GENERALIZED RANK ANNIHILATION METHOD: STANDARD ERRORS IN THE ESTIMATED EIGENVALUES IF THE INSTRUMENTAL ERRORS ARE HETEROSCEDASTIC AND CORRELATED

KLAAS FABER,[1]* AVRAHAM LORBER[1‡] AND BRUCE R. KOWALSKI[1]

[1] *Laboratory for Chemometrics*, *Department of Chemistry*, *Box 351700, University of Washington, Seattle, WA 98195, U.S.A.*

SUMMARY

The generalized rank annihilation method (GRAM) is a method for curve resolution and calibration that uses two data matrices simultaneously, i.e. one for the unknown and one for the calibration sample. The method is known to become an eigenvalue problem for which the eigenvalues are the ratios of the concentrations for the samples under scrutiny. Previously derived standard errors in the estimated eigenvalues of GRAM have very recently been shown to be based on unrealistic assumptions about the measurement errors. In this paper a systematic notation is introduced that enables the propagation of errors that are based on realistic assumptions concerning the data-generating process. The error propagation will be performed in detail for the case that one data order modulates the second one. Extensions to more complicated error models are indicated. © 1997 by John Wiley & Sons, Ltd.

Journal of Chemometrics, Vol. **11,** 95–109 (1997) (No. of Figures: 1 No. of Tables: 0 No. of Refs: 27)

KEY WORDS GRAM; eigenvalues; standard errors; heteroscedastic noise; correlated noise

INTRODUCTION

Building on the work of Ho *et al.*[1–3] and Lorber,[4] Sanchez and Kowalski[5] developed a method for quantitative and qualitative multicomponent analysis. Their method, the generalized rank annihilation method (GRAM), performs the task of calibrating for the desired analytes in the presence of unknown interferents. This property is the so-called second-order advantage. An important characteristic of GRAM is that the second-order advantage is obtained with only one calibration sample. The complete solution consists of the ratio of concentrations of the desired analytes for the unknown and calibration sample as well as the pure analyte profiles. The method is known to become an eigenvalue problem for which the eigenvalues are the concentration ratios and the eigenvectors define the transformation matrix that is needed to rotate the abstract profiles, found by a singular value decomposition (SVD), to the pure analyte profiles. Rigorous treatments of the properties of GRAM in the absence of measurement errors have recently been given by Leurgans *et al.*[6] and Kiers and Smilde.[7]

Since the introduction of the rank annihilation concept there has been great interest in determining the relibility of the quantitative results, i.e. the concentration ratios or eigenvalues, obtained by this principle. The statistical properties standard error and bias in the estimated eigenvalues have been studied by several researchers either by simulations[8, 9] or by a theoretical approach.[2, 10–14] In addition,

Correspondence to: Bruce R. Kowalski

* Current address: Netherlands Forensic Science Institute, Volmerlaan 17, 2288 GD Rijswijk, The Netherlands.

‡ On leave from Nuclear Research Centre-Negev, PO Box 9001, Beer-Sheva 84190, Israel.

CCC 0886–9383/97/020095–15 $17.50

Received 13 October 1995
Accepted 11 June 1996

Mitchell and Burdick[15] have recently compared GRAM and the alternating least squares (ALS) algorithm for the analysis of second-order data. The theoretical approach for determining the reliability of the obtained results is based on making certain assumptions about the measurement errors and propagating these (idealized) errors through the SVD and the eigenvalue problem.* Recently a reformulation of the eigenvalue problem has led to a simplification of the error propagation.[16] Irrespective of the differences in approach taken in these theoretical contributions, they all have in common that the measurement error is assumed to be uncorrelated and homoscedastic. The advantage of this assumption is that it leads to a relatively simple expression for the standard error in the estimated eigenvalue. A definite disadvantage of this assumption is that many analytical instruments do not produce data that can be adequately described in this way.

An example of data for which a more complicated error model is necessary is the data that are obtained by a 'hyphenated' type of instrument, e.g. high-performance liquid chromatography with ultraviolet detection (HPLC–UV) or gas chromatography combined with mass spectrometry (GC–MS). Such instruments can actually be interpreted as a combination of two instruments. In the above examples the first instrument is a chromatograph and the second one is a multichannel detector. At regular time intervals a complete spectrum is measured and the data are conveniently cast into a matrix. If the data are collected from such an instrument, the first order (e.g. a chromatograph) is said to 'modulate' the second one (e.g. a multichannel detector). A direct consequence is that disturbances in the chromatography result in a random fluctuation of a complete spectrum. Furthermore, independently of this error, the detector contributes to the total uncertainty in the data. Booksh and Kowalski[9] have shown by means of simulations that the uncertainty in the final result, i.e. the estimated eigenvalues, can no longer be quantified by means of propagating uncorrelated, homoscedastic errors through the model. The complete mechanism behind the total measurement error has to be accounted for. Thus the practical usefulness of existing expressions for the standard errors in the eigenvalues of GRAM is heavily compromised.

The goals of this paper are to (1) revisit previously derived equations[14] for the case where the instrumental errors are uncorrelated and homoscedastic, (2) introduce a consistent notation that allows for the generalization of these equations to cases where the instrumental errors are correlated and heteroscedastic and (3) solve for the case where the first instrumental order modulates the second one. The solution will be given for Lorber's rank annihilation method,[4] the generalization of Lorber's method by Sanchez and Kowalski[5] and the modification of Sanchez and Kowalski's method by Wilson *et al*.[17]

For reasons of clarity it will be assumed in the remaining part of this paper that a data matrix is measured by a hyphenated method and the rows and columns will be denoted by spectra and elution profiles respectively. This assumption will not affect the generality of any of the obtained results.

NOTATION AND CONVENTIONS

Boldface uppercase letters represent matrices, e.g. **A**. Column vectors will be indicated by boldface lowercase letters, e.g. **a**. Scalars are indicated by italic uppercase or lowercase letters, e.g. A and a. Transposition of a matrix or vector is indicated by a superscripted 'T', e.g. $\mathbf{A}^{\mathrm{T}}$ and $\mathbf{a}^{\mathrm{T}}$. For a given matrix **A** the matrices $\mathbf{A}^{-1}$ and $\mathbf{A}^{+}$ stand for its inverse and pseudoinverse respectively. The 'inverse transpose' and 'pseudoinverse transpose' matrices will be denoted by $\mathbf{A}^{-\mathrm{T}}(=(\mathbf{A}^{-1})^{\mathrm{T}}=(\mathbf{A}^{\mathrm{T}})^{-1})$ and $\mathbf{A}^{+\mathrm{T}}(=(\mathbf{A}^{+})^{\mathrm{T}}=(\mathbf{A}^{\mathrm{T}})^{+})$ respectively. The matrix element in row i and column j of **A** will be specified by a row and column index as A_{ij}. The ith row and jth column of **A** will be denoted by $\mathbf{A}_{i-\mathrm{row}}$ and $\mathbf{A}_{j-\mathrm{col}}$

* In a formal sense one deals with uncertainties in the method of error propagation. The terms error and uncertainty will be used interchangeably in this paper.

respectively. Additional notation is necessary in order to perform the error propagation in a straightforward way. First, one needs to make a distinction between the true, i.e. errorless, quantities on one hand and the measured and estimated quantities on the other hand. A measured quantity is symbolized by adding a 'tilde' to the unadorned symbol for the true quantity, e.g. $\tilde{\mathbf{A}}$. An estimated quantity is denote by a 'hat', e.g. $\hat{\mathbf{A}}$. Finally, in the error propagation of GRAM it is convenient to distinguish between the uncertainty in a measured quantity and the uncertainty in an estimated quantity. The prefix symbol d is used to denote the uncertainty in a measured quantity, e.g. $d\mathbf{A}$, whereas the prefix symbol Δ is used to denote the resulting uncertainty in an estimated quantity, e.g. $\Delta\mathbf{A}$. Much of the present confusion in previous derivations of standard errors in the estimated eigenvalues[13, 14] might have arisen by not making this distinction explicit in the notation.

LORBER'S METHOD

In Lorber's method it is assumed that the components contributing to the calibration sample are a subset of the substituents of the unknown sample.* The modifications involved in going from Lorber's method to generalizations thereof are not complicated and therefore only this case is treated in detail. Furthermore, in order to illustrate how the error propagation can be carried out for another error model, most equations are worked out for the simplest error model, i.e. uncorrelated and homoscedastic instrumental noise.

Model without error in the instrument responses

The $I \times J$ unknown sample data matrix $\mathbf{M}$ is given as

$$\mathbf{M} = \mathbf{X}\mathbf{Y}^{\mathrm{T}} \tag{1}$$

where $\mathbf{X}$ $(I \times K)$ contains the errorless elution profiles of the K chemical components and $\mathbf{Y}$ $(J \times K)$ contains the corresponding errorless spectra. It is seen that the rows are spectra measured at different times $(i = 1, \ldots, I)$ and the columns of $\mathbf{M}$ correspond to elution profiles at different wavelengths $(j = 1, \ldots, J)$. The spectra in $\mathbf{Y}$ are normalized so that the concentration dependence is absorbed in $\mathbf{X}$.

The $I \times J$ calibration sample data matrix $\mathbf{N}$ is given as

$$\mathbf{N} = \mathbf{X}\boldsymbol{\Pi}\mathbf{Y}^{\mathrm{T}} \tag{2}$$

where $\boldsymbol{\Pi}$ is a $K \times K$ diagonal matrix that contains the ratios of concentrations in the two samples, i.e. $\pi_k = \boldsymbol{\Pi}_{kk} = c_{N,k}/c_{M,k}$. From the assumption about the presence of the substituents it follows that the diagonal of $\boldsymbol{\Pi}$ may contain zeros.

Propagation of uncorrelated, homoscedastic instrumental errors

This is the case that has previously been considered in the literature.[2, 10–14] With uncorrelated, homoscedastic (i.e. constant variance) measurement noise the unknown sample data matrix can be expressed as

$$\tilde{\mathbf{M}} = \mathbf{M} + d\mathbf{M} \tag{3}$$

where $d\mathbf{M}$ denotes the $I \times J$ matrix of measurement errors in $\tilde{\mathbf{M}}$. The assumption of uncorrelated, homoscedastic measurement noise implies that

* In a strict sense Lorber[4] considers the case where the calibration sample contains only one analyte. The eigenvalue problem, however, remains essentially the same under the current, more general, assumption.

$$E[dM_{ij}dM_{i'j'}]=\sigma_M^2\delta_{ii'}\delta_{jj'} \tag{4}$$

where $E[\cdot]$ symbolizes taking the expectation, σ_M denotes the standard deviation of the measurement noise in $\tilde{\mathbf{M}}$ and δ is the well-known Kronecker delta. Analogously, the calibration sample data matrix is expressed as

$$\tilde{\mathbf{N}}=\mathbf{N}+d\mathbf{N} \tag{5}$$

where $d\mathbf{N}$ denotes the $I\times J$ matrix of measurement errors in $\tilde{\mathbf{N}}$ and

$$E[dN_{ij}dN_{i'j'}]=\sigma_N^2\delta_{ii'}\delta_{jj'} \tag{6}$$

where σ_N stands for the standard deviation of the measurement noise in $\tilde{\mathbf{N}}$. The assumption of uncorrelated, homoscedastic noise implies that the noise is sample-independent. It follows that $\sigma_N=\sigma_M$.

In order to derive the eigenvalue problem defining GRAM, $\tilde{\mathbf{M}}$ is decomposed according to the singular value decomposition (SVD)

$$\tilde{\mathbf{M}}=\hat{\mathbf{U}}\hat{\mathbf{\Theta}}\hat{\mathbf{V}}^{\mathrm{T}} \tag{7}$$

where $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$ are matrices that contain the estimated left and right singular vectors respectively and $\hat{\mathbf{\Theta}}$ is the diagonal matrix with estimated singular values. $\tilde{\mathbf{M}}$ may be approximated by the first A singular vector dyads or principal components (PCs):

$$\hat{\mathbf{M}}_A=\hat{\mathbf{U}}_A\hat{\mathbf{\Theta}}_A\hat{\mathbf{V}}_a^{\mathrm{T}} \tag{8}$$

where the subscript indicates that only the first A PCs are involved in the approximation. In fact, A is an estimate of K, the number of chemical components that (significantly) contribute to the signal. One makes the important assumption here that the K-term SVD fit of the experimental data matrix is optimal. This is certaintly true if the instrumental errors are uncorrelated and homoscedastic (see equation (4)), but some data preprocessing, e.g. scaling, may be necessary in the general case (see Discussion).

It is noted that several variations of this eigenvalue problem are possible. These transcriptions give rise to the same eigenvalues, but the corresponding eigenvectors are different. Since the eigenvectors are needed for the reconstruction of the column and row profiles, i.e. **X** and **Y**, the reconstruction equations will also be different. By using the SVD representation of $\hat{\mathbf{M}}_A$, the following eigenvalue problem is obtained:

$$(\hat{\mathbf{\Theta}}_A^{-1}\hat{\mathbf{U}}_A^{\mathrm{T}}\tilde{\mathbf{N}}\hat{\mathbf{V}}_A)\hat{\mathbf{T}}=\hat{\mathbf{T}}\hat{\mathbf{\Pi}} \tag{9}$$

where $\hat{\mathbf{\Pi}}$ and $\hat{\mathbf{T}}$ are the estimated eigenvalues and eigenvectors respectively. The elution profiles and spectra are found as

$$\hat{\mathbf{X}}=\hat{\mathbf{U}}_A\hat{\mathbf{\Theta}}_A\hat{\mathbf{T}} \tag{10}$$

$$\hat{\mathbf{Y}}=\hat{\mathbf{V}}_A\hat{\mathbf{T}}^{-\mathrm{T}} \tag{11}$$

These equations show that the eigenvector matrix rotates the abstract decomposition of $\tilde{\mathbf{M}}$, i.e. equation (7), into the physical decomposition, defined by equations (1) and (3). Further inspection of equations (9)–(11) shows that the solution for the concentration ratios can be written as*

$$\hat{\mathbf{\Pi}}=\hat{\mathbf{X}}^{+}\tilde{\mathbf{N}}\hat{\mathbf{Y}}^{+\mathrm{T}} \tag{12}$$

* This fact is easily established by noting that $\hat{\mathbf{X}}$ and $\hat{\mathbf{Y}}$ are products of a column orthogonal and an invertible matrix (see Reference 18, page 34, exercise 8). It is emphasized that the reverse order law for inverse matrices, $(\mathbf{AB})^{-1}=\mathbf{B}^{-1}\mathbf{A}^{-1}$, does not always hold for the pseudoinverse.

Reformulating the eigenvalue problem as a regression problem makes the error propagation more transparent, primarily because one has to deal with the physical decomposition of $\hat{\mathbf{M}}$ instead of the abstract one. Especially the derivation of the bias in the estimated eigenvalues is facilitated in this way.[14] There is, however, an important difference with the usual regression problem, e.g. encountered in multivariate calibration. In the usual regression problem the pseudoinverse of a matrix enters the calculation. Here one has the pseudoinverse of two matrices, both of which are estimated by GRAM. It follows that, as a first step in the error propagation, one has to focus on the *reconstruction errors* in $\hat{\mathbf{X}}$ and $\hat{\mathbf{Y}}$ instead of the *measurement errors* in $\tilde{\mathbf{X}}$ and $\tilde{\mathbf{Y}}$. (The relation with the total measurement error in $\tilde{\mathbf{M}}$ will become clear later.) The task of propagating the reconstruction errors is easily accomplished by recognizing that the following holds for the reconstructed unknown sample data matrix:

$$\hat{\mathbf{M}}_A = \hat{\mathbf{X}}\hat{\mathbf{Y}}^{\mathrm{T}} = (\mathbf{X} + \Delta\mathbf{X})(\mathbf{Y} + \Delta\mathbf{Y})^{\mathrm{T}} \tag{13}$$

By combining equations (12) and (13), the uncertainty in the estimated eigenvalues* is expanded in terms of the uncertainties in the *measured* calibration data matrix and the *reconstructed* profiles as

$$\hat{\boldsymbol{\Pi}} = \boldsymbol{\Pi} + \Delta\boldsymbol{\Pi} = (\mathbf{X} + \Delta\mathbf{X})^{+}(\mathbf{N} + d\mathbf{N})(\mathbf{Y} + \Delta\mathbf{Y})^{+\mathrm{T}} \tag{14}$$

Multiplying out the far right-hand side of equation (14) and neglecting terms that contain products of errors results in

$$\Delta\boldsymbol{\Pi} = \mathbf{X}^{+}(d\mathbf{N})\mathbf{Y}^{+\mathrm{T}} + (\Delta\mathbf{X}^{+})\mathbf{N}\mathbf{Y}^{+\mathrm{T}} + \mathbf{X}^{+}\mathbf{N}\Delta\mathbf{Y}^{+\mathrm{T}} \tag{15}$$

The first term is the contribution from the measured calibration data matrix $\tilde{\mathbf{N}}$ and the last two terms are the contributions from the reconstructed unknown data matrix $\hat{\mathbf{M}}_A$. It is possible to combine those two terms to obtain an expression where only $\Delta\mathbf{M}$ contributes using the following expression for the differential of the Moore–Penrose pseudoinverse of a matrix $\mathbf{A}$:[18]

$$\Delta\mathbf{A}^{+} = -\mathbf{A}^{+}(\Delta\mathbf{A})\mathbf{A}^{+} + \mathbf{A}^{+}\mathbf{A}^{+\mathrm{T}}(\Delta\mathbf{A}^{\mathrm{T}})(\mathbf{I} - \mathbf{A}\mathbf{A}^{+}) + (\mathbf{I} - \mathbf{A}^{+}\mathbf{A})(\Delta\mathbf{A}^{\mathrm{T}})\mathbf{A}^{+\mathrm{T}}\mathbf{A}^{+} \tag{16}$$

where $\mathbf{I}$ denotes an appropriately dimensioned identity matrix. For a full column rank matrix, e.g. $\hat{\mathbf{X}}$, the third term on the right-hand side is identical to zero and for a full row rank matrix, e.g. $\hat{\mathbf{Y}}^{\mathrm{T}}$, the second term vanishes. Now, using $\mathbf{A}\mathbf{A}^{+}\mathbf{A} = \mathbf{A}$ gives

$$\begin{aligned}(\Delta\mathbf{X}^{+})\mathbf{N}\mathbf{Y}^{+\mathrm{T}} &= -\mathbf{X}^{+}(\Delta\mathbf{X})\mathbf{X}^{+}\mathbf{X}\boldsymbol{\Pi}\mathbf{Y}^{\mathrm{T}}\mathbf{Y}^{+\mathrm{T}} + \mathbf{X}^{+}\mathbf{X}^{+\mathrm{T}}(\Delta\mathbf{X}^{\mathrm{T}})(\mathbf{I} - \mathbf{X}\mathbf{X}^{+})\mathbf{X}\boldsymbol{\Pi}\mathbf{Y}^{\mathrm{T}}\mathbf{Y}^{+\mathrm{T}}\\ &= -\mathbf{X}^{+}(\Delta\mathbf{X})\boldsymbol{\Pi}\end{aligned} \tag{17}$$

$$\begin{aligned}\mathbf{X}^{+}\mathbf{N}\Delta\mathbf{Y}^{+\mathrm{T}} &= -\mathbf{X}^{+}\mathbf{X}\boldsymbol{\Pi}\mathbf{Y}^{\mathrm{T}}\mathbf{Y}^{+\mathrm{T}}(\Delta\mathbf{Y}^{\mathrm{T}})\mathbf{Y}^{+\mathrm{T}} + \mathbf{X}^{+}\mathbf{X}\boldsymbol{\Pi}\mathbf{Y}^{\mathrm{T}}(\mathbf{I} - \mathbf{Y}^{+\mathrm{T}}\mathbf{Y}^{\mathrm{T}})(\Delta\mathbf{Y})\mathbf{Y}^{+}\mathbf{Y}^{+\mathrm{T}}\\ &= -\boldsymbol{\Pi}\mathbf{X}^{+}[\mathbf{X}(\Delta\mathbf{Y}^{\mathrm{T}})]\mathbf{Y}^{+\mathrm{T}}\end{aligned} \tag{18}$$

Equation (17) is further manipulated by noting that only the diagonal elements are of interest. The diagonal elements of a matrix do not change if the matrix is premultiplied by a diagonal matrix and postmultiplied by the corresponding inverse. Therefore

$$\mathrm{diag}[(\Delta\mathbf{X}^{+})\mathbf{N}\mathbf{Y}^{+\mathrm{T}}] = \mathrm{diag}[-\mathbf{X}^{+}(\Delta\mathbf{X})\boldsymbol{\Pi}] = \mathrm{diag}[-\boldsymbol{\Pi}\mathbf{X}^{+}[(\Delta\mathbf{X})\mathbf{Y}^{\mathrm{T}}]\mathbf{Y}^{+\mathrm{T}}] \tag{19}$$

where diag[·] stands for a vector that contains the diagonal elements of a matrix. The reconstruction error in $\hat{\mathbf{M}}_A$ is found by multiplying out equation (13) and neglecting the cross-term $\Delta\mathbf{X}\Delta\mathbf{Y}^{\mathrm{T}}$ as

* In the remainder of the paper focus is on *non-degenerate* eigenvalues. Degenerate eigenvalues lead to reconstructed profiles which are linear combinations of the true profiles, since the associated eigenvectors are not unique (rotation problem). Rigorous discussions of the circumstances under which GRAM gives a unique solution for the analyte of interest are given by Leurgans *et al.*[6] and Kiers and Smilde.[7]

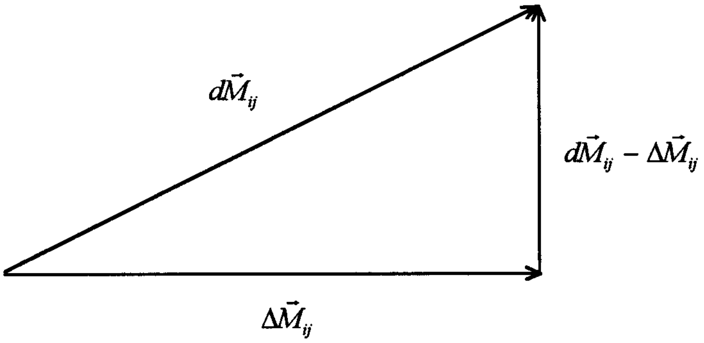

Figure 1. The total uncertainty dM_{ij} in an individual matrix element $\tilde{M}_{ij}$ can be represented as a vector, i.e. $d\vec{M}_{ij}$. It is seen that this vector is the sum of two distinct contributions. One part, i.e. $\Delta\vec{M}_{ij}$, lies in the space spanned by **M** and the other part, i.e. $d\vec{M}_{ij} - \Delta\vec{M}_{ij}$, is orthogonal to this 'embedded error'.[10] Only the imbedded error is explicitly propagated through the eigenvalue problem. The part of the total error that is projected out, i.e. the 'extracted' error, must, however, also be taken into consideration in order to establish the relation between the standard error in an estimated quantity and the total measurement error

$$\Delta\mathbf{M} = \hat{\mathbf{M}}_A - \mathbf{M} = (\Delta\mathbf{X})\mathbf{Y}^{\mathrm{T}} + \mathbf{X}\Delta\mathbf{Y}^{\mathrm{T}} \tag{20}$$

Combining equations (18)–(20) yields

$$\mathrm{diag}[(\Delta\mathbf{X}^{+})\mathbf{N}\mathbf{Y}^{+\mathrm{T}} + \mathbf{X}^{+}\mathbf{N}\Delta\mathbf{Y}^{+\mathrm{T}}] = \mathrm{diag}[-\mathbf{\Pi}\mathbf{X}^{+}(\Delta\mathbf{M})\mathbf{Y}^{+\mathrm{T}}] \tag{21}$$

Finally, equation (21) is seen to contain the projection of the reconstruction error **ΔM** onto the column and row space of **M** (true values). It follows that errors that are orthogonal to **ΔM** are eliminated by the projection:

$$\mathbf{X}^{+}(\Delta\mathbf{M})\mathbf{Y}^{+\mathrm{T}} = \mathbf{X}^{+}(d\mathbf{M})\mathbf{Y}^{+\mathrm{T}} - \mathbf{X}^{+}(d\mathbf{M} - \Delta\mathbf{M})\mathbf{Y}^{+\mathrm{T}} = \mathbf{X}^{+}(d\mathbf{M})\mathbf{Y}^{+\mathrm{T}} \tag{22}$$

This is illustrated in Figure 1. However, any error that is projected out in this way but is present in the original data must be added to the propagated reconstruction error. Thus equation (15) is reduced to*

$$\mathrm{diag}[\Delta\mathbf{\Pi}] = \mathrm{diag}[\mathbf{X}^{+}(d\mathbf{N})\mathbf{Y}^{+\mathrm{T}} - \mathbf{\Pi}\mathbf{X}^{+}(d\mathbf{M})\mathbf{Y}^{+\mathrm{T}}] \tag{23}$$

Equation (23) shows the relative importance of the uncertainty in the calibration and unknown sample data matrices. The minus sign is to be expected, since the unknown sample data matrix is 'inverted' in equation (9). It is emphasized that the total measurement error in $\tilde{\mathbf{M}}$ must be propagated although a substantial part is *not imbedded* in the data. Equation (23) gives the variation in the diagonal of the estimated eigenvalue matrix $\hat{\mathbf{\Pi}}$ resulting from variations in the measured response matrices. By introducing $\boldsymbol{\xi}_a^{\mathrm{T}} = (\mathbf{X}^{+})_{a-\mathrm{row}}$ and $\boldsymbol{\eta}_a = (\mathbf{Y}^{+\mathrm{T}})_{a-\mathrm{col}}$ the expression for an individual eigenvalue, $\hat{\pi}_a = \hat{\mathbf{\Pi}}_{aa}$, becomes

$$\Delta\pi_a = \boldsymbol{\xi}_a^{\mathrm{T}}(d\mathbf{N})\boldsymbol{\eta}_a - \pi_a\boldsymbol{\xi}_a^{\mathrm{T}}(d\mathbf{M})\boldsymbol{\eta}_a \tag{24}$$

This equation is conveniently worked out by applying the vec operator. The vec operator 'strings out' a matrix columnwise to yield a single column vector. Magnus and Neudecker[18] give an excellent treatment of the use of the vec operator in differential calculus. First one uses the identity[18]

* The situation is similar to ordinary least squares (OLS) where the residuals of the dependent variable **y** have to be corrected for the number of degrees of freedom in order to obtain an unbiased estimate of the total error in **y**, which is the sum of model error and measurement error. This estimate can then be propagated through the model in order to obtain an estimate for the standard error in the parameters.

$$\text{vec}(\mathbf{ABC}) = (\mathbf{C}^{\mathrm{T}} \otimes \mathbf{A})\text{vec}\,\mathbf{B} \tag{25}$$

where '$\otimes$' denotes the Kronecker product. Given an $n \times m$ matrix $\mathbf{A}$ and a $p \times q$ matrix $\mathbf{B}$, the Kronecker product builds an $np \times mq$ 'supermatrix' as follows:

$$\mathbf{A} \otimes \mathbf{B} = \begin{pmatrix} A_{11}\mathbf{B} & \dots & A_{1m}\mathbf{B} \\ \vdots & & \vdots \\ A_{n1}\mathbf{B} & \dots & A_{nm}\mathbf{B} \end{pmatrix} \tag{26}$$

Applying equation (25) to equation (24) results in

$$\begin{aligned} \Delta\pi_a &= \text{vec}(\Delta\pi_a) \\ &= \text{vec}(\boldsymbol{\xi}_a^{\mathrm{T}}(d\mathbf{N})\boldsymbol{\eta}_a) - \text{vec}(\pi_a\boldsymbol{\xi}_a^{\mathrm{T}}(d\mathbf{M})\boldsymbol{\eta}_a) \\ &= (\boldsymbol{\eta}_a^{\mathrm{T}} \otimes \boldsymbol{\xi}_a^{\mathrm{T}})\text{vec}d\mathbf{N} - (\boldsymbol{\eta}_a^{\mathrm{T}} \otimes \pi_a\boldsymbol{\xi}_a^{\mathrm{T}})\text{vec}d\mathbf{M} \end{aligned} \tag{27}$$

In order to find the standard error in the estimated eigenvalue, one must first derive the variance. Using the identity[18]

$$(\mathbf{A} \otimes \mathbf{B})^{\mathrm{T}} = (\mathbf{A}^{\mathrm{T}} \otimes \mathbf{B}^{\mathrm{T}}) \tag{28}$$

an approximate expression for the variance is found by taking the expectation of the squared deviation from the mean:

$$\begin{aligned} V(\hat{\pi}_a) &\approx E[(\Delta\pi_a)^2] \\ &= E[\Delta\pi_a\Delta\pi_a^{\mathrm{T}}] \\ &= (\boldsymbol{\eta}_a^{\mathrm{T}} \otimes \boldsymbol{\xi}_a^{\mathrm{T}})\mathbf{V}(\mathbf{N})(\boldsymbol{\eta}_a \otimes \boldsymbol{\xi}_a) + (\boldsymbol{\eta}_a^{\mathrm{T}} \otimes \pi_a\boldsymbol{\xi}_a^{\mathrm{T}})\mathbf{V}(\mathbf{M})(\boldsymbol{\eta}_a \otimes \boldsymbol{\xi}_a\pi_a) \end{aligned} \tag{29}$$

where $\mathbf{V}(\mathbf{N})$ and $\mathbf{V}(\mathbf{M})$ signify the covariance matrices of the measurement errors in $\tilde{\mathbf{N}}$ and $\tilde{\mathbf{M}}$ respectively. $\mathbf{V}(\mathbf{N})$ and $\mathbf{V}(\mathbf{M})$ are defined by

$$\mathbf{V}(\mathbf{N}) = E[\text{vec}d\mathbf{N}(\text{vec}d\mathbf{N})^{\mathrm{T}}] \tag{30}$$

$$\mathbf{V}(\mathbf{M}) = E[\text{vec}d\mathbf{M}(\text{vec}d\mathbf{M})^{\mathrm{T}}] \tag{31}$$

Since the measurement errors are assumed at this point to be uncorrelated and homoscedastic, the expressions simplify to $\mathbf{V}(\mathbf{N}) = \sigma_N^2\mathbf{I}$ and $\mathbf{V}(\mathbf{M}) = \sigma_M^2\mathbf{I}$, where $\mathbf{I}$ is the $IJ \times IJ$ identity matrix. Furthermore, cross-terms containing the covariance between the measurement errors in $\tilde{\mathbf{N}}$ and $\tilde{\mathbf{M}}$ are not present in equation (29). By using[18]

$$(\mathbf{A} \otimes \mathbf{B})(\mathbf{C} \otimes \mathbf{D}) = \mathbf{AC} \otimes \mathbf{BD} \tag{32}$$

the final result for the variance is obtained:

$$V(\hat{\pi}_a) \approx \|\boldsymbol{\xi}_a\|^2\|\boldsymbol{\eta}_a\|^2(\sigma_N^2 + \pi_a^2\sigma_M^2) \tag{33}$$

This result can be further simplified by using $\sigma_N = \sigma_M$. The approximate expression for the covariance between different estimated eigenvalues $\hat{\pi}_a$ and $\hat{\pi}_{a'}$,

$$C(\hat{\pi}_a, \hat{\pi}_{a'}) \approx E[\Delta\pi_a\Delta\pi_{a'}] \tag{34}$$

is worked out in a similar manner. The standard error in $\hat{\pi}_a$ follows as

$$\sigma(\hat{\pi}_a) = V(\hat{\pi}_a)^{1/2} \tag{35}$$

It is important to note that equation (33) gives an *approximation* to the variance in the estimated eigenvalue $\hat{\pi}_a$. It may not be equal to the true variance owing to the assumptions and approximations made in order to obtain it. Additional approximations result from the practical evalution of equation (33). It is seen that this equation contains quantities that are unknown and have to be replaced by estimates. Thus the practical result of theoretical error propagation is the estimated variance in the estimated eigenvalue, $\hat{V}(\hat{\pi}_a)$. It follows that the adequacy of equation (33) for predicting the true variance has to be evaluated. The adequacy of equation (33) has been thoroughly tested by means of Monte Carlo simulations for relevant noise levels.[14] An important rule of thumb emerging from that study is that for data with a high signal-to-noise ratio, which is achievable with modern spectrophotometers, the uncertainty introduced by inserting estimates for the standard deviations of the measurement errors (σ_N and σ_M) dominates the uncertainties associated with the other approximations. In other words, the limits for the successful application of theoretical error propagation results are not set by the approximations made during the derivation (only standard assumptions are used in this work) but by the precision to which the measurement errors are known.

Another consequence of substituting estimated quantities in equation (33) is that the estimated standard error $\hat{\sigma}(\hat{\pi}_a)$ and the estimated eigenvalue $\hat{\pi}_a$ are *not* independent. This is a requirement if a traditional test for significance (e.g. a *t*-test) is to be used. A so-called variance-stabilizing transform has been proposed to overcome this problem.[14]

Propagation of correlated, heteroscedastic instrumental errors

Booksh and Kowalski[9] have demonstrated that a realistic noise model for data obtained for a hyphenated method is as follows:

$$\tilde{\mathbf{M}} = (\mathbf{M} + d\mathbf{X}_M \mathbf{Y}^{\mathrm{T}}) + d\mathbf{M} = \mathbf{M} + d\mathbf{M}^{\mathrm{tot}} \tag{36}$$

$$\tilde{\mathbf{N}} = (\mathbf{N} + d\mathbf{X}_N \tilde{\boldsymbol{\Pi}} \mathbf{Y}^{\mathrm{T}}) + d\mathbf{N} = \mathbf{N} + d\mathbf{N}^{\mathrm{tot}} \tag{37}$$

where the matrices $d\mathbf{X}_M$ and $d\mathbf{X}_N$ denote the errors in the modulating order and $d\mathbf{M}^{\mathrm{tot}}$ and $d\mathbf{N}^{\mathrm{tot}}$ are the total uncertainties in the measured data matrices $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{N}}$ respectively. In the case of a spectro–chromatogram the modulating order is generated by the chromatograph. At discrete time intervals a complete spectrum is measured. A random error in the time domain will affect the complete spectrum measured at that particular time. On top of this error the detector noise is added, as shown by equations (36) and (37). The parentheses in equations (36) and (37) become operational if one of the two error contributions is proportional to the size of the data. Analogously to equations (30) and (31) the covariance matrices are defined by

$$\mathbf{V}(\mathbf{N}) = E[\mathrm{vec} d\mathbf{N}^{\mathrm{tot}} (\mathrm{vec} d\mathbf{N}^{\mathrm{tot}})^{\mathrm{T}}] \tag{38}$$

$$\mathbf{V}(\mathbf{M}) = E[\mathrm{vec} d\mathbf{M}^{\mathrm{tot}} (\mathrm{vec} d\mathbf{M}^{\mathrm{tot}})^{\mathrm{T}}] \tag{39}$$

Owing to the correlation of the measurement errors within one order, these matrices are not diagonal as for the case discussed earlier. The derivation of the covariance matrix $\mathbf{V}(\mathbf{N})$ follows. Making the reasonable assumptions that the errors in the first order as well as the detector noise are uncorrelated (see Reference 9 for more details) gives the following values for individual elements of $\mathbf{V}(\mathbf{N})$:

$$\begin{aligned}
\sigma(N^{\mathrm{tot}}_{iji'j'})^2 &= E[dN^{\mathrm{tot}}_{ij}\, dN^{\mathrm{tot}}_{i'j'}] \\
&= E[(dN^{\mathrm{tot}}_{ij})^{\mathrm{T}}\, dN^{\mathrm{tot}}_{i'j'}] \\
&= \mathbf{Y}_{j-\mathrm{row}} \Pi E[(d\mathbf{X}^{\mathrm{T}}_N)_{i-\mathrm{col}} (d\mathbf{X}_N)_{i'-\mathrm{row}}] \Pi (\mathbf{Y}^{\mathrm{T}})_{j'-\mathrm{col}} + E[(dN_{ij})^{\mathrm{T}}\, dN_{i'j'}] \\
&= [\mathbf{Y}_{j-\mathrm{row}} \Pi \mathbf{V}(\mathbf{X}^{ii'}_N) \Pi (\mathbf{Y}^{\mathrm{T}})_{j'-\mathrm{col}} + \sigma(N_{ij})^2 \delta_{jj'}]\delta_{ii'}
\end{aligned} \tag{40}$$

where $\mathbf{V}(\mathbf{X}_N^{ii'})$ denotes the covariance matrix for rows i and i' of the measured profile matrix $\tilde{\mathbf{X}}_\mathbf{N}$. This equation is as general as possible (within the context of this paper) with respect to allowing for heteroscedastic and correlated instrumental errors: the uncertainty in the first order is allowed to depend on the row index (it may even vary among substituents) and the detector noise is allowed to depend on both row and column index.

A pictorial representation of **V**(**N**) is

$$\mathbf{V}(\mathbf{N})=\begin{pmatrix} \sigma(N_{1111}^{\text{tot}})^2 & & \ldots & \mathbf{0}_{I-1} \\ \mathbf{0}_{I-1} & \sigma(N_{2121}^{\text{tot}})^2 & & \sigma(N_{I1IJ}^{\text{tot}})^2 \\ \sigma(N_{1211}^{\text{tot}})^2 & \mathbf{0}_{I-1} & \ldots & \vdots \\ \vdots & \sigma(N_{2221}^{\text{tot}})^2 & & \\ & \vdots & \ldots & \sigma(N_{IJIJ}^{\text{tot}})^2 \end{pmatrix} \tag{41}$$

where $\mathbf{0}_{I-1}$ is an $I-1$ zero vector. It is seen that **V**(**N**) is a sparse $IJ \times IJ$ matrix for which only a fraction $1/I$ elements is non-zero. For more complicated error models, which are beyond the scope of this paper, the fraction of zeros may be much smaller. This will certainly be true if the noise in adjacent channels of the detector are correlated. Extending the current derivation to this kind of measurement error models is straightforward, the only complication being that additional information (in the form of correlation coefficients) is necessary in order to evaluate the resulting expression in practice. The covariance matrix **V**(**M**) is obtained by substituting **M** for **N** in equation (40) and eliminating **Π**.

The variance in the estimated eigenvalues follows by working out the analog of equation (29). The result is

$$V(\hat{\pi}_a) \approx \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{j'=1}^{J} \xi_{a,i}^2 \eta_{a,j} \eta_{a,j'} \left[\sigma(N_{ijij'}^{\text{tot}})^2 + \pi_a^2 \sigma(M_{ijij'}^{\text{tot}})^2\right] \tag{42}$$

where the symbols are as defined above. Although more complicated than equation (33), this equation still provides useful insight into the relative importance of the individual error contributions.

At this point a considerable simplification results from assuming that the errors can be described by only two standard deviations, say σ_X and $\sigma_Y = \sigma_N = \sigma_M$. For example, the simulations performed by Booksh and Kowalski[9] are based on this assumption. (They use the symbols σ_X and σ_Y to denote the uncertainty in the modulating order and the detector noise respectively.) Several cases can be considered now, depending on whether the errors are assumed to be *additive*, i.e. independent of the data, or *relative*, i.e. proportional to the data. For example, with relative errors in the modulating order and additive detector noise, one obtains

$$V(\hat{\pi}_a) \approx 2\pi_a^2 \|\boldsymbol{\xi}_a\|^2 \sigma_X^2 + (1+\pi_a^2)\|\boldsymbol{\xi}_a\|^2 \|\boldsymbol{\eta}_a\|^2 \sigma_Y^2 \tag{43}$$

which shows that the standard error $\sigma(\hat{\pi}_a)$ relates to the different standard deviations σ_X and σ_Y as

$$\sigma(\hat{\pi}_a) \approx (\beta_1 \sigma_X^2 + \beta_2 \sigma_Y^2)^{1/2} \tag{44}$$

with β_1 and β_2 as defined in equation (43). Equation (44) is identical to the expression found by Booksh and Kowalski to describe their simulation results quantitatively (see equation (15) in Reference 9). Furthermore, from equation (43) it also becomes clear how the precision in the estimated eigenvalue will improve by increasing the number of data points in each order. It is easily verified that, for example, quadrupling both the number of rows and columns will decrease the term corresponding to σ_X by a factor of four and the term corresponding to σ_Y by a factor of 16. This translates directly to the simulation results given in Table 6a of Reference 9. These observations lend credence to the currently obtained results.

GENERALIZATION OF SANCHEZ AND KOWALSKI

The previous section considered the error propagation through the eigenvalue problem derived by Lorber.[4] This eigenvalue problem will result in a correct solution if the unknown sample contains all substituents that are present in the calibration sample in an amount that is sufficient to build the solution space. This procedure will, for example, work well in a standard addition situation. However, in the general case where both samples contain unique substituents, Lorber's method will fail and a modification is necessary. (An example is where the calibration sample contains analyte A and interferent I_1 but the unknown sample contains analyte A and interferent I_2. The unknown sample data matrix does not span the space of the calibration sample data matrix.) Sanchez and Kowalski[5] proposed to solve for the same eigenvalue problem after replacing $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{N}}$ by $\tilde{\mathbf{Q}}=\tilde{\mathbf{M}}+\tilde{\mathbf{N}}$ and $\tilde{\mathbf{M}}$ respectively. This will ensure that the solution space resulting from the SVD (see equation (7)) describes all substituents. Now the GRAM eigenvalues are related to the concentrations as $\pi_k=c_{M,k}/(c_{M,k}+c_{N,k})$.

Two cases will be considered here. The first case is the situation originally envisioned by Sanchez and Kowalski,[5] i.e. the situation where only one calibration sample is available. The second case is a generalization to the situation where multiple calibration samples are available, giving data matrices $\tilde{\mathbf{N}}_s$ $(s=1, \ldots, S)$. A straightforward generalization discussed by Leurgans *et al.*[6] is obtained by constructing the eigenvalue problem with $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{N}}$ replaced by $\tilde{\mathbf{Q}}=\tilde{\mathbf{M}}+\sum_{s=1}^{S}\tilde{\mathbf{N}}_s$ and $\tilde{\mathbf{M}}$ respectively. It should be noted that this generalization is different from the trilinear decomposition (TLD) method.[19] The error analysis of TLD is considered to be outside the scope of this paper.

One calibration sample

Inserting $\mathbf{Q}=\mathbf{M}+\mathbf{N}$ and $\mathbf{M}$ in equation (23) gives

$$\begin{aligned}\Delta\boldsymbol{\Pi}&=\mathbf{X}^{+}(d\mathbf{M})\mathbf{Y}^{+\mathrm{T}}-\boldsymbol{\Pi}\mathbf{X}^{+}(d\mathbf{Q})\mathbf{Y}^{+\mathrm{T}}\\&=(\mathbf{I}-\boldsymbol{\Pi})\mathbf{X}^{+}(d\mathbf{M})\mathbf{Y}^{+\mathrm{T}}-\boldsymbol{\Pi}\mathbf{X}^{+}(d\mathbf{N})\mathbf{Y}^{+\mathrm{T}}\end{aligned}\tag{45}$$

and the analogs of equations (29), (33), (42) and (43) follow by straightforward substitutions. It is easily verified that a variance reduction results from adding the two matrices, since the diagonalized matrix in equation (9) is now a 'ratio' of matrices that have correlated errors. (For more details see Reference 14.)

Multiple calibration samples

The corresponding equations for the case where multiple **N**-matrices are available are conveniently worked out by introducing the pertinent expression for **Q** in equation (45).

GENERALIZATION OF WILSON *ET AL.*

The consequences of the generalization of Wilson *et al.*[17] are easily recognized by observing that essentially the same eigenvalue problem is solved as in the previous methods.[14,16] The only difference lies in the way the left and right singular vectors are estimated. In the current method this is done by decomposing the column and row augmented matrices

$$\tilde{\mathbf{Q}}_{\mathrm{c}}=(\tilde{\mathbf{N}}|\tilde{\mathbf{M}}) \quad \text{and} \quad \tilde{\mathbf{Q}}_{\mathrm{r}}=\begin{pmatrix}\tilde{\mathbf{N}}\\ \overline{\tilde{\mathbf{M}}}\end{pmatrix}$$

as

$$\tilde{\mathbf{Q}}_{\mathrm{c}}=\hat{\mathbf{U}}\hat{\boldsymbol{\Theta}}_{\mathrm{c}}\hat{\mathbf{V}}_{\mathrm{c}}^{\mathrm{T}} \tag{46}$$

$$\tilde{\mathbf{Q}}_{\mathrm{r}}=\hat{\mathbf{U}}_{\mathrm{r}}\hat{\boldsymbol{\Theta}}_{\mathrm{r}}\hat{\mathbf{V}}^{\mathrm{T}} \tag{47}$$

It is seen that the matrices $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$ should span the common row and column space of both $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{N}}$. Next, $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{N}}$ are converted to square matrices $\hat{\mathbf{M}}_A^{UV}$ and $\hat{\mathbf{N}}_A^{UV}$ by projecting them onto the first A columns of $\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$ as

$$\hat{\mathbf{M}}_A^{UV}=\hat{\mathbf{U}}_A^{\mathrm{T}}\tilde{\mathbf{M}}\hat{\mathbf{V}}_A \tag{48}$$

$$\hat{\mathbf{N}}_A^{UV}=\hat{\mathbf{U}}_A^{\mathrm{T}}\tilde{\mathbf{N}}\hat{\mathbf{V}}_A \tag{49}$$

and the following eigenvalue problem is solved:

$$\hat{\mathbf{N}}_A^{UV}\hat{\mathbf{T}}=\hat{\mathbf{M}}_A^{UV}\hat{\mathbf{T}}\hat{\boldsymbol{\Pi}} \tag{50}$$

which can be recast into equation (12) using the reconstruction equations for the pure profiles.[14] In the original paper of Wilson *et al*. the matrices $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{N}}$ are subject to the projection, but the procedure can also be applied with $\tilde{\mathbf{M}}+\tilde{\mathbf{N}}$ and $\tilde{\mathbf{M}}$.[20] It follows that error propagation leads to the same equations as obtained above. There is, however, a subtle difference that is only noticed if the expressions are to be evaluated, i.e. when the unknown quantities are to be replaced by their estimates (see remark under equation (35)). Using the augmented matrices for the estimation of the left and right singular vectors ($\hat{\mathbf{U}}$ and $\hat{\mathbf{V}}$) may lead to more precise estimates for the elution profiles and spectra ($\hat{\mathbf{X}}$ and $\hat{\mathbf{Y}}$). This translates, for example, to more precise estimates for $\boldsymbol{\xi}_a$ and $\boldsymbol{\eta}_a$ in equation (33). As a consequence, evaluating this expression may lead to more precise estimates for the standard errors (i.e. $\hat{\sigma}(\hat{\pi}_a)$) when applying the method of Wilson *et al*. than when applying Lorber's method. The same argument holds for equations (42) and (43) and their analogs obtained for the generalization of Sanchez and Kowalski.

In summary, applying the generalization of Wilson *et al*. is of direct consequence for the precision of the reconstructed profiles ($\hat{\mathbf{X}}$ and $\hat{\mathbf{Y}}$). The consequences for the precision of the estimated standard error in the estimated eigenvalues ($\hat{\sigma}(\hat{\pi}_a)$) are merely indirect. Since analyzing $\tilde{\mathbf{M}}+\tilde{\mathbf{N}}$ and $\tilde{\mathbf{M}}$ always leads to a variance reduction (see remark under equation (45)), it seems best to use the hybrid method of Poe and Rutan[20] where $\tilde{\mathbf{M}}+\tilde{\mathbf{N}}$ and $\tilde{\mathbf{M}}$ are subject to the projections proposed by Wilson *et al.*

DISCUSSION

The foregoing gives a detailed account of the error propagation through the eigenvalue problem of GRAM. Some general comments with respect to the derived standard errors are in order here.

Comparison of 'old' and 'new' expressions

A comparison between the 'old' equation (33) and the 'new' equation (43) shows that the only difference consists of the term that accounts for the errors in the modulating order. Depending on the relative size of this term, the improvement may be negligible, thereby making the derivation a useless exercise in error propagation. However, the simulations of Booksh and Kowalski[9] have clearly demonstrated that the uncertainty in the estimated concentration ratios is more dependent on the first term of equation (43) than on the second term. This means that estimating the uncertainty in the quantitative results of GRAM using equation (33) alone is inadequate in many applications. The added value of this paper then should be that instead of the coefficients β_1 and β_2 of equation (44), one has expressions that relate data characteristics (figures of merit) to the uncertainty in estimated parameters. This result is, for example, important with respect to considerations of experimental design. Equation (42) is more difficult to interpret than equation (43), which is derived using an additional simplification. It is expected that qualitative statements based on equation (43) should still be valuable under more general error models.

Validation of the derived standard errors

No validation study will be presented in this paper in order to demonstrate the adequacy of the derived expressions. There are two reasons for not doing, for example, additional simulations. The first reason is that all new expressions presented here form a generalization of expressions that have been shown to work well.[14] The second reason is that the functional form of some of the newly derived standard errors is in accordance with previously published simulation results.[9] Since these expressions arise from a simplification of a more general form, it will be assumed that the presented derivations are valid within the assumptions and approximations made.

Assumptions and approximations

Three basic assumptions have been made in order to arrive at the final results. The first assumption concerns the model equations that are assumed to describe the errorless data, i.e. equations (1) and (2). It was assumed herein that the so-called bilinear model holds. If the bilinear model does not hold, the logical step is to move to other methods, e.g. non-bilinear rank annihilation (NBRA)[21] or residual bilinearization (RBL).[22] Furthermore, it is assumed that the pure profiles **X** and **Y** are the same for both samples. This may be a restrictive assumption in practice. For example, chromatographic data are characterized by synchronization problems, since the retention times are not absolute. It is important to note that Poe and Rutan[20] have shown that the hybrid method mentioned above is more sensitive to model errors than the generalization of Sanchez and Kowalski. The influence of model errors is not considered in this paper.

The second assumption concerns the truncated SVD that is used to construct the solution space. The truncated SVD gives an optimal approximation of a matrix if the residuals are homoscedastic and uncorrelated. If the errors do not follow this ideal behavior, the data points have to be weighted or scaled according to their esimated uncertainty. Since in this paper the focus is on a more realistic description of the effects of measurement noise, this topic will be discussed in more detail below.

The third assumption made is that the effect of the measurement noise can be quantified by first-order error propagation. Performing the derivation to first order is primarily a matter of convenience. The results will, however, only be satisfactory if the measurement errors are *sufficiently small*. For many modern instruments the signal-to-noise ratio is excellent and results obtained by the first-order approximation should work well. The validity of this assumption has previously been confirmed by simulations.[14]

Scaling prior to singular value decomposition

The problem of scaling prior to SVD has been recognized by Cochran and Horne in the context of pseudorank estimation.[23] They found that in the presence of heteroscedastic measurement errors, additional PCs are deemed significant. This means that reconstructing the data matrix using the 'correct' number of PCs will not lead to a satisfactory fit of the data and subsequently solving the GRAM eigenvalue problem will not give optimal results. The issue of scaling has been treated in detail by Paatero and Tapper,[24] who argue than an *ideal scaling* of the data matrix is not always feasible. Obviously, an ideal scaling of the data comes down to weighting each data point by its estimated uncertainty, i.e. a matrix element M_{ij} is replaced by $M_{ij}/\sigma(M_{ij})$. However, there is a disadvantage associated with this procedure. In general there will be no useful connection between the SVD of the raw and the scaled data matrix. Paatero and Tapper show that this will be the case if the rank of the matrix of estimated uncertainties, $\sigma(\mathbf{M})$, is larger than one.[24] Thus, if as an additional criterion for a successful scaling strategy one demands that there be such a connection, a conflict might arise. Unfortunately, this is the case if the SVD is used for calculating the ingredients for the GRAM eigenvalue problem (and in general for SVD- or PCA-based curve resolution). As a promising alternative, Paatero and Tapper propose a procedure which they call *balanced scaling*. In this procedure the weight matrix $\mathbf{W}$, with elements $W_{ij}=\sigma(M_{ij})^{-1}$, is decomposed as an outer product of two vectors which are subsequently used to build two diagonal scaling matrices $\mathbf{D}_\mathrm{l}$ and $\mathbf{D}_\mathrm{r}$. The scaled matrix $\tilde{\mathbf{M}}^\mathrm{scl}$ is calculated as

$$\tilde{\mathbf{M}}^\mathrm{scl}=\mathbf{D}_\mathrm{l}\tilde{\mathbf{M}}\mathbf{D}_\mathrm{r} \tag{51}$$

Using equation (51), the connection between the decompositions of $\tilde{\mathbf{M}}$ and $\tilde{\mathbf{M}}^\mathrm{scl}$ is easily established. In the procedure of Paatero and Tapper the diagonals of the scaling matrices are found by an alternating least squares (ALS) algorithm (rank one) which is heuristic in nature. An alternative would be to perform an SVD on the matrix $\mathbf{W}$ and use the first score and loading to build the scaling matrices in an analogous manner. Denote the first score and loading vectors by $\mathbf{f}$ and $\mathbf{g}$ respectively. Then the individual weights are approximated by

$$\hat{W}_{ij}=f_i g_j \tag{52}$$

and the diagonal elements of the scaling matrices are found as $D_{\mathrm{l},ii}=f_i$ $(i=1,\ldots,I)$ and $D_{\mathrm{r},jj}=g_j$ $(j=1,\ldots,J)$ respectively. The motivation for this modified procedure is as follows. With the current assumption about the measurement errors it is reasonable to assume that the matrix $\sigma(\mathbf{M})$ looks like a data matrix itself. (This will certainly be the case if the uncertainty in the modulating order is proportional: see Figure 1 in Reference 9.) Furthermore, in many applications of GRAM one has highly overlapping signal contributions of the individual substituents. For these multicomponent systems it is common to have one very large eigenvalue, which roughly accounts for the average of the variation in the data. (In Reference 13, examples are given where the first PC accounts for more than 90% of the total variation.) It is therefore reasonable to assume that in many situations the matrix $\sigma(\mathbf{M})$ can very well be estimated by the first PC and the same will then hold for the weight matrix $\mathbf{W}$. Investigating the merits of this assertion is certainly a topic for future research. It is important to note that the same scaling is applied to both matrices in order to make GRAM work. Finally, if scaling is applied, the expression for the estimated standard error has to be adapted in a straightforward fashion. Now one has to introduce the uncertainties in the scaled matrices, given by $\hat{W}_{ij}*\sigma(M_{ij})$ and $\hat{W}_{ij}*\sigma(N_{ij})$. Estimates for the true profiles are found by undoing the scaling of the profiles that are initially found.

Consequences for the estimated standard error in the reconstructed profiles and the estimated bias in the estimated eigenvalues

Expressions have been derived for the estimated standard error in the reconstructed profiles and the estimated bias in the estimated eigenvalues under the assumption of homoscedastic and uncorrelated measurement noise.[14] The consequences for these quantities are trivial if the scaling procedure discussed above is successful.

CONCLUSIONS

Expressions for predicting the standard error in the eigenvalues estimated by GRAM have been derived using realistic assumptions about the measurement errors. No simulation results are presented to illustrate the validity of the derived expressions, because the functional shape of these expressions (see equation (43)) is in accordance with previously published simulation results.

The consequence of allowing for more flexibility with respect to the assumptions is the need for more detailed information about the measurement errors if the expression is to be evaluated in practice. One cannot expect to obtain an error estimate for the estimated eigenvalues (concentration ratios) without knowing the noise characteristics of the instrument to a certain extent. Measuring instrumental errors and testing the practical usefulness of the derived expressions are the subject of future research in our laboratory. Applications currently under study include measurement of the errors associated with the recently developed fiber optic heavy metal sensor[25] and the flow probe sensor.[26]

Finally, it can be shown that there is a close relationship between different variations of GRAM and the multivariate regression technique known as principal covariates regression (PCovR).[27] In PCovR an eigenvalue problem is solved to calculate scores that reconstruct both the regressor matrix **X** as well as the regressand matrix **Y** as well as possible. Flexibility is introduced by allowing for a different weighting of **X** and **Y**. In this way PCovR builds a continuum of methods between multiple linear regression (MLR) and principal component regression (PCR). It should be straightforward to extend the derivations given in this paper to PCovR.

ACKNOWLEDGEMENTS

This work was supported by the Center for Process Analytical Chemistry (CPAC), a National Science Foundation Industry/University Cooperative Research Center at the University of Washington. The authors appreciate suggestions made by the reviewers to improve the presentation.

REFERENCES

1. C.-N. Ho, G. D. Christian and E. R. Davidson, *Anal. Chem.* **50**, 1108 (1978).
2. C.-N. Ho, G. D. Christian and E. R. Davidson, *Anal. Chem.* **52**, 1071 (1980).
3. C.-N. Ho, G. D. Christian and E. R. Davidson, *Anal. Chem.* **53**, 92 (1981).
4. A. Lorber, *Anal. Chim. Acta*, **164**, 293 (1984).
5. E. Sanchez and B. R. Kowalski, *Anal. Chem.* **58**, 496 (1986).
6. S. E. Leurgans, R. T. Ross and R. B. Abel, *SIAM J. Matrix Anal. Appl.* **14**, 1064 (1993).
7. H. A. L. Kiers and A. K. Smilde, *J. Chemometrics*, **9**, 179 (1995).
8. I. Scarminio and M. Kubista, *Anal. Chem.* **65**, 409 (1993).
9. K. Booksh and B. R. Kowalski, *J. Chemometrics*, **8**, 45 (1994).
10. C. J. Appelof and E. R. Davidson, *Anal. Chim. Acta*, **146**, 9 (1983).
11. E. Sanchez, *Ph.D. Dissertation*, University of Washington, Seattle (1987).
12. E. R. Malinowski, *Factor Analysis in Chemistry*, Wiley, New York (1991).
13. N. M. Faber, L. M. C. Buydens and G. Kateman, *J. Chemometrics*, **7**, 495 (1993).
14. N. M. Faber, L. M. C. Buydens and G. Kateman, *J. Chemometrics*, **8**, 181 (1994).
15. B. C. Mitchell and D. S. Burdick, *Chemometrics Intell. Lab. Syst.* **20**, 149 (1993).
16. N. M. Faber, L. M. C. Buydens and G. Kateman, *J. Chemometrics*, **8**, 147 (1994).
17. B. E. Wilson, E. Sanchez and B. R. Kowalski, *J. Chemometrics*, **3**, 493 (1989).
18. J. R. Magnus and H. Neudecker, *Matrix Differential Calculus with Applications in Statistics and Econometrics*, Wiley, Chichester (1988).
19. E. Sanchez and B. R. Kowalski, *J. Chemometrics*, **4**, 29 (1990).

20. R. B. Poe and S. C. Rutan, *Anal. Chim. Acta*, **283**, 845 (1993).
21. B. E. Wilson and B. R. Kowalski, *Anal. Chem.* **61**, 2277 (1989).
22. J. Öhman, P. Geladi and S. Wold, *J. Chemometrics*, **4**, 135 (1990).
23. R. N. Cochran and F. H. Horne, *Anal. Chem.* **49**, 846 (1977).
24. P. Paatero and U. Tapper, *Chemometrics Intell. Lab. Syst.* **18**, 183 (1993).
25. Z. Lin, K. S. Booksh, L. W. Burgess and B. R. Kowalski, *Anal. Chem.* **66**, 2552 (1994).
26. K. S. Booksh, Z. Lin, Z. Wang and B. R. Kowalski, *Anal. Chem.* **66**, 2561 (1994).
27. S. de Jong and H. A. L. Kiers, *Chemometrics Intell. Lab. Syst.* **14**, 155 (1992).