lowest non-zero eigenvalue, ξ^2, of D should become zero. In practice, because of errors in the data, this eigenvalue does not become exactly zero, but does reach a minimum. At that point, the multiplier β of the standard matrix is equal to the concentration of this compound in the mixture, relative to its concentration in the standard solution.

As shown earlier [2], the accuracy of this determination is proportional to the fluorescence intensity of the analyte at unit concentration (α_1^0), and to the uniqueness of the analyte relative to the other components in the mixture. The uniqueness of the emission spectrum of the first compound in the mixture is defined [2] as

$$q_x = (1 - s^T S^{-1} s) \tag{2}$$

where $s_p = \Sigma_i x_{ip} x_{i1} (p > 1)$ and $S_{pq} = \Sigma_i x_{ip} x_{iq}$ $(p, q > 1)$. The uniqueness of the excitation spectrum, q_y, is similarly defined. The uniqueness of either spectrum depends on the overlap of the analyte spectrum with the spectra of the other compounds in the mixture. If the spectrum of the analyte is very similar to a superposition of the other spectra, the uniqueness will be close to zero; if the spectrum is very different, the uniqueness will be close to one.

The error in the estimated concentration caused by errors in the measured data matrix is

$$\sigma^2(\beta) = (\partial^2 \xi^2 / \partial \beta^2)^{-1} \sigma^2 (M) \tag{3}$$

if the error in the data matrix does not depend on wavelength [3]. At the minimum of ξ^2,

$$\partial^2 \xi^2 / \partial \beta^2 = 2 (\alpha_1^0)^2 q_x q_y \tag{4}$$

so the sensitivity of the analysis by rank annihilation, which is defined here as $\partial^2 \xi^2 / \partial \beta^2$, is proportional to $q_x q_y$.

This paper presents an extension of this two-dimensional data-processing technique to the problem of processing a three-dimensional array of data

obtained, for example, from liquid chromatographic separation with video-fluorimetric detection (v.f./l.c.) [4]. These data are in the form of an array M_{ijk}, which measures the fluorescence emission intensity at wavelength λ_i when the sample collected at elution time t_k is excited by light with wavelength λ_j. For dilute solutions, this array, like its two-dimensional counterpart, is a linear combination of products of one-dimensional spectra:

$$M_{ijk} = \sum_{p=1}^{r} \alpha_p x_{ip} y_{jp} z_{kp} \tag{5}$$

where z_{kp} is the h.p.l.c. profile of compound p.

To start the calculation, the three covariance matrices $C(x)$, $C(y)$, and $C(z)$ are eigen-analyzed, where

$$C_{ii'}(x) = \sum_{jk} M_{ijk} M_{i'jk}$$

$$C_{jj'}(y) = \sum_{ik} M_{ijk} M_{ij'k} \tag{6}$$

$$C_{kk'}(z) = \sum_{ij} M_{ijk} M_{ijk'}$$

For error-free data, the number of non-zero eigenvalues of each matrix should be equal to the number of components in the sample. In practice, because of errors in the data, an estimate must be made of the number of eigenvalues which are large compared to zero. This number is then used as an estimate of the number of detectable components in the sample. Then, a procedure analogous to 2-D rank annihilation is followed. Some of a standard 3-D data array N for one component is subtracted from the mixture array: $E_{ijk} = M_{ijk} - \beta N_{ijk}$. Then one of the covariance matrices of E is formed and eigen-analyzed. The value of β which makes the lowest non-zero eigenvalue a minimum is the concentration of this component in the sample relative to its concentration in the standard solution. To reduce the calculations involved, the original data array and the standard array can be transformed to a reduced space as described earlier [1, 4].

As before [2], the accuracy of the calculated concentration will depend on how sensitive the lowest eigenvalue is to changes in concentration near the minimum. In two dimensions, this depends on the intensity of the standard at unit concentration and the uniqueness of the standard emission and excitation spectra (Eqn. 3). For three-dimensional data, there are three covariance matrix eigenvalues that could be examined, and these give three different sensitivity criteria:

$$\partial^2 \xi^2(x)/\partial \beta^2 = 2(\alpha_1^0)^2 \, q_x q_{yz} \tag{7}$$

$$\partial^2 \xi^2(y)/\partial \beta^2 = 2(\alpha_1^0)^2 \, q_y q_{xz} \tag{8}$$

$$\partial^2 \xi^2(z)/\partial \beta^2 = 2(\alpha_1^0)^2 \, q_z q_{xz} \tag{9}$$

The q's with a single subscript are the same uniqueness indices which are cal-

culated for 2-D data, and just depend on the overlap of the standard emission, excitation, and chromatograms with those of the other components in solution. However, the doubly subscripted q's are calculated from two-dimensional overlaps. For q_{yz} in Eqn. (7) for instance,

$$q_{yz} = (1 - s^T[yz]\ S^{-1}[yz]\ s[yz]) \tag{10}$$

where $s_p(yz) = (\sum_j y_{jp} y_{j1})(\sum_k z_{kp} z_{k1})$

and $S_{pq}(yz) = (\sum_j y_{jp} y_{jq})(\sum_k z_{kp} z_{kq})$ $(p, q > 1)$

The factor q_{yz} in Eqn. (7) will always be larger than q_y in Eqn. (4) and q_{xz} in Eqn. (8) will be larger than q_x in Eqn. (4), so the use of 3-D processing may provide better sensitivity than a 2-D processing if the signal intensity is the same.

It is usually the case that the lowest eigenvalue of one of the three covariance matrices is less sensitive to errors in the data than the other two. Therefore, the concentration corresponding to the minimum of that eigenvalue is more accurate than that determined by the other two. In extreme cases, when the spectrum of the analyte totally overlaps the spectra of the other components in one of the dimensions (e.g., the emission spectrum), a change in the subtracted concentration will have no effect on the lowest eigenvalue of the corresponding covariance matrix. In that case, one of the other covariance matrices must be used to determine the true concentration of the standard. In this way, a 3-D array gives more flexibility in choosing the most appropriate covariance matrix, and in the extreme case, can provide results when a 2-D procedure would fail completely.

Equation (9) can be evaluated only if all of the components in the solution are known. Because rank annihilation is designed to work without this knowledge, an alternative formula must be used to estimate the sensitivity. For noise-free data, when the minimum of the lowest eigenvalue has been found, the uniqueness can be calculated from the overlap of the eigenvector corresponding to this eigenvalue with the corresponding eigenvector from the standard analyte spectrum. For instance, if there are K components in the mixture, the Kth eigenvalue should reach a minimum; and for a standard with normalized spectral vectors x, y, and z,

$$q_x = \sum_i x_i v_{iK} \tag{11}$$

where $\mathbf{v}_K$ is the Kth eigenvector of the emission covariance matrix of E. To calculate q_{yz}, it is necessary to form the covariance matrix:

$$D_{(jk)(j'k')}(yz) = \sum_i E_{i(jk)} E_{i(j'k')} \tag{12}$$

This matrix has a row index of (jk) and a column index of $(j'k')$. For a K-component mixture, the Kth eigenvalue of this matrix will reach a minimum for the same choice of β. The Kth eigenvector, $\mathbf{w}_K$, can be used to compute

$$q_{yz} = \sum_{jk} y_j z_k w_{(jk)\,K} \tag{13}$$

The uniqueness indices q_y, q_{xz}, q_z, and q_{xy} can be calculated in a similar way.

RESULTS AND DISCUSSION

The excitation and emission spectra for perylene, fluoranthene, tetracene and 9,10-dimethylanthracene were obtained by eigenanalysis of EEM's taken with a video-fluorimeter [2]. Then, h.p.l.c. elution profiles for each compound were simulated as strongly overlapping Gaussian peaks and random noise was added. The noise had a uniform distribution with a maximum magnitude which was 10% of the root-mean-square value of the data array. Table 1 lists the true relative concentrations of the components in the simulated data, as well as the concentrations calculated by rank annihilation.

Table 2 compares the uniqueness indices calculated by Eqn. (2) to those calculated by Eqn. (10). As can be seen, the 3-D indices for excitation and emission are larger than the corresponding 2-D index. Even a partial h.p.l.c. separation into overlapping peaks improves the sensitivity. Table 2 also compares the 3-D uniqueness indices calculated by Eqn. (10) to those calculated by Eqns. (11—13). The difference between these last two columns of Table 2 is caused by the random noise.

When the uniqueness indices in this table are considered, it can be seen, for example, that to obtain the most sensitive determination of fluoranthene, the lowest non-zero eigenvalue of the excitation covariance matrix should be minimized. It is also seen that because of the large overlap put into the h.p.l.c. simulation, the h.p.l.c. covariance matrix is predicted to give the least sensitive results. These predictions are supported by Table 1.

TABLE 1

Relative concentrations of compounds for the simulated data matrix

	True	Calculated	Error
Perylene	5.638	5.691[a]	+0.053
		5.694[b]	+0.056
		5.729[c]	+0.091
Fluoranthene	1.122	1.140[a]	+0.018
		1.177[b]	+0.055
		1.236[c]	+0.114
Tetracene	11.13	11.12[a]	—0.01
		11.10[b]	—0.03
		11.15[c]	+0.02
9,10-Dimethyl-anthracene	22.26	22.27[a]	+0.01
		22.29[b]	+0.03
		22.28[c]	+0.02

[a]Calculated from excitation covariance matrix. [b]Calculated from emission covariance matrix. [c]Calculated from h.p.l.c. covariance matrix.

TABLE 2

Square of uniqueness indices for the simulated data matrix

	2-D	3-D	
		True[a]	Calculated[b]
Perylene	0.0677	0.1827[c]	0.1708[c]
		0.1962[d]	0.2107[d]
		0.2251[e]	0.2098[e]
Fluoranthene	0.0750	0.2915[c]	0.2775[c]
		0.1016[d]	0.0938[d]
		0.0918[e]	0.0829[e]
Tetracene	0.1274	0.2536[c]	0.2527[c]
		0.3055[d]	0.2718[d]
		0.0729[e]	0.0821[e]
9,10-Dimethyl-anthracene	0.2647	0.5615[c]	0.5527[c]
		0.4257[d]	0.4237[d]
		0.1960[e]	0.2066[e]

[a]Eqns (7)—(9) and Eqn. (10). [b]Eqns. (7)—(9) and (11)—(13). [c]Calculated for excitation covariance matrix. [d]Calculated for emission covariance matrix. [e]Calculated for h.p.l.c. covariance matrix.

The authors thank Leon Hershberger and Chu-Ngi Ho for providing the experimental data. This work was supported by NIH grant number GM-22311.

REFERENCES

1 C. N. Ho, G. D. Christian and E. R. Davidson, Anal. Chem., 50 (1978) 1108.
2 C. N. Ho, G. D. Christian and E. R. Davidson, Anal. Chem., 52 (1980) 1071.
3 I. W. Warner, G. D. Christian and E. R. Davidson, Anal. Chem., 49 (1977) 564.
4 L. W. Hershberger, J. B. Callis and G. D. Christian, Anal. Chem., 51 (1979) 1444.