

MULTIVARIATE RESOLUTION OF RANK-DEFICIENT SPECTROPHOTOMETRIC DATA FROM FIRST-ORDER KINETIC DECOMPOSITION REACTIONS

J. SAURINA, S. HERNÁNDEZ-CASSOU, R. TAUER* AND A. IZQUIERDO-RIDORSA
Departament de Química Analítica, Facultat de Química, Universitat de Barcelona, Diagonal 647, E-08028 Barcelona, Spain

SUMMARY

The effect of a rank deficiency upon curve resolution in simple kinetic reaction-based systems is studied. Firstly, simulated rank-deficient spectrophotometric data of a mixture of two reagents, each one yielding its own reaction product by a first-order kinetic reaction, are analysed. Four different situations are considered according to the differences in the spectral responses of the reaction constituents and to the differences in the rate constants between the two kinetic processes. A variation of the rate constant between runs for a certain kinetic process is also taken into account. Secondly, the resolution of a real rank-deficient data system, corresponding to the study of the pH-dependent decomposition of 1,2-naphthoquinone-4-sulphonate, is investigated. All these studies were carried out using a multivariate curve resolution method based on the alternating least squares optimization of the kinetic and spectral profiles of the species present in the system. © 1998 John Wiley & Sons, Ltd.

KEY WORDS rank deficiency; rank overlap; trilinearity; second-order data; kinetic decomposition reactions; multivariate curve resolution

INTRODUCTION

Multivariate data analysis methods have been proposed to expand the capabilities of kinetic methods for the simultaneous determination of several analytes in unknown mixtures,^{1,2} especially for those cases where full selectivity cannot be achieved and traditional univariate data analysis methods fail. Multivariate data analysis methods proposed for kinetic data can be classified into two main groups. The first group is represented by the so-called hard-modelling data analysis methods,^{3–5} based on least squares curve-fitting regression methods or on Kalman filters. Using these methods, a chemical model including the reaction chemical order, the stoichiometry of the reaction and the rate constant (rate law) is initially proposed. For quantitative analysis, however, hard-modelling methods are too restrictive, because the experimental conditions to keep strictly the constraints imposed by the fulfilment of the initially proposed models cannot be easily achieved and also because in many circumstances the kinetic model is unknown. Usually, these are the methods of choice to investigate the kinetic model and to estimate the kinetic parameters of the chemical reactions.

The second group of methods comprises the so-called soft-modelling data analysis methods. Among them, first-order multivariate calibration methods such as the partial least squares regression

* Correspondence to: R. Tauler, Departament de Química Analítica, Facultat de Química, Universitat de Barcelona, Diagonal 647, E-08028 Barcelona, Spain. E-mail: roma@quimio.qui.ub.es
Contract/grant sponsor: DGICYT (Spain); Contract/grant number: PB96-0377

(PLSR) method and the principal component regression (PCR) method have been proposed.^{6–8} They have been successfully used in the quantitative determination of mixtures without a prior knowledge of the reaction order or the rate constants involved. More recently, artificial neural networks (ANNs) have also been proposed for the quantitative analysis of mixtures, and their performance has been compared with that of first-order calibration methods.^{9,10} Neither for first-order calibration methods nor for ANN methods does a kinetic model have to be initially proposed, and no kinetic restrictions are involved. However, in contrast with hard-modelling methods, neither first-order calibration nor ANN methods are able to recover the qualitative information of the system, i.e. the resolution and recovery of the kinetic and pure spectral profiles of the components in the mixture, and they are mostly applied for quantitative prediction purposes.

Apart from the quantitative analysis, the resolution of a kinetic system includes the determination of the number of chemical species formed during the reaction, the recovery of the concentration and pure spectral profiles of these species and the estimation of the kinetic parameters involved in the reaction, such as the reaction order and the rate constant. The resolution of a kinetic system can only be achieved by methods that take advantage simultaneously of both the kinetic and spectral differences between the analytes. Second-order calibration and resolution methods¹¹ and three-way data analysis methods^{12–15} allow the resolution of a kinetic system without the prior postulation of a kinetic model. In these cases, for each sample, the kinetic data are arranged in a two-way data matrix. For instance, in the case of spectrometric multiwavelength diode array detection the experimental data matrix contains in its rows the spectra measured at preselected successive times during the evolution of the reaction, and in its columns the kinetic profiles measured at preselected successive wavelengths. Second-order multivariate calibration methods are capable of analysing one or several analytes in a sample in the presence of interferences with only the additional input of the response matrix of the desired analyte or of the response matrix of a mixture containing it at a known concentration (second-order advantage¹⁶). Examples of second-order calibration methods are the generalized rank annihilation method (GRAM)^{12,17} and its extension, the direct trilinear decomposition (DTD) method.^{12,14,18} Multivariate calibration and resolution can also be performed using the PARAFAC^{19–21} and restricted Tucker three-way data analysis²² methods and the alternating least squares (ALS) multivariate curve resolution (MCR) method.^{23,24} Whereas GRAM and DTD require the experimental data to be trilinear (i.e. every chemical species must be defined in the different data matrices by the same spectral and concentration profiles (vectors)), the Tucker and curve resolution ALS methods are less restrictive and are suitable for those cases which do not fulfil the trilinearity requirement.¹⁸ Multivariate curve resolution of a data matrix recovers a profile in each order (mode) of the data matrix for each detected species: a kinetic concentration profile (concentration versus time) and a species spectrum (light absorption versus wavelength). Both profiles define the species of the reaction system and identify it.

A data matrix is rank-deficient when the number of significant contributions to the data variance, estimated by using singular value decomposition or other related factor analysis techniques, is lower than the real number of chemical components present in the system. Such a situation occurs quite frequently when, for instance, at the beginning of a chemical reaction or of a chemical process, more than one component already exists. In this case, instead of detecting all the components present at the beginning of the reaction, only one can be mathematically derived, with its spectrum being a linear combination of the pure spectra of all the components initially present. The complete resolution of a such a rank deficient data matrix is therefore hindered.^{24,25} Recently, it has been shown that the rank deficiency problem can be overcome by using different strategies.^{24,25} Two methods have been proposed to break rank deficiency and augment the rank of a rank-deficient data matrix: (i) matrix augmentation by simultaneous analysis of multiple process runs under linearly independent conditions (initial concentrations); (ii) matrix perturbation by addition of a single component or of

mixtures of components during the process. The first of these two strategies is extensively discussed in this work for the resolution of rank-deficient first-order kinetics. Two possible ways of matrix augmentation, row-wise (time kinetic order) and column-wise (wavelength spectral order), will be analysed.

Moreover, as has been previously shown, the investigation of augmented data matrices leads to an improved resolution of the system and eventually to the elimination of the rotational and intensity ambiguities¹⁶ inherently associated with the factor analysis of a two-way data matrix. The intensity ambiguity is always present in factor analysis and curve resolution solutions and it means that the estimated concentrations and spectra can be multiplied (scaled) respectively by an unknown factor and its reciprocal without changing the result. The rotational ambiguities mean that the estimated spectral or concentration profiles recovered in the data analysis are an unknown linear combination of the true ones. These ambiguities occur especially when the experimental data are not selective enough for some or all of the species formed. Whether the final solutions are equal or not to the true ones depends on the selectivity and local rank of the analysed systems.^{16,26}

For the analysis of three-way data sets, trilinearity is an essential property in order to avoid factor analysis ambiguities associated with the resolution of the system. If trilinearity is present and the system is full rank, the estimation of the correct profiles without intensity and rotational ambiguities is always possible. When full trilinearity is not present, the resolution depends on the selectivity and on the local rank of the augmented data matrices.¹⁶ In the analysis of kinetic data the spectral direction can be considered constant between matrices (i.e. between kinetic runs), since the spectral measurements are usually carried out at the same wavelengths and the pure species spectral profiles do not change between kinetic runs. In contrast, the shapes of the pure species kinetic profiles can vary from run to run. Under these circumstances the system lacks a trilinear structure.¹⁸

In this work the effect of a rank deficiency in the resolution of kinetic decomposition reactions is evaluated. Furthermore, the influence of a lack of trilinearity in the study of these rank-deficient kinetic data is also investigated. Firstly, simulated rank-deficient spectrophotometric data of a mixture of two independent first-order kinetic reactions are analysed. The data correspond to the spectrophotometric monitoring of the kinetic processes at preselected time intervals, which yields ordered two-way data matrices. Four different situations are considered according to the presence or not of rank overlap in the concentration profiles and/or in the spectra. There is rank overlap when two chemical species are characterized by the same profile in any of the two orders. The simultaneous analysis of several of these ordered data matrices, which gives three-way data matrices, is also performed. Secondly, a set of real data, obtained in the kinetic decomposition of 1,2-naphthoquinone-4-sulphonate (NQS) at different pH values, is analysed. NQS is a reagent of interest in the analysis of amino acids, since it has been applied to their pre- and post-column derivatization and their determination by means of continuous flow and flow injection methods.^{27–30}

EXPERIMENTAL

Reagents and solutions

All reagents were of analytical grade. Sodium 1,2-naphthoquinone-4-sulphonate (NQS; Aldrich) was used to prepare a 6.7×10^{-4} M aqueous solution, which also contained 0.1 M hydrochloric acid (Merck). From sodium carbonate and sodium hydroxide (both Merck) the following buffer solutions were prepared: solution R₁, Na₂CO₃ 0.1 M + NaOH 0.02 M; solution R₂, Na₂CO₃ 0.1 M + NaOH 0.084 M; solution R₃, Na₂CO₃ 0.1 M + NaOH 0.18 M.

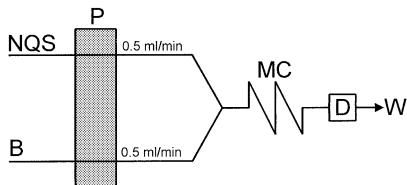


Figure 1. Continuous flow manifold: P, peristaltic pump; MC, mixing coil (length 2 m, i.d. 0.5 mm); D, detector (diode array spectrophotometer); NQS, NQS channel; B, buffer channel; W, waste

Apparatus

Detection was performed using a Hewlett-Packard HP 8452A diode array spectrophotometer with a Hellma flow cell of path length 10 mm and volume 18 μl . Spectra were recorded from 290 to 590 nm, every 2 nm. Data acquisition and generation of ASCII files were done with a Hewlett-Packard Vectra N2 4/50 computer using the standard HP kinetic software.

The pH was measured with an Orion ROSS combined glass electrode connected to a Radiometer PHM84 pH-meter.

Stopped flow methodology

The kinetic decomposition of NQS was studied using the continuous flow manifold shown in Figure 1. All solutions were pumped by means of a peristaltic pump (Scharlau HP4) using standard tygon tubing. NQS and buffer solutions were mixed in a PTFE reaction coil of length 2 m and inner diameter (i.d.) 0.5 mm. The flow rate of each individual channel was around 0.5 ml min^{-1} .

The kinetic decomposition process was monitored by stopping the flow after the system had achieved the steady state and simultaneously starting the measurement process. In this way the flow system used allowed us to control the initial conditions of each kinetic run in a simple and reproducible manner, which assured time synchronization between runs. For each kinetic run, spectra were recorded at the same time intervals (every 10 min during a period of 190 min) and at the same wavelengths. At the end of the chosen time interval the decomposition process was already extensively developed.

Data sets under study

Simulated data sets

Spectrophotometric data sets from the kinetic study of a mixture of two first-order kinetic processes have been simulated. A 1:1 reagent:product stoichiometry has been assumed for both reactions:



In all the simulated data sets a rank deficiency caused by closure is present. Furthermore, four different cases have been considered in these simulations according to the presence or not of rank overlap in the concentration profiles and/or in the spectra. There is rank overlap in the concentration profiles when both kinetic processes have the same rate constant. In this paper it has been considered that there is rank overlap in the spectral profiles when the two reagents have the same spectrum ($\mathbf{s}_A = \mathbf{s}_C$) and so have the two decomposition products ($\mathbf{s}_B = \mathbf{s}_D$). Mixed situations are not studied in the present paper, but the conclusions derived here can be easily adapted to them. The four cases analyzed are the following, according to the causes of the rank deficiencies present in the simulated data.

- Case I. Rank deficiency caused by closure. There is no rank overlap (different species spectra and different rate constants).
- Case II. Rank deficiency caused by closure and by rank overlap in the concentration profiles (different species spectra and equal rate constants).
- Case III. Rank deficiency caused by closure and by rank overlap in the spectral profiles (equal species spectra and different rate constants).
- Case IV. Rank deficiency caused by closure and by rank overlap in the concentration and spectral profiles (equal species spectra and equal rate constants).

It has been assumed that the four species absorb in the whole working wavelength range (no selective wavelengths exist) and that the spectrum of the product is always different from the one of the initial reagent it comes from.

Simulations of kinetic runs of a single kinetic process (either A → B or C → D) have also been studied. These simulations have been performed both with the same rate constants as they have in the simulation of the mixture (trilinear data) and with different ones (non-trilinear data), since in real experiments rate constants can change from run to run owing to variations in the experimental conditions.

For each kinetic run the simulated data are arranged in a matrix **D**, which has a number of columns (n) equal to the number of wavelengths used and a number of rows (m) equal to the number of time intervals.

Matrix **D** was calculated using the Lambert–Beer law in matrix form, as described below:

$$\mathbf{D} = \sum \mathbf{C}_i \mathbf{S}_i^T$$

The dimensions of these matrices are $\mathbf{D}(m \times n)$, $\mathbf{C}_i(m \times 2)$ and $\mathbf{S}_i(n \times 2)$. The superscript T indicates the transposed matrix. Matrix \mathbf{S}_i contains the individual spectra of reagent i and of its reaction product. The spectra used are the first and last spectra of runs A and C of the decomposition study of NQS. Matrix \mathbf{C}_i contains the kinetic concentration profiles of reagent i and of its reaction product. At each time value a closure condition controlled by the mass balance equation has to be fulfilled: the total concentration must be the sum of the concentrations of reagent and of its product. If first-order kinetic processes are assumed, from the known rate constant and for a certain total initial concentration of reagent i , the concentrations of remaining reagent i and of its product at each time value are calculated from

$$\begin{aligned} [\text{reagent } i]_t &= [\text{reagent } i]_0 \exp(-k_i t) \\ [\text{product } i]_t &= [\text{reagent } i]_0 - [\text{reagent } i]_t \end{aligned}$$

In order to check for rank deficiency problems from a theoretical point of view, neither random error nor systematic error was added to this data set.

Kinetic decomposition of NQS

The effect of the pH on the NQS decomposition is studied. Data sets labelled R₁, R₂ and R₃ correspond to the NQS decomposition at pH values of 9.4, 10.5 and 13.3 respectively. The pH was measured at the waste (see Figure 1), once the flow system had achieved the steady state and before the flow was stopped. It corresponded to the pH value of the solution where the decomposition process was to be monitored. This solution was obtained by mixing the corresponding buffer solution with the NQS solution in acidic medium (see Figure 1).

NQS presents an acid–base equilibrium with a pK_a value of 10.7. This equilibrium is schematized in Figure 2 together with the decomposition process for both the acidic and basic forms of the NQS

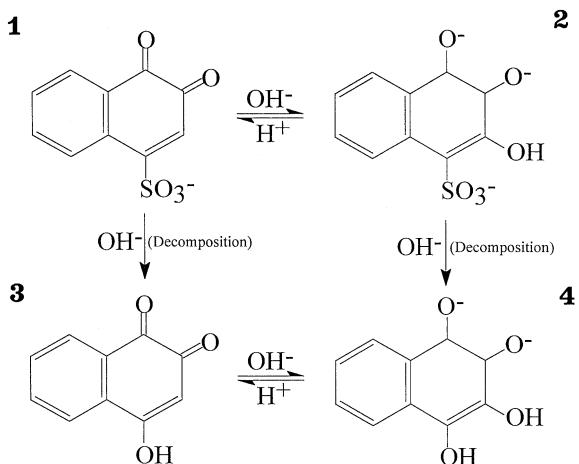


Figure 2. Scheme of proposed NQS species. Species assignment: 1, acidic NQS species; 2, basic NQS species; 3, acidic NQS decomposition product; 4, basic NQS decomposition product

reagent. The chemical structure of the decomposition product has not yet been elucidated, and the one included in Figure 2 is only a probable one.²⁹ Two different situations can be considered according to the pH values at which the decomposition study was performed.

1. pH values 9.4 and 13.3 (data sets R₁ and R₃). In these two cases a single analyte is present at the beginning of the run, the acidic (pH 9.4) and the basic (pH 13.3) form of NQS respectively, which in each case yields a single decomposition product. Both the initial reagent and its decomposition product absorb in the working wavelength range.
2. pH value 10.5. In this case a mixture of the protonated and the deprotonated NQS reagent is initially present in the solution. Both of them absorb in the working wavelength range and so do their decomposition products. A total number of four components is then expected.

The data obtained for each kinetic run consisted of the absorbance values measured between 290 and 590 nm at successive regular time intervals. Spectral data obtained in every run were subjected to a data reduction, taking absorbances in steps of 6 nm.

DATA ANALYSIS

Multivariate curve resolution was applied both to the study of single kinetic runs (i.e. individual data matrices) and to the simultaneous study of multiple kinetic runs (i.e. augmented data matrices). Although more detailed descriptions of the method have been given elsewhere,^{13,31–33} the main steps can be summarized as follows.

Data arrangement

For both the real and simulated data the individual matrices share their row and column matrix indices (absorbance data were obtained at the same time intervals and wavelength values for all the individual kinetic runs). In this way the multivariate curve resolution method can be applied both to the analysis of an individual data matrix and to the simultaneous analysis of several correlated data matrices. In these cases the individual matrices are arranged either in an augmented column-wise matrix (wavelength-wise, keeping wavelengths in common) or in an augmented row-wise matrix (time-wise,

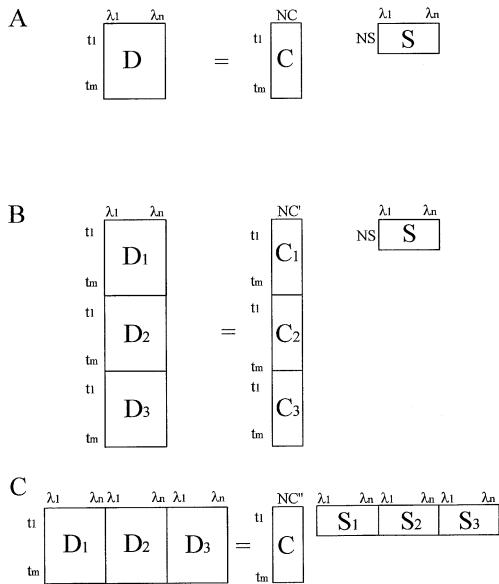


Figure 3. Arrangement of kinetic data in individual and augmented column-wise and row-wise matrices. NC , NC' refer to the numbers of significant factors in each arrangement, which can coincide or not (see text). Augmented matrices are indicated using Matlab notation: $[R_1; R_2; R_3]$ and $[R_1, R_2, R_3]$ refer to column-wise and row-wise augmented matrices respectively

keeping time intervals in common), as shown in Figure 3. The augmented column-wise matrices have a number of rows equal to the total number of acquired spectra in the different kinetic runs (number of time intervals times number of data matrices simultaneously analysed) and a number of columns equal to the number of wavelengths present in each single matrix. On the other hand, the augmented row-wise matrices have a number of rows equal to the number of time intervals present in each single matrix and a number of columns equal to the number of wavelengths times the number of data matrices simultaneously analysed. The notation used in this paper to indicate column-wise or row-wise augmentation is explained in Figure 3.

Rank analysis of individual and augmented data matrices

Rank analysis by inspection of the singular values is used for the determination of the chemical rank. It is assumed that the chemical components have associated singular values much larger than other possible contributions such as instrumental drift or experimental error. In this paper the number of chemical components ('chemical' or 'pseudo' rank of the matrix) is estimated by visual inspection of the plots of singular values.

Since all the individual or augmented data matrices analysed in this paper can be considered as the product of two matrices ($\mathbf{D} = \mathbf{CS}^T$), the rank obtained for these \mathbf{D} matrices can be understood from the following mathematical property of the rank: when a matrix \mathbf{D} can be decomposed into the product of two matrices \mathbf{C} and \mathbf{S}^T , then $\text{rank}(\mathbf{D}) \leq \min(\text{rank}(\mathbf{C}), \text{rank}(\mathbf{S}^T))$. Thus, if one of the two matrices is full rank, it is sufficient to look into the rank of the other for analysing the rank of \mathbf{D} .

In the case of reacting mixtures a rank deficiency appears in the concentration matrix \mathbf{C} , since the concentration profiles of the absorbing species are not linearly independent owing to the underlying

reaction network and to the mass balance equations derived from these reactions (rank deficiency caused by closure). It has been proven for constant-volume batch reaction systems with constant density²⁵ that $\text{rank}(\mathbf{C}) = \min(R+1, Q)$, where R is the number of independent reactions present in the system and Q is the total number of species. This proposition, which is fulfilled in rank-deficient systems consisting of acid–base titrations of mixtures of acids,^{24,34} is also valid for the simulated kinetic systems under study in this paper, which consist of a mixture of two independent first-order kinetic reactions.

A rank deficiency in a data matrix can also be caused by rank overlap. There is rank overlap when different chemical species have common profiles in one of the two orders of measurement. In the presence of rank overlap in the spectral profiles the rank of the \mathbf{D} matrix is equal to the number of independent spectral responses; similarly, in the case of rank overlap in the concentration profiles the rank of the \mathbf{D} matrix is equal to the number of independent concentration responses.

In rank analysis of augmented matrices, another important aspect to take into account is the effect of trilinearity on the rank of the row-wise and column-wise augmented matrices. When the data are trilinear and no rank deficiency is present, the rank of both augmented matrices, row- and column-wise, should be equal and not larger than the total number of species present in the system. In contrast, when the system lacks trilinearity, the ranks of the row- and column-wise augmented data matrices are different and one of them can be larger than the total number of species present in the system.^{16,35} For instance, in the very common case in chemistry of a lack of trilinearity because the concentration profiles (kinetic profiles in the study of kinetic processes) have different shapes in the different runs (different data matrices simultaneously analysed), the rank of the row-wise augmented matrix is larger than the rank of the column-wise augmented matrix and also larger than the total number of species present in the system. Therefore in this case the estimated rank will be ‘in excess’ with respect to the number of species present in the system.

The rank deficiency due to closure can be broken by column-wise matrix augmentation, as has already been demonstrated by Amrhein *et al.*²⁵ However, the effect of both column-wise and row-wise matrix augmentation in rank-deficient systems due to rank overlap has not been extensively analysed yet. It can be demonstrated that rank overlap in the concentration profiles can be successfully broken by column-wise matrix augmentation. On the other hand, rank overlap in the spectral profiles can be broken by row-wise matrix augmentation. Consequently, depending upon the origin of the rank deficiency, either column-wise augmentation or row-wise augmentation should be chosen. In all cases the success of rank augmentation is linked with a good experimental design of which individual matrices should be included in the augmented one.

Initial estimations

Initial estimations are important in multivariate curve resolution using alternating least squares optimization,^{16,24} because they help search for the optimal solutions and avoid the occurrence of local minima. In the present study, spectral estimations were obtained from techniques based on the detection of the purest variables^{36,37} of a data matrix.

Alternating least squares (ALS) optimization

The ALS optimization of the kinetic concentration profiles and of the pure spectra is based on compliance with Beer’s law. The ALS optimization method for MCR is explained in detail elsewhere.^{16,23,29,31} In the analysis of a single matrix the following constraints were applied in order to improve the resolution and to limit the number of possible solutions: (i) the pure spectra of each component have to be non-negative; (ii) the kinetic concentration profiles have to be non-negative; (iii) the system is closed with respect to the total analytical concentration.

Table 1. Rank analysis of individual data matrices for simulated kinetic systems

	Spectral overlap	Concentration overlap	Rank	Number of species	Number of reactions
Case I	No	No	3	4	2
Case II	No	Yes	2	4	2
Case III	Yes	No	2	4	2
Case IV	Yes	Yes	2	4	2

When multiple kinetic runs were simultaneously analysed by MCR-ALS, additional constraints were applied: (iv) the pure spectra of common components in the different kinetic runs are equal; (v) (optional in the case of trilinearity) the pure kinetic profiles of the common components in the different kinetic runs have the same shape; (vi) correspondence of common components between runs; (vii) selectivity constraint—the absent components in a run or in a time interval are forced to have a zero concentration value. Constraint (iv) is fulfilled whenever there are no changes in physical properties such as solvent composition or temperature between runs. In contrast, constraint (v) is only fulfilled for strictly first-order kinetic conditions. Constraint (vi) is fulfilled for the augmented data matrices coming from different kinetic runs which have common and uncommon components. The selectivity constraint within a kinetic run can only be implemented in the analysis of a single reaction system and for the first point of the kinetic concentration profiles, since at this point the reagent is the only species present. However, this constraint is unrealistic in real conditions, since the first experimental spectrum is measured when the reaction has already started and it does not correspond to a pure component spectrum. Accordingly, in the present work no selectivity constraints were applied during the ALS optimization.

RESULTS AND DISCUSSION

Rank analysis

Simulated data sets

Rank analysis of individual data matrices (single kinetic runs) Rank analysis of the individual simulated data matrices was performed on pure data, free from random error. Therefore the rank of the matrices could be unambiguously determined since it was equal to the number of singular values different from zero. All these matrices corresponded to systems of two first-order independent kinetic processes ($A \rightarrow B$ and $C \rightarrow D$), i.e. they consisted of mixtures of two initial reagents (A and C), each one independently yielding its own reaction product (B and D). The rank analysis results obtained are shown in Table 1.

The number of singular values larger than zero or the rank of the individual matrices is in all cases lower than four, which is the total number of species known to be present.

In Cases I and II the four components have different spectra and thus the matrix of spectra, \mathbf{S}^T , is full rank. The rank of the experimental matrix is then related to the rank of the concentration matrix \mathbf{C} . Matrix \mathbf{C} is rank-deficient in both cases, since the number of independent reactions (two) is lower than the number of components in the system (four). In Case I the rank of the \mathbf{C} matrix is three owing to the rank deficiency caused by closure, and thus the \mathbf{D} matrix is rank three. On the other hand, in Case II, apart from that rank deficiency problem, there is rank overlap in the concentration profiles (both kinetic reactions have the same rate constant), which is equivalent to saying that there is only one independent reaction in the system. Consequently, matrix \mathbf{D} is rank two.

Table 2. Rank analysis of column-wise augmented data matrices for simulated kinetic processes: \mathbf{D} , data matrix of system consisting of two kinetic processes (1 and 2); \mathbf{D}_1 , data matrix of system consisting of single kinetic process 1; \mathbf{D}_2 , data matrix of system consisting of single kinetic process 2

	Spectral overlap	Concentration overlap	Trilinearity ^a	[$\mathbf{D}; \mathbf{D}_1$]	[$\mathbf{D}; \mathbf{D}_1; \mathbf{D}_2$]
Case I	No	No	Yes	4	4
			No	4	4
Case II	No	Yes	Yes	4	4
			No	4	4
Case III	Yes	No	Yes	2	2
			No	2	2
Case IV	Yes	Yes	Yes	2	2
			No	2	2

^a Trilinearity is present when the rate constant for a certain kinetic process is invariant between matrices. On the other hand, there is no trilinearity when the rate constant for a certain kinetic process changes between matrices.

In Cases III and IV, not only is the rank deficiency present in matrix \mathbf{C} (which is rank two or three depending on whether there is rank overlap in the concentrations or not), but also there is spectral rank overlap in matrix \mathbf{S}^T . Matrix \mathbf{S}^T is rank two, since there are only two different spectral responses: one equal for both reagents and another equal for both products. Consequently, matrix \mathbf{D} is rank two in both cases.

Therefore in the analysis of an individual data matrix the existence of a rank deficiency caused by closure in all the systems analysed, as well as the presence of rank overlap in some of them, makes impossible the total resolution of the mixture, even in the optimal case (Case I), in which spectral and kinetic differences exist between the two kinetic reactions (there is no rank overlap). In order to break the rank deficiency and obtain full rank systems, matrix augmentation is attempted. The two possible ways to perform matrix augmentation are in the time (row-wise) and wavelength (column-wise) directions.

Rank analysis of augmented data matrices (multiple kinetic runs) The mathematical deduction of the rank of the augmented matrices under closure, rank overlap and trilinearity situations (Cases I–IV) is provided in the Appendix.

Tables 2 and 3 contain the results obtained in the rank analysis of the column-wise and row-wise augmented matrices respectively. All the augmented matrices consisted of the following individual matrices: one corresponding to the mixture of the two kinetic processes (1 and 2) and one or two more corresponding to a single kinetic process. Furthermore, two different situations were studied: (a) the rate constant for a given process was invariant between runs (trilinearity); (b) the rate constant for a given process changed from run to run (lack of trilinearity).

Wavelength-wise (column-wise) augmented matrices (Table 2) Wavelength-wise (column-wise) augmented matrices are decomposed into an augmented \mathbf{C} matrix, which consists of independent matrices of concentration profiles, and an individual \mathbf{S}^T matrix (see Figure 3). Therefore, while the rank of the \mathbf{S}^T matrix is kept constant when this kind of augmentation is performed, the rank of the augmented \mathbf{C} matrix is increased when new information is included in the different individual concentration matrices.

In Cases I and II there are spectral differences between the four components (no spectral overlap) and the \mathbf{S}^T matrix is full rank (rank four). Thus the rank of the augmented \mathbf{D} matrices depends on the

Table 3. Rank analysis of row-wise augmented data matrices for simulated kinetic processes: \mathbf{D} , data matrix of system consisting of two kinetic processes (1 and 2); \mathbf{D}_1 , data matrix of system consisting of single kinetic process 1; \mathbf{D}_2 , data matrix of system consisting of single kinetic process 2

	Spectral overlap	Concentration overlap	Trilinearity ^a	[\mathbf{D}, \mathbf{D}_1]	[$\mathbf{D}, \mathbf{D}_1, \mathbf{D}_2$]
Case I	No	No	Yes	3	3
			No	4	5
Case II	No	Yes	Yes	2	2
			No	3	4
Case III	Yes	No	Yes	3	3
			No	3	4
Case IV	Yes	Yes	Yes	2	2
			No	3	4

^a Trilinearity is present when the rate constant for a certain kinetic process is invariant between matrices. On the other hand, there is no trilinearity when the rate constant for a certain kinetic process changes between matrices.

rank of the augmented \mathbf{C} matrices. Both in the presence and in the absence of trilinearity in the kinetic profiles a full-rank augmented matrix (rank four) is already obtained when an additional data matrix corresponding to the kinetic process of only one of the two reagents in the mixture is included in the augmented one, the other reagent being considered as an unknown interferent (see Appendix).

While in Case I the rank deficiency in the individual matrix of the mixture is only due to the presence of closure, in Case II a rank overlap in the concentrations is also present. However, column-wise matrix augmentation has yielded full-rank matrices in both cases. It can be demonstrated (see Appendix) that rank deficiency caused by rank overlap in the concentration profiles can be successfully broken by column-wise matrix augmentation whenever the different individual matrices simultaneously analysed are not degenerate (i.e. they are not proportional).

In Cases III and IV the \mathbf{S}^T matrix is rank-deficient (rank two), since both kinetic processes have equal spectral responses (spectral overlap). Accordingly, this is the rank-limiting matrix, and the rank of the augmented \mathbf{D} matrix is two independently of the number of individual matrices simultaneously analysed by matrix augmentation (see Appendix). The rank deficiency caused by spectral overlap cannot be broken by column-wise matrix augmentation.

Time-wise (row-wise) augmented matrices (Table 3) Time-wise (row-wise) augmented matrices are decomposed into an augmented \mathbf{S}^T matrix, which consists of independent matrices of spectral profiles, and an individual \mathbf{C} matrix. A mixture of two simultaneous kinetic processes, where each reagent and its product concentrations are linearly dependent owing to the kinetic reaction, gives a rank-deficient \mathbf{C} matrix. Consequently, the rank of the row-wise augmented \mathbf{D} matrix cannot be higher than that of the rank-deficient \mathbf{C} matrix. It can be concluded that when row augmentation is performed, the rank deficiency caused by closure is not broken.

In time-wise augmentation, two species are the same if they have the same kinetic profile. Consequently, with trilinearity between runs the total number of different species is four (expected rank equals four for full-rank matrices). However, for non-trilinear data the expected rank increases to six or eight for full-rank matrices, depending on the number of runs of individual kinetic processes (each individual kinetic process implies the inclusion of two new kinetic profiles, that of the reagent and that of the product) included in the augmented matrix.

The rank values obtained for the time-wise augmented matrices in the different cases are lower than the expected rank values in the absence of rank deficiencies, both in the presence and in the

Table 4. Rank analysis of individual and augmented data matrices for NQS decomposition at different pH values:
 R_1 , kinetic run at pH 9.4; R_2 , kinetic run at pH 10.5; R_3 , kinetic run at pH 13.3

R_1	R_2	R_3	[$R_1; R_2$]	[$R_2; R_3$]	[R_1, R_2]	[R_2, R_3]	[R_1, R_2, R_3]
15.7573	16.4196	14.8729	22.5529	21.8443	22.7459	22.1465	27.1750
0.8854	1.0251	0.5954	3.2878	3.8102	1.4974	1.2899	1.5884
0.0870	0.0575	0.0400	0.4603	0.5605	0.3783	0.2860	0.5350
0.0103	0.0064	0.0259	0.2795	0.4565	0.0345	0.0456	0.1060
0.0044	0.0060	0.0094	0.0723	0.0293	0.0074	0.0407	0.0482
			0.0550	0.0161	0.0065	0.0066	0.0081

absence of trilinearity. The mathematical proof of the rank values obtained for the different row-wise augmented matrices is included in the Appendix. In the cases in which there is no spectral overlap (Cases I and II), the rank of the row-wise augmented matrices is understood by looking at the rank of the related matrices **C**. For trilinear systems the rank of matrix **C** is equal to $R + 1$, where R is the number of independent reactions in the system (i.e. the number of different rate constants in the system). Thus in the presence of trilinearity the rank is three when the rate constants of the two reactions are different (Case I) and two when the rate constants of the two reactions are equal (Case II), independently of the number of individual runs included in the augmented matrix. In contrast, in the absence of trilinearity there is an increase in the rank of one unit for each individual matrix added to the system, although the rank is still lower than that expected in the absence of rank deficiencies.

From all these studies it can be concluded that in the simulated systems analysed in this paper, wavelength-wise augmentation is the best procedure to break the rank deficiencies caused by closure and concentration overlap, whenever the spectral responses are independent for every component present in the system (there is no spectral overlap). On the other hand, time-wise augmentation cannot provide full-rank augmented matrices in any of the analysed cases, since the reaction dependence between the components of the system (closure) takes place in the time direction. Furthermore, the rank of the time-wise augmented matrices is highly affected by the lack of trilinearity in the data. Since the lack of trilinearity is only introduced in the concentration profiles (the spectral profiles have been considered to be always invariant between runs), it does not affect the rank of the wavelength-wise augmented matrices.

Real data sets

Table 4 shows the singular values obtained for each individual kinetic data matrix of the NQS decomposition process. For all of them, only two larger or significant contributions were observed. They could be attributed to the presence of an initial NQS species and of its corresponding decomposition product formed over time. However, according to the acid-base equilibrium of NQS known from previous studies,²⁹ for the kinetic run developed at pH values around pK_a (data set R_2), two initial NQS species (the acidic and the basic one) should be present and each one should suffer its own degradation process (see Figure 2). Therefore for this kinetic system the real number of chemical components present is four, although only two significant contributions are detected by rank analysis. The reagent is a mixture of the acidic and the basic form of NQS, and the decomposition product is also a mixture of an acidic and a basic form. The two forms of the reagent are linked by an acid-base equilibrium and so are those of the decomposition product. Furthermore, since the pH is kept constant along the whole run, the proportion between acidic and basic forms is constrained to be constant along the run for both the reagent and the decomposition product. Consequently, the composition of

the reagent is constant during the whole run and so is that of the product. This system is mathematically equivalent to that of a mixture of two independent decomposition processes that have equal rate constants (Case II of the simulated data), which implies that both the proportion of the two reagents and that of the two products are kept constant during the process. Thus the conclusions about rank augmentation obtained for the simulated Case II can be extended to this real system.

In order to achieve rank augmentation of this rank-deficient data matrix, a simultaneous analysis of the matrices which contain kinetic runs of the same system at different pH values was carried out. The study was performed on column-wise augmented matrices, since it has already been shown for Case II of the simulated systems that full-rank augmentation can be achieved by column-wise matrix augmentation (in the spectral order). However, in order to evaluate whether the data are trilinear or not, the row-wise augmented matrices have also been analysed. Results obtained in the rank analysis of the column-wise and row-wise data matrices are also shown in Table 4. Rank analysis of the column-wise augmented matrices which contain that of the mixture and one with a single kinetic run (that of the acidic form, augmented matrix $[\mathbf{R}_2; \mathbf{R}_1]$, or that of the basic form, augmented matrix $[\mathbf{R}_2; \mathbf{R}_3]$) indicates that in both cases the rank deficiency is successfully broken. On the other hand, rank analysis of the row-wise $[\mathbf{R}_2, \mathbf{R}_1]$, $[\mathbf{R}_2, \mathbf{R}_3]$ and $[\mathbf{R}_2, \mathbf{R}_1, \mathbf{R}_3]$ gives a value of three, three and four respectively, which agrees with a system in which the rate constants vary from run to run (lack of trilinearity, Case II; see Table 3). From these results it can be concluded that the rate constants change between experiments at different pH values and therefore that the system lacks a trilinear structure.

Kinetic interpretation of rank analysis results

The NQS decomposition in aqueous medium is produced by the substitution of the sulphonate group in NQS by hydroxide ion (see Figure 2). Therefore this decomposition is expected to be a process which depends on both the concentrations of NQS and hydroxide ions, and the decomposition rate can be expressed

$$\frac{d[NQS]}{dt} = -k[NQS]^n[OH]^m \quad (1)$$

where n and m are the partial kinetic orders with respect to NQS and OH respectively.

Since during each kinetic run the pH was buffered in order to keep it constant during the process, the OH concentration did not vary in the run, and under these circumstances equation (1) can be rewritten as

$$\frac{d[NQS]}{dt} = -k'[NQS]^n \quad (2)$$

where k' is a pseudoconstant which depends on the concentration of hydroxide.

The rank results obtained are in agreement with the proposed kinetic model (equation (1)), which shows the dependence of the decomposition process on the hydroxide concentration (i.e. on the pH). Therefore the pseudoconstant in equation (2) will change from run to run and so will the shape of the kinetic profiles (lack of trilinearity).

Species resolution

Individual rank analysis of matrices \mathbf{R}_1 , \mathbf{R}_2 and \mathbf{R}_3 showed that in the three cases only two components can be resolved. The kinetic profiles and the pure UV spectra obtained from the individual ALS curve resolution treatment of these matrices are shown in Figure 4. Matrices \mathbf{R}_1 and \mathbf{R}_3 were obtained at pH values where a single initial NQS species predominated in the solution (the acidic or the basic one respectively). Thus these individual matrices are full rank, and full recovery of the kinetic profiles is achieved for these matrices. On the other hand, matrix \mathbf{R}_2 is rank-deficient, since it corresponded to an initial mixture of the two species (acidic and basic). Although it is a rank-

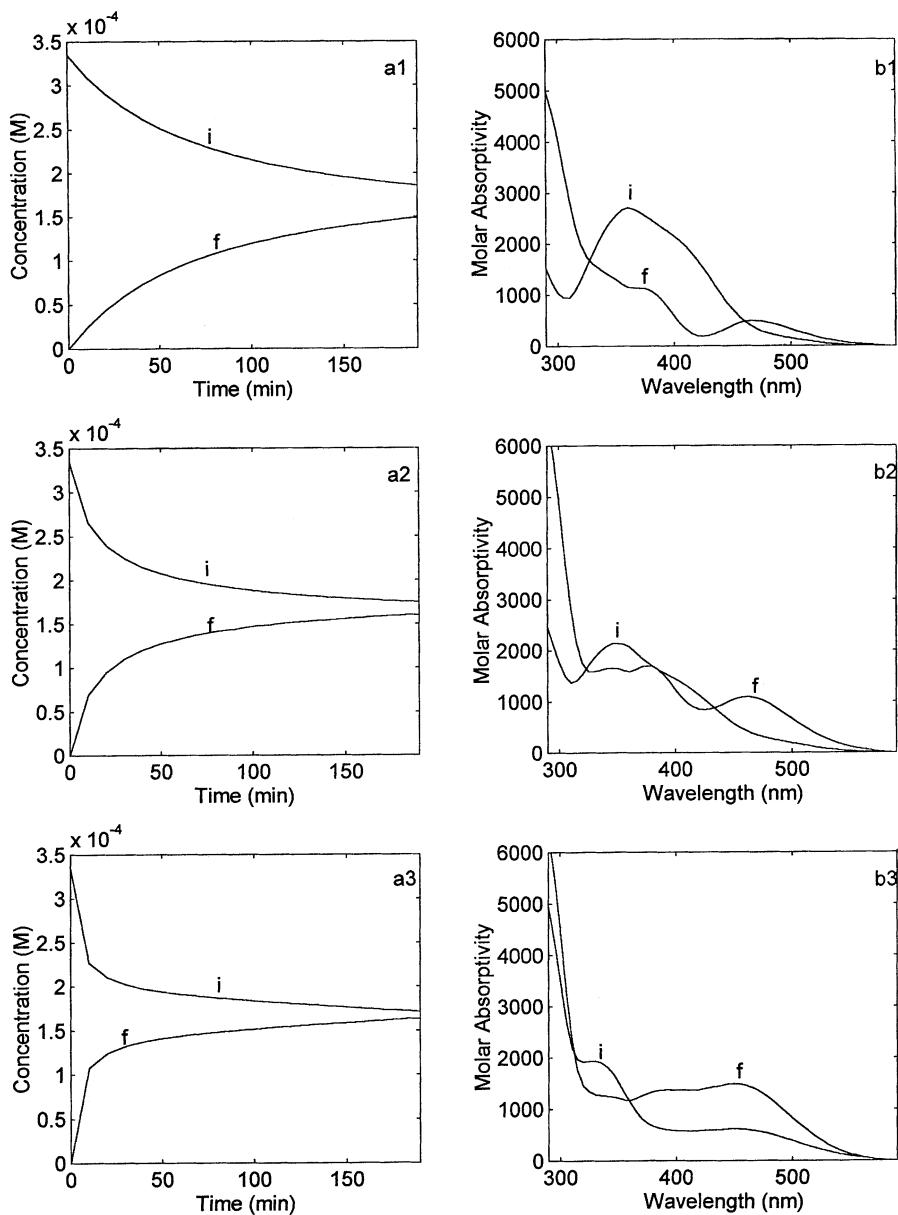


Figure 4. Kinetic profiles (a1 – a3) and pure spectra (b1 – b3) recovered in individual analysis of matrices \mathbf{R}_1 (pH 9.4), \mathbf{R}_2 (pH 10.5) and \mathbf{R}_3 (pH 13.3) by multivariate curve resolution. Species assignment: i, initial NQS species; f, NQS decomposition product

deficient system, all the constraints described for the analysis of a single matrix can still be applied. Only two concentration profiles are resolved for this system, which are related respectively to a linear combination of the concentration profiles of the acidic and basic species of the reagent, whose composition is constant along the process, and to a linear combination of the concentration profiles of the two degradation products, whose composition is also kept constant. Thus the reagent mixture has

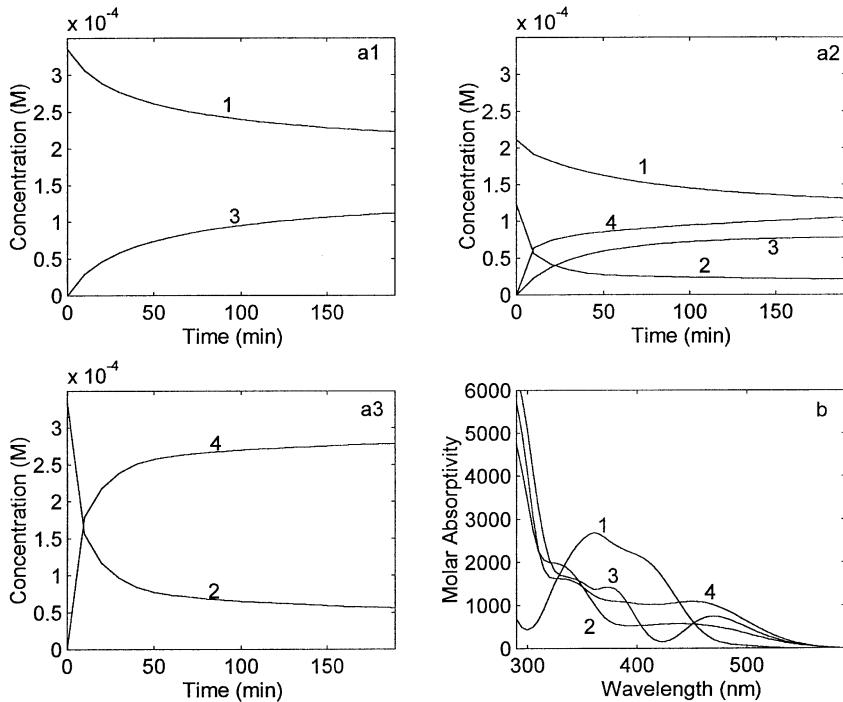


Figure 5. Kinetic profiles (a1 – a3) and pure spectra (b) recovered in simultaneous analysis of wavelength-wise augmented matrix $[\mathbf{R}_1; \mathbf{R}_2; \mathbf{R}_3]$ by multivariate curve resolution. Species assignment: 1, acidic NQS species; 2, basic NQS species; 3, acidic NQS decomposition product; 4, basic NQS decomposition product

an associated UV spectrum that is a linear combination of those corresponding to the acidic and basic species of NQS. Since the individual UV spectra are positive in all the working wavelength range, so is that of the mixture. The same conclusion can be deduced for the UV spectrum associated with the product mixture. Thus the constraint of positivity in both the spectral and concentration orders can also be applied to this rank-deficient system.

In order to obtain the correct resolution of matrix \mathbf{R}_2 , i.e. to evaluate the contribution of the two initial forms and of the two degradation products, a simultaneous analysis of data matrices \mathbf{R}_1 , \mathbf{R}_2 and \mathbf{R}_3 was carried out. Column-wise matrix augmentation was used to eliminate the rank deficiency present in this system. Figure 5 shows the concentration profiles and individual spectra obtained from the analysis of the full-rank augmented matrix. In this case, resolution is accomplished for all the compounds in the different kinetic runs. For the kinetic run \mathbf{R}_2 the first concentration profile that was obtained in the individual analysis is now resolved into two: that of the acidic NQS form and that of the basic NQS form. The same happens with the concentration profiles of the two decomposition products. The four spectra obtained correspond to those of the acidic and basic NQS reagent species and to those of the acidic and basic degradation products, and they coincide with the spectra obtained from the individual analysis of matrices \mathbf{R}_1 and \mathbf{R}_3 (see Figure 4). Since at the three working pH values there is a lack of selectivity over the whole monitored time interval (when the monitoring of the process begins, there is already a certain amount of decomposition product, and when monitoring is stopped, it cannot be assured that the process has already finished), both rotational and intensity ambiguities can still be present in the results. Owing to the possible presence of these ambiguities, the concentration profiles obtained should be considered as a qualitative picture of the evolution of the

different species with time under the different pH conditions. However, semiquantitative results can still be extracted from these profiles. For the reagent, its pK_a value has been calculated from the concentration profiles obtained for the acidic and basic species of the reagent at a working pH value of 10.5 (see Figure 5(a2)). A pK_a value of around 10.7 has been obtained, which agrees with the one obtained from the literature.²⁹ Furthermore, from the concentration profiles obtained for the acidic and basic species of the decomposition product (Figure 5(a2)), the pK_a value of this product can also be calculated: $pK_a \approx 10.4$. This pK_a value is similar to that already known for the initial NQS reagent, as expected owing to the high structural similarity between the reagent and the decomposition product.

CONCLUSIONS

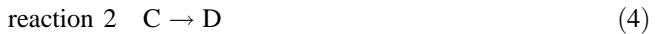
The resolution of simulated mixtures of two first-order kinetic processes in the presence of rank deficiencies caused by closure and rank overlap in any of the two orders (time or spectra) has been analysed. The rank deficiency caused by closure and rank overlap in the concentration profiles can be successfully eliminated by analysing an augmented wavelength-wise matrix which contains the individual matrix of the mixture and the individual matrix of a single kinetic process. However, this matrix augmentation cannot break the rank deficiency caused by spectral overlap. It has been demonstrated that the rank deficiency due to spectral overlap can only be broken by time-wise augmentation. Consequently, both types of matrix augmentation should be considered in order to completely eliminate rank deficiencies. Further work is being carried out in relation to the simultaneous study of both row- and column-wise augmented matrices in order to achieve resolution of different rank-deficient systems.

The conclusions obtained from the simulated systems have been extrapolated to an experimental case: the kinetic decomposition of NQS. Resolution of the rank deficient kinetic process (carried out at a pH value at which NQS is partially dissociated) has been achieved by wavelength-wise matrix augmentation. From the concentration profiles obtained at this pH condition, it has been possible to extract thermodynamic information related to acidic constants for both the NQS reagent and its decomposition product.

APPENDIX: RANK AUGMENTATION BY MATRIX AUGMENTATION. COMBINED EFFECTS: CLOSED REACTION SYSTEMS, RANK OVERLAP AND NON-TRILINEARITY

The effect of matrix augmentation (both column-wise and row-wise) in rank augmentation is studied for the four situations analysed in the paper.

Two simple kinetic reactions are studied:



Let us assume that matrix \mathbf{D}_1 corresponds to the simultaneous study of the two reactions and can be initially expressed as the sum of four rank-one contributions, i.e.

$$\mathbf{D}_1 = \mathbf{c}_{a1}\mathbf{s}_a^T + \mathbf{c}_{b1}\mathbf{s}_b^T + \mathbf{c}_{c1}\mathbf{s}_c^T + \mathbf{c}_{d1}\mathbf{s}_d^T \quad (5)$$

and that matrix \mathbf{D}_2 corresponds to the study of only one kinetic reaction, i.e.

$$\mathbf{D}_2 = \mathbf{c}_{a2}\mathbf{s}_a^T + \mathbf{c}_{b2}\mathbf{s}_b^T \quad (6)$$

where \mathbf{c}_a , \mathbf{c}_b , \mathbf{c}_c and \mathbf{c}_d are the concentrations of species A, B, C and D respectively and \mathbf{s}_a , \mathbf{s}_b , \mathbf{s}_c and \mathbf{s}_d are their corresponding spectra. Subindices 1 and 2 refer to matrices \mathbf{D}_1 and \mathbf{D}_2 respectively. In

expressions (5) and (6) it has been assumed that the spectra of the common species are the same in all the individual data matrices.

The column-wise augmented matrix $[\mathbf{D}_1; \mathbf{D}_2]$ can be expressed as

$$[\mathbf{D}_1; \mathbf{D}_2] = [\mathbf{c}_{a1}; \mathbf{c}_{a2}] \mathbf{s}_a^T + [\mathbf{c}_{b1}; \mathbf{c}_{b2}] \mathbf{s}_b^T + [\mathbf{c}_{c1}; \mathbf{0}] \mathbf{s}_c^T + [\mathbf{c}_{d1}; \mathbf{0}] \mathbf{s}_d^T \quad (7)$$

where $\mathbf{0}$ is a column vector of zeros. This column-wise augmented matrix is rank four.

The row-wise augmented matrix $[\mathbf{D}_1, \mathbf{D}_2]$ can be expressed as

$$[\mathbf{D}_1, \mathbf{D}_2] = \mathbf{c}_{a1} [\mathbf{s}_a^T, \mathbf{0}^T] + \mathbf{c}_{b1} [\mathbf{s}_b^T, \mathbf{0}^T] + \mathbf{c}_{c1} [\mathbf{s}_c^T, \mathbf{0}^T] + \mathbf{c}_{d1} [\mathbf{s}_d^T, \mathbf{0}^T] + \mathbf{c}_{a2} [\mathbf{0}^T, \mathbf{s}_a^T] + \mathbf{c}_{b2} [\mathbf{0}^T, \mathbf{s}_b^T] \quad (8)$$

This row-wise augmented matrix is rank six.

In the presence of trilinearity the following expressions are fulfilled for the concentration profiles of the same species in different data matrices:

$$\mathbf{c}_{a2} = k_a \mathbf{c}_{a1}, \quad \mathbf{c}_{b2} = k_b \mathbf{c}_{b1} \quad (9)$$

where k_a and k_b are scalars.

Introducing the trilinearity constraint, expressions (7) and (8) are transformed into

$$[\mathbf{D}_1; \mathbf{D}_2] = [\mathbf{c}_{a1}; k_a \mathbf{c}_{a1}] \mathbf{s}_a^T + [\mathbf{c}_{b1}; k_b \mathbf{c}_{b1}] \mathbf{s}_b^T + [\mathbf{c}_{c1}; \mathbf{0}] \mathbf{s}_c^T + [\mathbf{c}_{d1}; \mathbf{0}] \mathbf{s}_d^T \quad (10)$$

$$[\mathbf{D}_1, \mathbf{D}_2] = \mathbf{c}_{a1} [\mathbf{s}_a^T, k_a \mathbf{s}_a^T] + \mathbf{c}_{b1} [\mathbf{s}_b^T, k_b \mathbf{s}_b^T] + \mathbf{c}_{c1} [\mathbf{s}_c^T, \mathbf{0}^T] + \mathbf{c}_{d1} [\mathbf{s}_d^T, \mathbf{0}^T] \quad (11)$$

In trilinear systems both column-wise and row-wise augmented matrices are rank four. Conversely, when trilinearity is not present, the rank of the row-wise augmented matrix increases to six (equation (8)), while the rank of the column-wise augmented matrix is still four (equation (7)). Therefore the rank of the row-wise augmented matrix is ‘in excess’ with respect to the number of chemical components of the system.

Case I. Rank deficiency caused by closure

The two closure conditions for the first matrix are

$$t_{ab1} \mathbf{1} = \mathbf{c}_{a1} + \mathbf{c}_{b1} \quad (12)$$

$$t_{cd1} \mathbf{1} = \mathbf{c}_{c1} + \mathbf{c}_{d1} \quad (13)$$

and the closure condition for the second matrix is

$$t_{ab2} \mathbf{1} = \mathbf{c}_{a2} + \mathbf{c}_{b2} \quad (14)$$

where t_{ab1} , t_{cd1} and t_{ab2} are scalars, $\mathbf{1}$ is a column vector of ones and

$$t_{ab1} = k_x t_{cd1} = k_y t_{ab2} \quad (15)$$

Because of the two closure conditions, matrix \mathbf{D}_1 is rank-deficient and its rank is three (see text and References 24, 25 and 34, rank=NR+1=3). On the other hand, matrix \mathbf{D}_2 is full rank and its rank is two (see text and References 24, 25 and 34, rank=NR+1=2).

Column-wise matrix augmentation

When the closure constraints are included, expression (7) is transformed into

$$[\mathbf{D}_1; \mathbf{D}_2] = [\mathbf{c}_{a1}; \mathbf{c}_{a2}] \mathbf{s}_a^T + [(t_{ab1} \mathbf{1} - \mathbf{c}_{a1}); (t_{ab2} \mathbf{1} - \mathbf{c}_{a2})] \mathbf{s}_b^T + [\mathbf{c}_{c1}; \mathbf{0}] \mathbf{s}_c^T + [(t_{cd1} \mathbf{1} - \mathbf{c}_{c1}); \mathbf{0}] \mathbf{s}_d^T \quad (16)$$

This expression is valid for both trilinear and non-trilinear systems. The rank of the augmented column-wise matrix in the presence of closure is four, equal to the number of chemical components in

the system. Thus total resolution of the system will be possible.

Row-wise matrix augmentation

In the presence of closure the row-wise augmented matrix can be written as follows.

Trilinear systems Equation (11) is now

$$[\mathbf{D}_1, \mathbf{D}_2] = \mathbf{c}_{a1}([\mathbf{s}_a^T, k_a \mathbf{s}_a^T] - [\mathbf{s}_b^T, k_b \mathbf{s}_b^T]) + \mathbf{c}_{c1}([\mathbf{s}_c^T, \mathbf{0}^T] - [\mathbf{s}_d^T, \mathbf{0}^T]) + t_{ab1}\mathbf{1}([\mathbf{s}_b^T, k_b \mathbf{s}_b^T] + [k_x \mathbf{s}_d^T, \mathbf{0}^T]) \quad (17)$$

The resulting row-wise augmented matrix is rank three, lower than the number of chemical components in the system (rank deficiency). Total resolution of the system will not be possible.

Non-trilinear systems Equation (8) is now

$$\begin{aligned} [\mathbf{D}_1, \mathbf{D}_2] = & \mathbf{c}_{a1}([\mathbf{s}_a^T, \mathbf{0}^T] - [\mathbf{s}_b^T, \mathbf{0}^T]) + \mathbf{c}_{c1}([\mathbf{s}_c^T, \mathbf{0}^T] - [\mathbf{s}_d^T, \mathbf{0}^T]) + \mathbf{c}_{a2}([\mathbf{0}^T, \mathbf{s}_a^T] - [\mathbf{0}^T, \mathbf{s}_b^T]) \\ & + t_{ab1}\mathbf{1}([\mathbf{s}_b^T, \mathbf{0}^T] + [k_x \mathbf{s}_d^T, \mathbf{0}^T] + [\mathbf{0}^T, k_y \mathbf{s}_b^T]) \end{aligned} \quad (18)$$

The resulting row-wise augmented matrix is rank four, equal to the number of chemical components in the system. However, the resolved profiles will not be the true species profiles, but a combination of them.

Case II. Rank deficiency caused by closure and by rank overlap in concentration profiles

It is assumed that there is rank overlap in the concentration profiles between the two reagents (A and C) and between the two products (B and D). Accordingly, the following expressions can be written:

$$\mathbf{c}_{c1} = \alpha \mathbf{c}_{a1}, \quad \mathbf{c}_{d1} = \beta \mathbf{c}_{b1} \quad (19)$$

This corresponds to the case where the two kinetic processes have the same rate constant.

Column-wise matrix augmentation

If the rank overlap situation (equation (19)) is included, expression (16) is transformed into

$$[\mathbf{D}_1, \mathbf{D}_2] = [\mathbf{c}_{a1}; \mathbf{c}_{a2}] \mathbf{s}_a^T + [(t_{ab1}\mathbf{1} - \mathbf{c}_{a1}); (t_{ab2}\mathbf{1} - \mathbf{c}_{a2})] \mathbf{s}_b^T + [\alpha \mathbf{c}_{a1}; \mathbf{0}] \mathbf{s}_c^T + [(t_{cd1}\mathbf{1} - \alpha \mathbf{c}_{a1}); \mathbf{0}] \mathbf{s}_d^T \quad (20)$$

The column-wise augmented matrix is rank four. Total resolution of the system is possible.

Row-wise matrix augmentation

Trilinear systems If the rank overlap situation (equation (19)) is included, expression (17) is transformed into

$$[\mathbf{D}_1, \mathbf{D}_2] = \mathbf{c}_{a1} \{ [\mathbf{s}_a^T, k_a \mathbf{s}_a^T] - [\mathbf{s}_b^T, k_b \mathbf{s}_b^T] + \alpha ([\mathbf{s}_c^T, \mathbf{0}] - [\mathbf{s}_d^T, \mathbf{0}^T]) \} + t_{ab1}\mathbf{1}([\mathbf{s}_b^T, k_b \mathbf{s}_b^T] + [k_x \mathbf{s}_d^T, \mathbf{0}^T]) \quad (21)$$

The augmented matrix is rank two and total resolution of the system is not possible.

Non-trilinear systems

$$\begin{aligned} [\mathbf{D}_1, \mathbf{D}_2] = & \mathbf{c}_{a1} \{ [\mathbf{s}_a^T, \mathbf{0}^T] - [\mathbf{s}_b^T, \mathbf{0}^T] + \alpha ([\mathbf{s}_c^T, \mathbf{0}^T] - [\mathbf{s}_d^T, \mathbf{0}^T]) \} + \mathbf{c}_{a2}([\mathbf{0}^T, \mathbf{s}_a^T] - [\mathbf{0}^T, \mathbf{s}_b^T]) \\ & + t_{ab1}\mathbf{1}([\mathbf{s}_b^T, \mathbf{0}^T] + [k_x \mathbf{s}_d^T, \mathbf{0}^T] + [\mathbf{0}^T, k_y \mathbf{s}_b^T]) \end{aligned} \quad (22)$$

The augmented matrix is rank three and total resolution of the system is not possible.

Case III. Rank deficiency caused by closure and by rank overlap in spectral profiles

An overlap in the spectral profiles between the two reagents (A and C) and between the two products (B and D) is considered in this case. Accordingly, the following expressions can be written:

$$\mathbf{s}_a^T = \mathbf{s}_c^T, \quad \mathbf{s}_b^T = \mathbf{s}_d^T \quad (23)$$

Column-wise matrix augmentation

If the rank overlap constraint in the spectral profiles is included, expression (16) is transformed into

$$[\mathbf{D}_1; \mathbf{D}_2] = ([\mathbf{c}_{a1}; \mathbf{c}_{a2}] + [\mathbf{c}_{c1}; \mathbf{0}])\mathbf{s}_a^T + [(t_{ab1}\mathbf{1} - \mathbf{c}_{a1}); (t_{ab2}\mathbf{1} - \mathbf{c}_{a2})] + [(t_{cd1}\mathbf{1} - \mathbf{c}_{c1}); \mathbf{0}]\mathbf{s}_b^T \quad (24)$$

The augmented matrix is rank two. Total resolution of the system is not possible.

Row-wise matrix augmentation

Trilinear systems If rank overlap in the spectral profiles is now considered, expression (17) is transformed into

$$[\mathbf{D}_1, \mathbf{D}_2] = \mathbf{c}_{a1}([\mathbf{s}_a^T, k_a \mathbf{s}_a^T] - [\mathbf{s}_b^T, k_b \mathbf{s}_b^T]) + \mathbf{c}_{c1}([\mathbf{s}_a^T, \mathbf{0}^T] - [\mathbf{s}_b^T, \mathbf{0}^T]) + t_{ab1}\mathbf{1}([\mathbf{s}_b^T, k_b \mathbf{s}_b^T] + [k_x \mathbf{s}_b^T, \mathbf{0}^T]) \quad (25)$$

The rank of the row-wise augmented matrix is three. Total resolution of the system is not possible. The higher rank of this augmented matrix compared with the column-wise augmented one is due to the breaking of the rank deficiency caused by rank overlap in the spectral profiles when row-wise augmentation is performed.

Non-trilinear systems

$$[\mathbf{D}_1, \mathbf{D}_2] = (\mathbf{c}_{a1} + \mathbf{c}_{c1})([\mathbf{s}_a^T, \mathbf{0}^T] - [\mathbf{s}_b^T, \mathbf{0}^T]) + \mathbf{c}_{a2}([\mathbf{0}^T, \mathbf{s}_a^T] - [\mathbf{0}^T, \mathbf{s}_b^T]) + t_{ab1}\mathbf{1}([\mathbf{s}_b^T, \mathbf{0}^T] + [k_x \mathbf{s}_b^T, \mathbf{0}^T] + [\mathbf{0}^T, k_y \mathbf{s}_b^T]) \quad (26)$$

The rank of the row-wise augmented matrix is three. Total resolution of the system is not possible.

Case IV. Rank deficiency caused by closure and by rank overlap in concentration and spectral profiles

Considering closure and rank overlap in both the concentration and spectral profiles, the following expressions can be easily deduced.

Column-wise matrix augmentation

$$[\mathbf{D}_1; \mathbf{D}_2] = ([\mathbf{c}_{a1}; \mathbf{c}_{a2}] + [\alpha \mathbf{c}_{a1}; \mathbf{0}])\mathbf{s}_a^T + [(t_{ab1}\mathbf{1} - \mathbf{c}_{a1}); (t_{ab2}\mathbf{1} - \mathbf{c}_{a2})] + [(t_{cd1}\mathbf{1} - \alpha \mathbf{c}_{a1}); \mathbf{0}]\mathbf{s}_b^T \quad (27)$$

The rank of the column-wise augmented matrix is two. Total resolution of the system is not possible.

Row-wise matrix augmentation

Trilinear system

$$[\mathbf{D}_1, \mathbf{D}_2] = \mathbf{c}_{a1}\{[\mathbf{s}_a^T, k_a \mathbf{s}_a^T] - [\mathbf{s}_b^T, k_b \mathbf{s}_b^T] + \alpha([\mathbf{s}_a^T, \mathbf{0}^T] - [\mathbf{s}_b^T, \mathbf{0}^T])\} + t_{ab1}\mathbf{1}([\mathbf{s}_b^T, k_b \mathbf{s}_b^T] + [k_x \mathbf{s}_b^T, \mathbf{0}^T]) \quad (28)$$

The rank of the augmented matrix is two. Total resolution of the system is not possible.

Non-trilinear system

$$[\mathbf{D}_1, \mathbf{D}_2] = \mathbf{c}_{a1}(1 + \alpha)([\mathbf{s}_a^T, \mathbf{0}^T] - [\mathbf{s}_b^T, \mathbf{0}^T]) + \mathbf{c}_{a2}([\mathbf{0}^T, \mathbf{s}_a^T] - [\mathbf{0}^T, \mathbf{s}_b^T]) + t_{ab1}([\mathbf{s}_b^T, \mathbf{0}^T] + [k_x \mathbf{s}_b^T, \mathbf{0}^T] + [\mathbf{0}^T, k_y \mathbf{s}_b^T]) \quad (29)$$

The rank of the augmented matrix is three. Total resolution of the system is not possible.

The expressions deduced in this appendix allow us to find out in what situations full rank recovery of rank-deficient systems is achieved by data matrix augmentation. In cases of full rank recovery the resolution of the species profiles will be possible. However, in these cases the total resolution of the system is determined by other data structure aspects such as local rank,²⁶ the presence or not of trilinearity, the resolution method used or the presence of noise.

In cases where full rank recovery is not possible by data matrix augmentation and thus rank deficiency still persists, partial resolution of some of the species can still be possible. The expressions deduced in the appendix can help to infer which profiles could be recovered using a resolution method.

ACKNOWLEDGEMENT

This research has been funded by DGICYT (Spain) project PB96-0377.

REFERENCES

1. D. Perez-Bendito and M. Silva, *Trends Anal. Chem.* **15**, 232–240 (1996).
2. V. Cerdà, A. Cladera and J. M. Estela, *Quim. Anal.* **15**, 341–350 (1996).
3. A. Navas-Díaz and J. A. González-García, *Anal. Chem.* **66**, 988–993 (1994).
4. B. M. Quencer and S. R. Crouch, *Anal. Chem.* **66**, 458–463 (1994).
5. M. Gui and S. C. Rutan, *Anal. Chem.* **66**, 1513–1519 (1994).
6. M. Blanco, J. Coello, H. Iturriaga, S. MasPOCH and J. Riba, *Anal. Chem.* **66**, 2905–2911 (1994).
7. Y. Xie, J. J. Baeza-Baeza and G. Ramis-Ramos, *Chemometrics Intell. Lab. Syst.* **27**, 211–220 (1995).
8. Z. Gu and X. Wang, *Talanta*, **42**, 205–210 (1995).
9. S. Ventura, M. Silva and D. Pérez-Bendito, *Anal. Chem.* **67**, 4458–4461 (1995).
10. M. Blanco, J. Coello, H. Iturriaga, S. MasPOCH and M. Redón, *Anal. Chem.* **67**, 4477–4483 (1995).
11. A. Cladera, E. Gómez, J. M. Estela and V. Cerdà, *Anal. Chem.* **65**, 707–715 (1993).
12. J. M. Henshaw, L. W. Burgess, K. S. Booksh and B. R. Kowalski, *Anal. Chem.* **66**, 3328–3336 (1994).
13. R. Tauler, A. K. Smilde, J. M. Henshaw, L. W. Burgess and B. R. Kowalski, *Anal. Chem.* **66**, 3337–3344 (1994).
14. M. Gui, S. C. Rutan and A. Agbodjan, *Anal. Chem.* **67**, 3293–3299 (1995).
15. J. Saurina, S. Hernández-Cassou and R. Tauler, *Anal. Chem.* **69**, 2329–2336 (1997).
16. R. Tauler, A. Smilde and B. R. Kowalski, *J. Chemometrics*, **9**, 31–58 (1995).
17. E. Sanchez, L. S. Ramos and B. R. Kowalski, *J. Chromatogr.* **385**, 151–164 (1987).
18. E. Sanchez and B. R. Kowalski, *J. Chemometrics*, **4**, 29–45 (1990).
19. P. Kroonenberg and J. DeLeeuw, *Psychometrika*, **45**, 69–97 (1980).
20. P. J. Gemperline, K. H. Miller, T. L. West, J. E. Weinstein, J. G. Hamilton and J. T. Bray, *Anal. Chem.* **64**, 523A–531A (1992).
21. K. S. Booksh, J. M. Henshaw, L. W. Burgess and B. R. Kowalski, *J. Chemometrics*, **9**, 263–282 (1995).
22. A. K. Smilde, Y. Wang and B. R. Kowalski, *J. Chemometrics*, **8**, 21–36 (1994).
23. R. Tauler, A. Izquierdo-Ridorsa and E. Casassas, *Chemometrics Intell. Lab. Syst.* **18**, 293–300 (1993).
24. A. Izquierdo-Ridorsa, J. Saurina, S. Hernández-Cassou and R. Tauler, *Chemometrics Intell. Lab. Syst.* **38**, 183–196 (1997).
25. M. Amrhein, B. Srinivasan, D. Bonvin and M. M. Schumacher, *Chemometrics Intell. Lab. Syst.* **33**, 17–33 (1996).
26. R. Manne, *Chemometrics Intell. Lab. Syst.* **27**, 89–94 (1995).
27. J. Saurina and S. Hernández-Cassou, *J. Chromatogr. A*, **676**, 311 (1994).
28. J. Saurina and S. Hernández-Cassou, *Anal. Chim. Acta*, **281**, 593 (1993).

29. J. Saurina, S. Hernández-Cassou and R. Tauler, *Anal. Chim. Acta*, **312**, 189–198 (1995).
30. J. Saurina, S. Hernández-Cassou and R. Tauler, *Anal. Chem.* **67**, 3722–3726 (1995).
31. R. Tauler, A. Izquierdo-Ridorsa, R. Gargallo and E. Casassas, *Chemometrics Intell. Lab. Syst.* **27**, 163–174 (1995).
32. E. Casassas, R. Gargallo, A. Izquierdo-Ridorsa and R. Tauler, *React. Polym.* **27**, 1–14 (1995).
33. E. Casassas, R. Tauler and I. Marqués, *Macromolecules*, **27**, 1729–1737 (1994).
34. L. Norgaard and C. Ridder, *Chemometrics Intell. Lab. Syst.* **23**, 107–114 (1994).
35. R. Tauler, *Chemometrics Intell. Lab. Syst.* **30**, 133–146 (1995).
36. W. Windig, *Chemometrics Intell. Lab. Syst.* **16**, 1–23 (1992).
37. W. Windig and D. A. Stephenson, *Anal. Chem.* **64**, 2735–2742 (1992).