

Second-order calibration: bilinear least squares regression and a simple alternative

Marie Linder^{*}, Rolf Sundberg

Mathematical Statistics, Stockholm University, S-106 91 Stockholm, Sweden

Received 6 June 1997; accepted 12 February 1998

Abstract

We consider calibration of second-order, or hyphenated instruments generating bilinear two-way data for each specimen. The bilinear regression model is to be estimated from a number of specimens of known composition. We propose a simple estimator and study how it works on real and simulated data. The estimator, which we call the SVD (singular value decomposition) estimator is usually not much less efficient than bilinear least squares. The advantages of our method over bilinear least squares are that it is faster and more easily computed, its standard errors are explicit (and derived in the paper), and it has a simpler correlation structure. © 1998 Elsevier Science B.V. All rights reserved.

Keywords: Multivariate calibration; Bilinear regression; Hyphenated instruments; Matrix data; Multilinearity; Singular value decomposition (SVD); SVD estimator

1. The bilinear model

A hyphenated instrument, consisting of two instruments in series, yields a matrix of data for each specimen. Given a number of calibration specimens with several constituents of known concentration, how do we calibrate such instruments? How do we resolve the pure response of each constituent?

We will use LC–UV as a model example. Each pure constituent is characterized by an elution profile and a UV-spectrum. The LC separation gives rise to a new spectrum for each chromatographic time, i.e., we get a matrix of absorbances for each specimen when we collect the response in the UV-instrument. This matrix will be indexed by time (from the LC profile) and wavelength (from the UV spectra). We use the notation $z_{ij}^{(k)}$ for the absorbance z at time i and wavelength j in specimen k ; constituent is indexed by r . We prefer the word specimen, instead of sample, because they are thought of as generated from designed mixtures of known concentrations and not from random sampling.

We assume proportionality in concentration for each of the instruments separately. For a UV instrument this can be motivated by Lambert–Beer's law. For an LC instrument the elution peak area is expected to increase linearly with (moderately) increasing concentration of the constituent. Examples of other instrument combina-

^{*} Corresponding author.

tions which yield bilinear data matrices are excitation-emission fluorescence instruments and LC- or GC-MS, whereas MS-MS and 2D-NMR do not yield bilinear data matrices. We also assume that responses for both instruments are additive over constituents, in the sense that the response for a mixture is the sum of the individual constituent responses. From above we note that the concentrations must not be too high for proportional linearity to be realistic. The final assumption behind our model is that the background has been separately eliminated. If the background had not been eliminated, we would have to add some type of intercept-parameter to our model to make it valid.

From the assumption of linearity in each instrument response, we conclude that data should follow a bilinear regression model. In its most intuitive form, this model is written as $\mathbf{Z}^{(k)} = \sum_{r=1}^R c_{rr}^{(k)} \boldsymbol{\alpha}_r \boldsymbol{\beta}_r^T$, with elements $Z_{ij}^{(k)} = \sum_r \alpha_{ir} c_{rr}^{(k)} \beta_{jr}$ for specimen k . In the LC-UV example $c_{rr}^{(k)}$ is the concentration of constituent r in specimen k , $\boldsymbol{\alpha}_r$ is the elution profile of constituent r , and $\boldsymbol{\beta}_r$ is the spectrum of constituent r . The model can be written in condensed matrix form as

$$\text{Model for data: } \mathbf{Z}^{(k)} = \mathbf{A} \mathbf{C}^{(k)} \mathbf{B}^T + \mathbf{E}^{(k)}, \quad (1)$$

where $\mathbf{Z}^{(k)}$ is the observed data matrix for specimen k , \mathbf{A} is the pure response of the first instrument, $\mathbf{C}^{(k)}$ is a diagonal matrix with the concentrations for calibration specimen k on the diagonal, \mathbf{B} is the pure response of the second instrument and $\mathbf{E}^{(k)}$ is the noise matrix for specimen k . At present we make no specific assumptions about the noise, in Section 3 it will be assumed uncorrelated and homoscedastic. For an example of simulated data matrices, see Fig. 1.

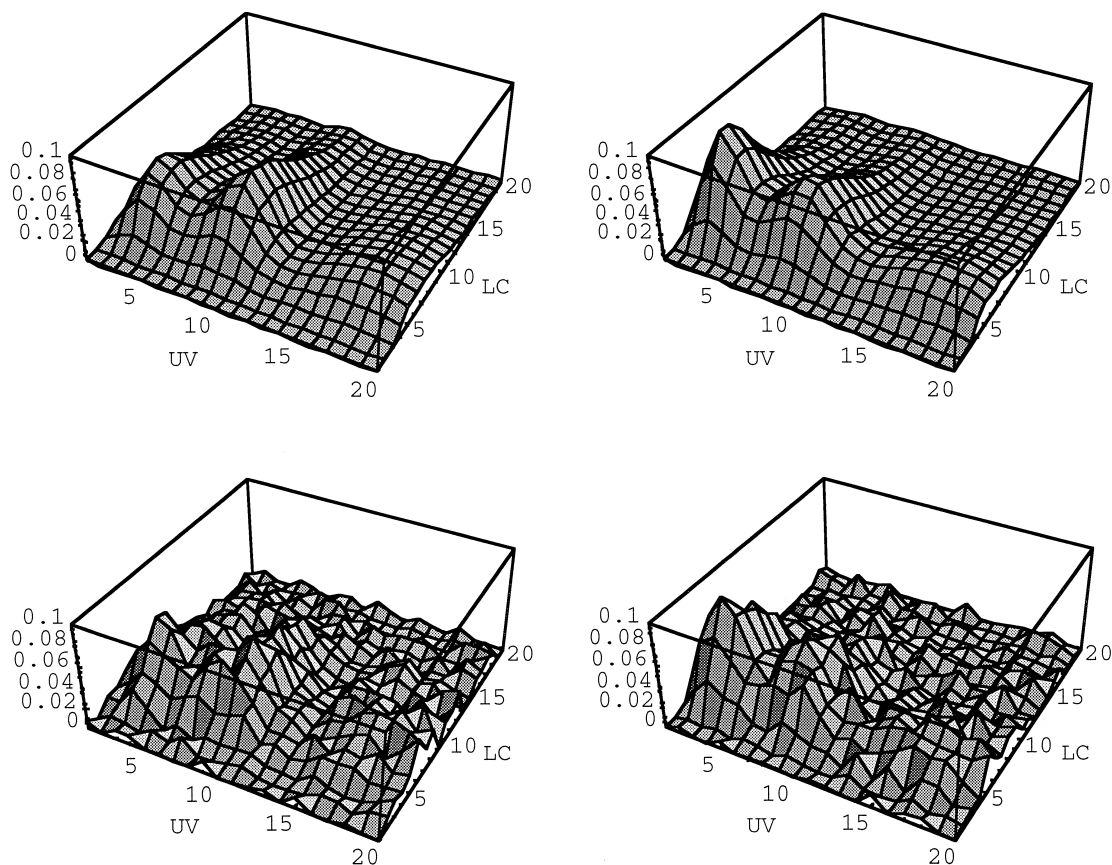


Fig. 1. Pure and noisy data structures for two specimens with the same constituents but with different concentrations.

The model description above is not unique since multiplication of α_r and division of β_r by the same scalar will not change their product $\alpha_r \beta_r^T$. In order to make it unique in a symmetric way we write $\mathbf{Z}^{(k)} = \sum_{r=1}^R c_{rr}^{(k)} \gamma_{rr} \alpha_r \beta_r^T$, or equivalently

$$\text{Symmetrically constrained model: } \mathbf{Z}^{(k)} = \mathbf{A} \mathbf{C}^{(k)} \mathbf{\Gamma} \mathbf{B}^T + \mathbf{E}^{(k)}, \quad |\alpha_r| = |\beta_r| = 1, \quad (2)$$

where columns α_r and β_r for constituent r of matrices \mathbf{A} and \mathbf{B} represent the profiles from the first and the second instrument, respectively, normalized to unit length. Here $\mathbf{\Gamma}$ is a diagonal matrix of scale parameters with elements γ_{rr} , the products of the original lengths of α_r and β_r . We need to estimate the parameters $\{\mathbf{A}, \mathbf{B}, \mathbf{\Gamma}\}$ of the calibration model, which we can then use for future predictions. One possibility is bilinear least squares (BLLS), which we try in this paper. We also present an alternative, a simple estimator, which we call the singular value decomposition (SVD) estimator. This estimator is formed by singular value decomposition of suitably weighted sums of the observed data matrices and can be interpreted as obtained from reweighted least squares normal equations. For the SVD estimator we can explicitly write down the variance–covariance matrix. We make numerical comparisons of efficiency with the BLLS estimator, and try both methods on a real data example.

Among alternative methods for solving the bilinear regression problem, we first note the Generalized Rank Annihilation Method (GRAM), see Sanchez and Kowalski [1], even though it is based on just one calibration specimen which is compared with one unknown specimen. Other methods are based on the PARAFAC model, which is obtained by regarding also the concentration matrices $\mathbf{C}^{(k)}$ in model 1 as unknown parameters, see Bro [2] for a review. The parameters in this model can be estimated by, for example, Alternating Least Squares (ALS) or by the methods known as TriLinear Decomposition methods (TLD), see Sanchez and Kowalski [3] or Burdick et al. [4]. Note that trilinear here means a bilinear instrument with more than one specimen used. TLD somewhat arbitrarily uses the GRAM idea on two selected specimens (or two selected linear combinations). The estimated concentrations in the PARAFAC model are finally regressed on the known concentrations to build a calibration model. PLS-based methods have been proposed and used by Wold et al. [5] and Bro [6]. Wold et al. use unfolding to transform three-way data to two-way data, whereas Bro refrains from that. Bro's method has recently been discussed by Smilde [7] and de Jong [8]. For more general views on bilinear (and multilinear) models and methods for decompositions of such data sets, see Geladi [9] and Leurgans and Ross [10]. Not much is known about how different bilinear (multilinear) methods compare in precision.

2. Estimation procedures

We assume that the calibration is done with at least as many specimens as constituents. This condition is not really necessary for the quantitative prediction of the same constituents in a new specimen, as shown by the GRAM method, and it is not always necessary for the existence of a unique solution to the estimation problem, see Kruskal [11].

2.1. The BLLS estimator

The least squares method for fitting a model to data, be it linear or nonlinear, requires that the sum of squared deviations between measured and expected responses is made as small as possible. An ideal situation for least squares is that the noise is normally distributed, uncorrelated and homoscedastic. Then the least squares method is equivalent to the maximum likelihood method. To obtain the least squares solution we minimize the residual sum of squares with respect to the parameters. This can be done by differentiation with respect to the parameters. Equating these derivatives to zero we get the normal equations. Solving them yields the parameter estimates.

The residual sum of squares for bilinear models and three-dimensional arrays is naturally expressed by summing over specimens and using the trace of the resulting matrix, or in formulas

$$\text{Tr} \left(\sum_{k=1}^K (\mathbf{Z}^{(k)} - \mathbf{A}\mathbf{C}^{(k)}\mathbf{\Gamma}\mathbf{B}^T)^T (\mathbf{Z}^{(k)} - \mathbf{A}\mathbf{C}^{(k)}\mathbf{\Gamma}\mathbf{B}^T) \right). \quad (3)$$

This has to be minimized with respect to $\{\mathbf{A}, \mathbf{B}, \mathbf{\Gamma}\}$, under the constraints $|\boldsymbol{\alpha}_r| = |\boldsymbol{\beta}_r| = 1$ from the assumed normalization. Using Lagrange's multiplier method for minimization under constraints we find the normal equations

$$\begin{aligned} \gamma_{rr} \mathbf{A}\mathbf{D}^{(r)}\mathbf{\Gamma}\mathbf{B}^T\boldsymbol{\beta}_r - \gamma_{rr}^2 d_{rr}^{(r)}\boldsymbol{\alpha}_r - \gamma_{rr} \mathbf{T}^{(r)}\boldsymbol{\beta}_r - \boldsymbol{\lambda}_r \boldsymbol{\alpha}_r &= 0 \\ \gamma_{rr} \mathbf{B}\mathbf{\Gamma}\mathbf{D}^{(r)}\mathbf{A}^T\boldsymbol{\alpha}_r - \gamma_{rr}^2 d_{rr}^{(r)}\boldsymbol{\beta}_r - \gamma_{rr} (\mathbf{T}^{(r)})^T \boldsymbol{\alpha}_r - \boldsymbol{\mu}_r \boldsymbol{\beta}_r &= 0 \\ \boldsymbol{\beta}_r^T \mathbf{B}\mathbf{\Gamma}\mathbf{D}^{(r)}\mathbf{A}^T\boldsymbol{\alpha}_r - \boldsymbol{\alpha}_r^T \mathbf{T}^{(r)}\boldsymbol{\beta}_r &= 0, \end{aligned}$$

where $\boldsymbol{\lambda}_r$ and $\boldsymbol{\mu}_r$ are the Lagrangian multipliers, and $\mathbf{T}^{(r)}$ and $\mathbf{D}^{(r)}$ are defined by

$$\mathbf{T}^{(r)} = \sum_{k=1}^K c_{rr}^{(k)} \mathbf{Z}^{(k)}, \quad \mathbf{D}^{(r)} = \sum_{k=1}^K c_{rr}^{(k)} \mathbf{C}^{(k)}.$$

The matrix statistics $\mathbf{T}^{(r)}$ and the coefficient matrices $\mathbf{D}^{(r)}$ are of great importance for the sequel. They are weighted sums of data matrices and concentration matrices, respectively. Note that each constituent, $r = 1, \dots, R$ requires its own weights, namely its concentrations in the different specimens. The matrices $\mathbf{T}^{(r)}$ carry all information in data about the bilinear structure, in a sense that may be specified in terms of the concept of sufficiency.

After simplification of the normal equations we get

Result 2.1: Normal equations for BLLS.

The normal equations, over $r = 1, \dots, R$, can be written:

$$\begin{aligned} \mathbf{A}\mathbf{D}^{(r)}\mathbf{\Gamma}\mathbf{B}^T\boldsymbol{\beta}_r &= \mathbf{T}^{(r)}\boldsymbol{\beta}_r \\ \mathbf{B}\mathbf{\Gamma}\mathbf{D}^{(r)}\mathbf{A}^T\boldsymbol{\alpha}_r &= (\mathbf{T}^{(r)})^T \boldsymbol{\alpha}_r. \end{aligned} \quad (4)$$

Remark: In the special case when we have only one constituent, the matrices \mathbf{A} and \mathbf{B} reduce to vectors and the diagonal matrices $\mathbf{\Gamma}$ and \mathbf{D} reduce to scalars. Using the property $|\boldsymbol{\alpha}| = |\boldsymbol{\beta}| = 1$, the system of normal equations reduces to the eigen-problem

$$\begin{aligned} \mathbf{T}\mathbf{T}^T\boldsymbol{\alpha} &= (\gamma d)^2 \boldsymbol{\alpha} \\ \mathbf{T}^T\mathbf{T}\boldsymbol{\beta} &= (\gamma d)^2 \boldsymbol{\beta}. \end{aligned}$$

This reduction cannot be performed when we have more than one constituent. The following argument explains why. For one constituent the term $\mathbf{A}^T\mathbf{A} = \boldsymbol{\alpha}^T\boldsymbol{\alpha} = 1$, but more generally $\mathbf{A}^T\mathbf{A}$ is a matrix in which only the diagonal elements are fixed and equal to one. The unknown off-diagonal elements will typically not be zero, since they represent the angles between the profile vectors for the different constituents. ■

The system of normal equations can be solved iteratively in analogy with the PARAFAC model. The following alternating least squares procedure, with starting values for either \mathbf{A} or \mathbf{B} , is natural and usually works slowly

but safely. We express \mathbf{A} as a function of \mathbf{B} and vice versa in the following reformulation of Eq. (4), cf. Model Eq. (1):

$$\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \vdots \\ \boldsymbol{\alpha}_R \end{pmatrix} = \begin{pmatrix} d_{11} \boldsymbol{\beta}_1^T \boldsymbol{\beta}_1 & \cdots & d_{1R} \boldsymbol{\beta}_1^T \boldsymbol{\beta}_R \\ \vdots & \ddots & \vdots \\ d_{R1} \boldsymbol{\beta}_R^T \boldsymbol{\beta}_1 & \cdots & d_{RR} \boldsymbol{\beta}_R^T \boldsymbol{\beta}_R \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{T}^{(1)} \boldsymbol{\beta}_1 \\ \vdots \\ \mathbf{T}^{(R)} \boldsymbol{\beta}_R \end{pmatrix} \quad (5)$$

$$\begin{pmatrix} \boldsymbol{\beta}_1 \\ \vdots \\ \boldsymbol{\beta}_R \end{pmatrix} = \begin{pmatrix} d_{11} & \cdots & d_{1R} \boldsymbol{\alpha}_1^T \boldsymbol{\alpha}_R \\ \vdots & \ddots & \vdots \\ d_{R1} \boldsymbol{\alpha}_R^T \boldsymbol{\alpha}_1 & \cdots & d_{RR} \end{pmatrix}^{-1} \begin{pmatrix} (\mathbf{T}^{(1)})^T \boldsymbol{\alpha}_1 \\ \vdots \\ (\mathbf{T}^{(R)})^T \boldsymbol{\alpha}_R \end{pmatrix}. \quad (6)$$

Here we have incorporated γ_{rr} in $\boldsymbol{\beta}_r$ and only assume $\boldsymbol{\alpha}_r$ normalized to unit length in each step, for uniqueness. The iteration procedure works by alternating updating of \mathbf{A} from Eq. (5) and \mathbf{B} from Eq. (6). When the iteration procedure has converged we get the symmetrically constrained solution by setting $\hat{\gamma}_{rr} = |\hat{\boldsymbol{\beta}}_r|$ and renormalizing $\hat{\boldsymbol{\beta}}_r$ by its length. This BLS algorithm has only linear convergence and can be quite slow when the calibration is imprecise, so that the function (3) is relatively flat in some directions from the minimum.

2.2. The SVD estimator

We now introduce and motivate our proposed alternative to the BLS estimator, the SVD (single value decomposition) estimator.

Result 2.2: Construction of the SVD estimator.

The SVD estimator is constructed from singular value decomposition of weighted data matrices by

$$\hat{\gamma}_{rr} \hat{\boldsymbol{\alpha}}_r \hat{\boldsymbol{\beta}}_r^T = \text{SVD}_1 \left[\sum_{r_1=1}^R (\mathbf{D}^{-1})_{rr_1} \mathbf{T}^{(r_1)} \right] = \mathbf{u} \mathbf{v}^T,$$

which yields $\hat{\boldsymbol{\alpha}}_r = \mathbf{u}$, $\hat{\boldsymbol{\beta}}_r = \mathbf{v}$ and $\hat{\gamma}_{rr} = s$. Here $\text{SVD}_1[\dots]$ means the first singular value component of a matrix to be constructed below.

Justification: To derive and motivate this estimator we start from the same matrices $\mathbf{T}^{(r)}$ and $\mathbf{D}^{(r)}$ as in the least squares procedure. As mentioned in Section 2.1, the matrices $\mathbf{T}^{(r)}$ carry all information in data about the bilinear structure, in a sense that may be specified in terms of the concept of sufficiency. First, note that

$$\mathbf{T}^{(r)} = \sum_{k=1}^K c_{rr}^{(k)} \mathbf{Z}^{(k)} = \sum_{k=1}^K c_{rr}^{(k)} (\mathbf{A} \mathbf{C}^{(k)} \boldsymbol{\Gamma} \mathbf{B}^T + \mathbf{E}^{(k)}) = \mathbf{A} \mathbf{D}^{(r)} \boldsymbol{\Gamma} \mathbf{B}^T + \sum_{k=1}^K c_{rr}^{(k)} \mathbf{E}^{(k)},$$

so $\mathbf{T}^{(r)}$ has the expected value $\mathbf{A} \mathbf{D}^{(r)} \boldsymbol{\Gamma} \mathbf{B}^T = \sum_{r_1=1}^R d_{rr_1} \gamma_{r_1 r_1} \boldsymbol{\alpha}_{r_1} \boldsymbol{\beta}_{r_1}^T$. Now we simply identify $\mathbf{T}^{(r)}$ with its expected value. In matrix-like form for all constituents simultaneously, where the vectors have matrix elements, this yields the estimating equations

$$\begin{pmatrix} \mathbf{T}^{(1)} \\ \vdots \\ \mathbf{T}^{(R)} \end{pmatrix} = \begin{pmatrix} d_{11} & \cdots & d_{1R} \\ \vdots & \ddots & \vdots \\ d_{R1} & \cdots & d_{RR} \end{pmatrix} \begin{pmatrix} \gamma_{11} \boldsymbol{\alpha}_1 \boldsymbol{\beta}_1^T \\ \vdots \\ \gamma_{RR} \boldsymbol{\alpha}_R \boldsymbol{\beta}_R^T \end{pmatrix},$$

or in brief notation $\mathbf{T} = \mathbf{D} \boldsymbol{\omega}$. Inverting the matrix \mathbf{D} to solve for $\boldsymbol{\omega}$, we obtain $\gamma_{rr} \boldsymbol{\alpha}_r \boldsymbol{\beta}_r^T = (\mathbf{D}^{-1})_{r_1} \mathbf{T}^{(1)} + \dots + (\mathbf{D}^{-1})_{r_R} \mathbf{T}^{(R)}$. Thus if the $\mathbf{T}^{(r)}$ were error-free, the right-hand side would be exactly rank one. Allow-

ing noise in the $\mathbf{T}^{(r)}$ we estimate $\gamma_{rr} \boldsymbol{\alpha}_r \boldsymbol{\beta}_r^T$ by the first component of the singular value decomposition, which is the natural rank one approximation. This means that $\hat{\gamma}_{rr} \hat{\boldsymbol{\alpha}}_r \hat{\boldsymbol{\beta}}_r^T = \text{SVD}_1 \left[(\mathbf{D}^{-1} \mathbf{t})_r^T \right] = \mathbf{u} \mathbf{v}^T$. Since we always choose $\boldsymbol{\alpha}_r$ and $\boldsymbol{\beta}_r$ as vectors of unit length, we get the estimates $\hat{\boldsymbol{\alpha}}_r = \mathbf{u}$, $\hat{\boldsymbol{\beta}}_r = \mathbf{v}$ and $\hat{\gamma}_{rr} = s$ by identifying terms in the two expressions. ■

We propose the following algorithm for calculating the SVD estimates.

SVD algorithm:

$$D_{r_1 r_2} = \sum_{k=1}^K c_{r_1 r_1}^{(k)} c_{r_2 r_2}^{(k)} \quad \text{Compute the matrix } \mathbf{D}.$$

$$\mathbf{T}^{(r)} = \sum_{k=1}^K c_{rr}^{(k)} \mathbf{Z}^{(k)} \quad \text{Compute the vector } \mathbf{t} \text{ of matrices } \mathbf{T}^{(r)}.$$

$$\boldsymbol{\omega} = \mathbf{D}^{-1} \mathbf{t} \quad \text{Invert the matrix } \mathbf{D} \text{ and solve for the vector of matrices } \boldsymbol{\omega}.$$

$$s_{rr} \mathbf{u}_r \mathbf{v}_r^T = \text{SVD}_1(\boldsymbol{\omega}_r) \quad \text{Select the first singular value component for each matrix component of } \boldsymbol{\omega}.$$

$$\hat{\boldsymbol{\alpha}}_r = \mathbf{u}_r, \hat{\boldsymbol{\beta}}_r = \mathbf{v}_r, \hat{\gamma}_{rr} = s_{rr} \quad \text{Identify components in the singular value decomposition.}$$

In the SVD step the singular vectors \mathbf{u}_1 and \mathbf{v}_1 can be rotated to point in the opposite direction, in principle. However, since $\boldsymbol{\alpha}_r$ and $\boldsymbol{\beta}_r$ should have non-negative components there should be only one natural choice. If this is not the case we must question the model for this data set. A few small negative estimated components may be replaced by zeros or kept. Another indication of model invalidity is that the matrices forming $\mathbf{D}^{-1} \mathbf{T}$ are not close to rank one. ■

Remark: With only one constituent present it is easy to show that the SVD and BLS estimators are identical. In the case of two constituents, which is taken to represent the general case, we find for SVD

$$\boldsymbol{\alpha}_1 \text{ and } \boldsymbol{\beta}_1 \text{ are the first singular vectors of } (d_{22} \mathbf{T}^{(1)} - d_{12} \mathbf{T}^{(2)})$$

$$\boldsymbol{\alpha}_2 \text{ and } \boldsymbol{\beta}_2 \text{ are the first singular vectors of } (-d_{21} \mathbf{T}^{(1)} + d_{11} \mathbf{T}^{(2)}).$$

In this way the SVD estimator puts explicit weights on the $\mathbf{T}^{(r)}$ matrices. The weights for the BLS estimator, on the other hand, are parameter dependent as seen from the normal equations, Eq. (4), and therefore only implicitly determined. Hence the BLS estimator is generally not identical with the SVD estimator. ■

3. Precision

For the new SVD estimator explicit, approximate error-bounds can be written down, see below. We will compare with the BLS estimator, for which we use approximate least squares results for the linear case with linear constraints. For this section we assume uncorrelated homoscedastic noise for simplicity. Uncorrelated means lack of correlation between measurements taken for different specimens, at different wavelengths or at different chromatographic times. Assuming that the number of calibration specimens is large, or more realistically that σ is small, so that the estimated model falls near the true model, we may linearize to find the approximate precision.

3.1. The BLS estimator

We briefly indicate how an approximate variance–covariance matrix for the BLS estimator can be computed. Under the above assumption the bilinear model may be linearized in terms of the full parameter $\boldsymbol{\theta} =$

$(\alpha_1, \dots, \alpha_R, \beta_1, \dots, \beta_R, \gamma_{11}, \dots, \gamma_{RR})^T$ in a neighborhood of the true θ that includes the estimate $\hat{\theta}$. If it were not for the constraints on the parameters, the ordinary theory of linear regression models could then be applied. It tells that if we have a linear problem $(y, \mathbf{X}\theta, \sigma^2\mathbf{I})$, where y is the vector of responses, \mathbf{X} is the matrix of explanatory variables, and the noise is homoscedastic and uncorrelated, then the variance of the least-squares estimate $\hat{\theta}$ is $\text{Var}(\hat{\theta}) = \sigma^2(\mathbf{X}^T\mathbf{X})^{-1}$. In order to cope with the parameter constraints, we linearize them as well, to read $\mathbf{H}^T\theta = \eta$, for some matrix \mathbf{H} and vector η . The following result (see section 4a.9 of Ref. [12]) tells how to combine the linearized design matrix \mathbf{X} with the linearized constraints matrix \mathbf{H} to obtain the variance–covariance matrix for θ .

Result 3.1: *Asymptotic variance–covariance matrix of the BLLS estimator.*

The approximate variance–covariance matrix for BLLS is given by

$$\text{Var}(\hat{\theta}) = \sigma^2 \left\{ \begin{pmatrix} \mathbf{X}^T\mathbf{X} & \mathbf{H} \\ \mathbf{H}^T & 0 \end{pmatrix}^{-1} \right\}_{11}$$

The interpretation of this formula is that we first take the inverse of the given matrix and then extract its upper left part, of the same size as $\mathbf{X}^T\mathbf{X}$.

For a mathematical justification of Result 3.1, see Appendix A.

3.2. The SVD estimator

In this section we first (approximately) solve the eigenvector problem induced by the singular value decomposition. We then linearize the expressions for the SVD estimator found in this way and conclude by finding the asymptotic variance–covariance matrix.

Straightforward computation shows that the SVD estimator for the pure response $\gamma_{rr} \alpha_r \beta_r^T$ of constituent r is given by the first term of the singular value decomposition of

$$\sum_{r_1=1}^R (\mathbf{D}^{-1})_{rr_1} \mathbf{T}^{(r_1)} = \gamma_{rr} \alpha_r \beta_r^T + \sum_{k=1}^K \sum_{r_1=1}^R (\mathbf{D}^{-1})_{rr_1} c_{r_1 r_1}^{(k)} \mathbf{E}^{(k)} = \gamma_{rr} \alpha_r \beta_r^T + \xi^{(r)},$$

where $\xi^{(r)}$ is a noise term consisting of a linear combination of the noise matrices $\mathbf{E}^{(k)}$ from the different specimens. From the definition of singular value decomposition, $\hat{\alpha}_r$ and $\hat{\beta}_r$ are also the first eigenvectors u and v of the eigen-problems

$$\left(\gamma_{rr}^2 \alpha_r \alpha_r^T + \gamma_{rr} \alpha_r \beta_r^T (\xi^{(r)})^T + \gamma_{rr} \xi^{(r)} \beta_r \alpha_r^T + \xi^{(r)} (\xi^{(r)})^T \right) u = \lambda u \tag{7}$$

and

$$\left(\gamma_{rr}^2 \beta_r \beta_r^T + \gamma_{rr} \beta_r \alpha_r^T \xi^{(r)} + \gamma_{rr} (\xi^{(r)})^T \alpha_r \beta_r^T + (\xi^{(r)})^T \xi^{(r)} \right) v = \lambda v, \tag{8}$$

respectively. In Appendix B we solve these equations in detail to find that the SVD estimates can be expressed as

$$\begin{aligned} \hat{\alpha}_r &\approx \alpha_r (1 - \alpha_r^T \xi^{(r)} \beta_r / \gamma_{rr}) + \xi^{(r)} \beta_r / \gamma_{rr} \\ \hat{\beta}_r &\approx \beta_r (1 - \alpha_r^T \xi^{(r)} \beta_r / \gamma_{rr}) + (\xi^{(r)})^T \alpha_r / \gamma_{rr} \\ \hat{\gamma}_{rr} &\approx \gamma_{rr} + \alpha_r^T \xi^{(r)} \beta_r, \end{aligned}$$

where $\xi^{(r)} = \sum_{k=1}^K \sum_{r_1=1}^R (\mathbf{D}^{-1})_{rr_1} c_{r_1 r_1}^{(k)} \mathbf{E}^{(k)}$ as before. This linearization of the estimates allows us to calculate variances and covariances. In the case of uncorrected homoscedastic noise the result is particularly simple and is given by the following result.

Result 3.2: Precision of the SVD estimator.

For uncorrelated homoscedastic noise with variance σ^2 the approximate covariances and variances of the SVD estimator are

$$\text{Cov}[\hat{\alpha}_{r_1}, \hat{\alpha}_{r_2}] = \sigma^2 (\mathbf{D}^{-1})_{r_1 r_2} (\mathbf{I} - \alpha_{r_1} \alpha_{r_1}^T) \beta_{r_1}^T \beta_{r_2} (\mathbf{I} - \alpha_{r_2} \alpha_{r_2}^T) / \gamma_{r_1 r_1} \gamma_{r_2 r_2}$$

$$\text{Cov}[\hat{\alpha}_{r_1}, \hat{\beta}_{r_2}] = \sigma^2 (\mathbf{D}^{-1})_{r_1 r_2} (\mathbf{I} - \alpha_{r_1} \alpha_{r_1}^T) \alpha_{r_2} \beta_{r_1}^T (\mathbf{I} - \beta_{r_2} \beta_{r_2}^T) / \gamma_{r_1 r_1} \gamma_{r_2 r_2}$$

$$\text{Cov}[\hat{\alpha}_{r_1}, \hat{\gamma}_{r_2 r_2}] = \sigma^2 (\mathbf{D}^{-1})_{r_1 r_2} (\mathbf{I} - \alpha_{r_1} \alpha_{r_1}^T) \beta_{r_1}^T \beta_{r_2} \alpha_{r_2} / \gamma_{r_1 r_1}$$

$$\text{Cov}[\hat{\beta}_{r_1}, \hat{\beta}_{r_2}] = \sigma^2 (\mathbf{D}^{-1})_{r_1 r_2} (\mathbf{I} - \beta_{r_1} \beta_{r_1}^T) \alpha_{r_1}^T \alpha_{r_2} (\mathbf{I} - \beta_{r_2} \beta_{r_2}^T) / \gamma_{r_1 r_1} \gamma_{r_2 r_2}$$

$$\text{Cov}[\hat{\beta}_{r_1}, \hat{\gamma}_{r_2 r_2}] = \sigma^2 (\mathbf{D}^{-1})_{r_1 r_2} (\mathbf{I} - \beta_{r_1} \beta_{r_1}^T) \alpha_{r_1}^T \alpha_{r_2} \beta_{r_2} / \gamma_{r_1 r_1}$$

$$\text{Cov}[\hat{\gamma}_{r_1 r_1}, \hat{\gamma}_{r_2 r_2}] = \sigma^2 (\mathbf{D}^{-1})_{r_1 r_2} \alpha_{r_1}^T \beta_{r_1}^T \beta_{r_2} \alpha_{r_2},$$

where \mathbf{I} is the identity matrix of suitable size. In particular the variances and the zero covariances within constituent are:

$$\text{Var}[\hat{\alpha}_r] = \sigma^2 (\mathbf{D}^{-1})_{rr} (\mathbf{I} - \alpha_r \alpha_r^T) / \gamma_{rr}^2$$

$$\text{Var}[\hat{\beta}_r] = \sigma^2 (\mathbf{D}^{-1})_{rr} (\mathbf{I} - \beta_r \beta_r^T) / \gamma_{rr}^2$$

$$\text{Var}[\hat{\gamma}_{rr}] = \sigma^2 (\mathbf{D}^{-1})_{rr}$$

$$\text{Cov}(\hat{\alpha}_r, \hat{\beta}_r) = \text{Cov}(\hat{\alpha}_r, \hat{\gamma}_{rr}) = \text{Cov}(\hat{\beta}_r, \hat{\gamma}_{rr}) = 0.$$

Remark: The matrices $\text{Var}(\hat{\alpha}_r)$ and $\text{Var}(\hat{\beta}_r)$ are singular due to the constraints $|\hat{\alpha}_r| = |\hat{\beta}_r| = 1$. For derivation of the variances and covariances, see Appendix B. Note that within constituent many covariances are zero, so the correlation structure is relatively simple. Also covariances between constituents will typically be relatively small. However, all covariances will vanish only if all profiles are mutually orthogonal, a situation not likely to occur in practice. ■

4. A simulation study

We carried out an extensive simulation study to answer the following important questions. How does the SVD estimator perform relative to the BLLS estimator? Do the approximate precision formulas work well with few specimens and large or moderate noise? Can the theoretical formulas be verified empirically? Which factors have the highest influence on the precision of the estimates?

4.1. Design

To set up a design for the simulation study we needed those factors which influence the precision. From Result 3.2 for the SVD estimator we see that the precision depends on the following factors:

- (i) σ , i.e., size of noise in data,
- (ii) \mathbf{D} , i.e., choice of concentrations, number of specimens and quality of the calibration set, and
- (iii) $\{\alpha_r, \beta_r, \gamma_{rr}\}$, i.e., the true profiles, which are influenced by the number of wavelengths, the number of chromatographic times, and the degree of overlap between profiles.

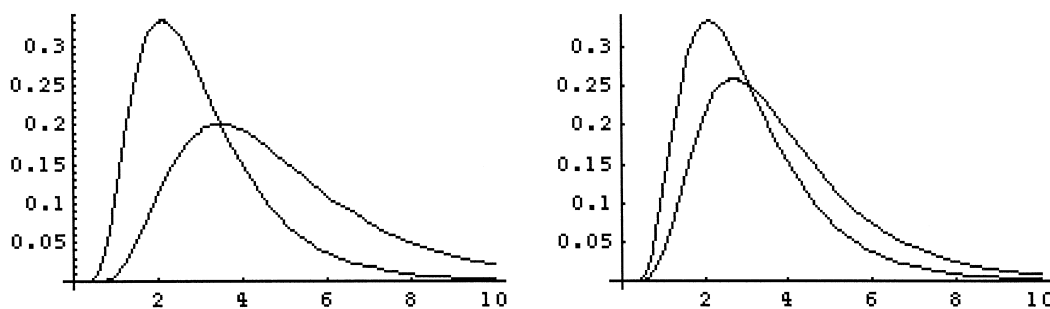


Fig. 2. Factor 4: elution profiles, moderately and much overlapping, respectively.

To illustrate the performance of the SVD estimator versus the BLLS estimator we simulated an LC–UV example with two constituents ($R = 2$). In accordance with the above discussion, we chose to investigate the following factors, at two levels each in a 2^{6-1} factorial design.

Factor	– level	+ level
(1) number of chromatographic times and wavelengths	20	10
(2) number of specimens (K)	4	2
(3) size of noise in data (σ)	0.0025 \approx 10%	0.0075 \approx 30%
(4) degree of overlap in true profiles	moderately overlapping	much overlapping
(5) degree of overlap in true spectra	moderately overlapping	much overlapping
(6) quality of the calibration design	more informative	less informative

Factor 1 determines the size of the data matrix, which is number of times \times number of wavelengths (20×20 or 10×10). For factor 2, the + level was the minimum number of specimens possible to investigate, since we had two constituents, see Section 2. The relative size of noise in data was of magnitude 10% and 30% of the average response, respectively. Figs. 2 and 3 illustrate factors 4 and 5, the form and degree of overlap in the true, pure responses.

Since we had only two constituents, the choice of concentration combinations can be illustrated in the xy -plane, see Fig. 4. For the more informative set (– level) we first chose two points well apart for the $K = 2$ specimens level, and then for each such point (x, y) also the point $(\theta x, \theta y)$ with $\theta = 2$ to form the set for four specimens. The less informative set (+ level) was formed in the same way, but starts with two points closer together. The advantage of this choice is that we can predict the difference in precision between two and four specimens. The \mathbf{D} -matrices with $K = 2$ and $K = 4$, \mathbf{D}_2 and \mathbf{D}_4 are simply related by $\mathbf{D}_4 = (1 + \theta^2)\mathbf{D}_2$. We se-

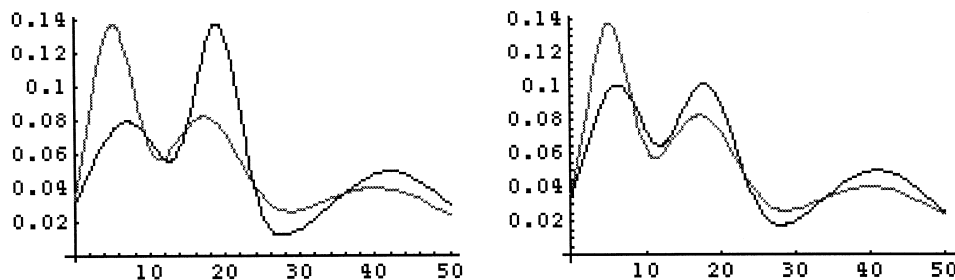


Fig. 3. Factor 5: spectra, moderately and much overlapping, respectively.

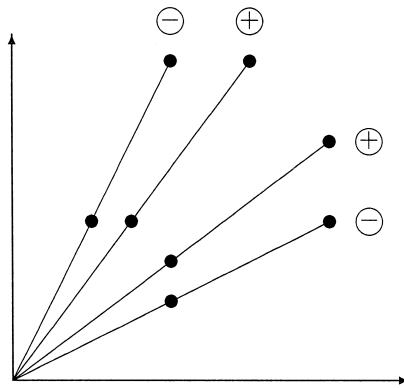


Fig. 4. Calibration designs.

lected the points (1,2) and (2,1) for the informative set and the points (1.5,2) and (2,1.5) for the less informative set, with $\theta = 2$.

The design chosen for the study was a factorial 2^{6-1} design with the usual confounding pattern for half-fractions, that is the highest order interaction aliased with the identity. We made 100 replicates in each point of the factorial experiment, and as much as feasible we let the noise matrices $\mathbf{E}^{(k)}$ be based on the same random numbers over different design points.

4.2. Response

As measures of closeness between estimated and true elution profiles we use the statistics

$$m_r^\alpha = \frac{1}{p} |\hat{\boldsymbol{\alpha}}_r - \boldsymbol{\alpha}_r|^2 = \frac{1}{p} \sum_{i=1}^p (\hat{\alpha}_{ir} - \alpha_{ir})^2$$

for each of the two elution profiles $\boldsymbol{\alpha}_1$ and $\boldsymbol{\alpha}_2$, where p is the number of chromatographic times. Analogously we use the statistics $m_r^\beta = |\hat{\boldsymbol{\beta}}_r - \boldsymbol{\beta}_r|^2/q = \sum_{j=1}^q (\hat{\beta}_{jr} - \beta_{jr})^2/q$ for each of the two spectra $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$, where q is number of wavelengths, and $m_r^\gamma = (\hat{\gamma}_{rr} - \gamma_{rr})^2/\gamma_{rr}^2$ for each of the two lengths γ_{11} and γ_{22} . Using Result 3.2 we can calculate the expected values of m_r^α , m_r^β and m_r^γ for the SVD estimator. For example the expected value of m_r^α is

$$\begin{aligned} E(m_r^\alpha) &= \frac{1}{p} E\left(\sum_{i=1}^p (\hat{\alpha}_{ir} - \alpha_{ir})^2\right) \approx \frac{1}{p} \sum_{i=1}^p \text{Var}(\hat{\alpha}_{ir}) = \frac{1}{p} \text{Tr}\left(\frac{\sigma^2(\mathbf{D}^{-1})_{rr}}{\gamma_{rr}^2} (\mathbf{I}_{p \times p} - \boldsymbol{\alpha}_r \boldsymbol{\alpha}_r^T)\right) \\ &= \frac{\sigma^2(\mathbf{D}^{-1})_{rr}}{\gamma_{rr}^2} \frac{p - \boldsymbol{\alpha}_r^T \boldsymbol{\alpha}_r}{p} = \frac{p-1}{p} \frac{\sigma^2(\mathbf{D}^{-1})_{rr}}{\gamma_{rr}^2}. \end{aligned} \quad (9)$$

In the same way the other expected values are found as

$$E(m_r^\beta) = \frac{q-1}{q} \frac{\sigma^2(\mathbf{D}^{-1})_{rr}}{\gamma_{rr}^2}, \quad E(m_r^\gamma) = \frac{\sigma^2(\mathbf{D}^{-1})_{rr}}{\gamma_{rr}^2}. \quad (10)$$

For BLS we do not have an explicit asymptotic covariance matrix, but we can compute $\Sigma \text{Var}(\hat{\alpha}_{ir})/p$, $\Sigma \text{Var}(\hat{\beta}_{ir})/q$ and $\text{Var}(\hat{\gamma}_{rr})/\gamma_{rr}^2$ numerically from Result 3.1.

The mean responses in the factorial experiment are listed for constituent No. 1 in Tables 1 and 2, as $12 + \ln(\bar{m}_1^\alpha)$, $12 + \ln(\bar{m}_1^\beta)$ and $12 + \ln(\bar{m}_1^\gamma)$, where \bar{m} is the mean value over the 100 replicates of the experiment. The constant 12 was added for convenience to make all response values positive. The logarithmic transformation changed the previously multiplicative character of the precision indices into an additive one, so that any substantial interactions were neither expected nor found for the logarithmic responses. The logarithms of the approximate expected values (Eqs. (9) and (10)) were used for theoretical comparison with the observed responses in the factorial experiment, by inserting the known values from the design. The effects in the factorial experiment were estimated as signed means and thus correspond to half the effect of changing the current factor from its minus level to its plus level. The theoretical effects are calculated from the asymptotic variances according to Eqs. (9) and (10). Theoretical and observed results are discussed in Section 4.3 below.

Table 1
Simulation results for elution profiles and for spectra

No.	Elution profile						Spectrum					
	BLLS		SVD		ln(ratio)/2		BLLS		SVD		ln(ratio)/2	
	Observed	Theoretical	Observed	Theoretical	Observed	Theoretical	Observed	Theoretical	Observed	Theoretical	Observed	Theoretical
1	0.8	1.1	1.1	1.1	-0.13	0.00	0.6	0.9	1.1	1.1	-0.28	-0.12
2	1.0	1.2	1.1	1.1	-0.04	0.04	0.9	1.1	1.1	1.1	-0.10	-0.01
3	2.7	3.0	2.7	2.7	-0.03	0.12	2.2	2.5	2.8	2.7	-0.27	-0.10
4	2.4	2.6	2.7	2.7	-0.15	-0.08	2.6	2.7	2.8	2.7	-0.09	-0.03
5	2.4	2.7	2.4	2.4	-0.04	0.12	1.9	2.3	2.4	2.4	-0.25	-0.08
6	2.1	2.2	2.4	2.4	-0.17	-0.11	2.3	2.4	2.4	2.4	-0.08	-0.03
7	3.8	4.0	4.1	4.0	-0.17	-0.02	3.6	3.8	4.0	4.0	-0.22	-0.11
8	4.0	4.1	4.1	4.0	-0.04	0.03	3.9	4.1	4.0	4.0	-0.06	0.01
9	2.0	2.3	2.4	2.4	-0.20	-0.06	0.7	1.0	2.4	2.4	-0.87	-0.69
10	1.3	1.4	2.4	2.4	-0.56	-0.51	1.6	1.7	2.4	2.4	-0.43	-0.34
11	2.9	3.2	4.0	4.0	-0.57	-0.42	2.4	2.7	4.1	4.0	-0.84	-0.67
12	3.7	3.7	4.0	4.0	-0.16	-0.14	3.3	3.3	4.1	4.0	-0.39	-0.34
13	2.6	2.9	3.8	3.7	-0.60	-0.40	2.2	2.5	3.7	3.7	-0.76	-0.63
14	3.3	3.4	3.8	3.7	-0.23	-0.16	3.0	3.1	3.7	3.7	-0.37	-0.31
15	5.0	5.2	5.4	5.3	-0.21	-0.06	3.8	4.0	5.3	5.3	-0.73	-0.66
16	4.2	4.2	5.4	5.3	-0.61	-0.55	4.7	4.7	5.3	5.3	-0.32	-0.32
17	3.2	3.5	3.3	3.3	-0.03	0.12	2.7	3.1	3.3	3.3	-0.28	-0.10
18	3.0	3.1	3.3	3.3	-0.14	-0.08	3.1	3.2	3.3	3.3	-0.10	-0.03
19	4.7	4.9	4.9	4.9	-0.13	0.00	4.5	4.7	5.0	4.9	-0.25	-0.12
20	5.0	5.0	4.9	4.9	0.01	0.04	4.9	4.9	5.0	4.9	-0.06	-0.01
21	4.3	4.6	4.6	4.6	-0.16	-0.02	4.2	4.4	4.6	4.6	-0.23	-0.11
22	4.6	4.7	4.6	4.6	-0.02	0.03	4.5	4.6	4.6	4.6	-0.06	0.01
23	6.3	6.5	6.3	6.2	-0.01	0.12	5.9	6.1	6.2	6.2	-0.16	-0.08
24	6.0	6.0	6.3	6.2	-0.13	-0.11	6.2	6.2	6.2	6.2	-0.02	0.03
25	3.6	3.8	4.6	4.6	-0.53	-0.42	3.0	3.3	4.6	4.6	-0.83	-0.67
26	4.3	4.3	4.6	4.6	-0.16	-0.14	3.9	3.9	4.6	4.6	-0.38	-0.34
27	6.1	6.1	6.3	6.2	-0.10	-0.06	5.0	4.8	6.3	6.2	-0.63	-0.69
28	5.4	5.2	6.3	6.2	-0.41	-0.51	5.7	5.5	6.3	6.2	-0.30	-0.34
29	5.6	5.8	6.0	5.9	-0.18	-0.06	4.5	4.6	5.9	5.9	-0.72	-0.66
30	4.9	4.8	6.0	5.9	-0.55	-0.55	5.3	5.3	5.9	5.9	-0.31	-0.32
31	6.8	6.7	7.6	7.5	-0.42	-0.40	6.5	6.3	7.6	7.5	-0.52	-0.63
32	7.2	7.2	7.6	7.5	-0.22	-0.16	7.1	6.9	7.6	7.5	-0.21	-0.31
	0.03		0.03		0.02		0.03		0.03		0.02	
	0.09		0.05		0.05		0.10		0.05		0.05	

The table gives values of $12 + \ln$ of the precision indices m of Section 4.2, and in the last two lines lower and upper bounds for their empirical standard errors. For the factorial design, see Table 2.

Table 2

Simulation results for the scale parameter, γ_{11} , in analogy with Table 1

No.	The factorial design						BLLS		SVD		ln(ratio)/2	
	σ	Design	No. of points	No. of specimen	Form 1	Form 2	Observed	Theoretical	Observed	Theoretical	Observed	Theoretical
1	-	-	-	-	-	-	1.2	1.1	1.3	1.2	-0.05	-0.04
2	-	-	-	-	+	+	1.3	1.1	1.3	1.2	-0.01	0.00
3	-	-	-	+	-	+	2.9	2.7	2.9	2.8	0.00	-0.01
4	-	-	-	+	+	-	2.9	2.7	2.9	2.8	0.00	-0.01
5	-	-	+	-	-	+	2.5	2.5	2.5	2.5	-0.03	-0.01
6	-	-	+	-	+	-	2.5	2.5	2.5	2.5	-0.03	-0.02
7	-	-	+	+	-	-	4.2	4.1	4.3	4.1	-0.06	-0.04
8	-	-	+	+	+	+	4.3	4.1	4.3	4.1	-0.02	0.00
9	-	+	-	-	-	+	2.6	2.4	2.6	2.5	-0.01	-0.02
10	-	+	-	-	+	-	2.6	2.4	2.6	2.5	-0.03	-0.03
11	-	+	-	+	-	-	4.1	3.9	4.2	4.1	-0.05	-0.07
12	-	+	-	+	+	+	4.2	4.0	4.2	4.1	-0.01	-0.01
13	-	+	+	-	-	-	3.6	3.7	3.8	3.8	-0.11	-0.08
14	-	+	+	-	+	+	3.8	3.8	3.8	3.8	-0.03	-0.01
15	-	+	+	+	-	+	5.5	5.4	5.6	5.4	-0.07	-0.02
16	-	+	+	+	+	-	5.5	5.4	5.6	5.4	-0.07	-0.03
17	+	-	-	-	-	+	3.5	3.3	3.5	3.4	-0.02	-0.01
18	+	-	-	-	+	-	3.5	3.3	3.5	3.4	-0.03	-0.01
19	+	-	-	+	-	-	5.1	4.9	5.2	5.0	-0.04	-0.04
20	+	-	-	+	+	+	5.2	5.0	5.1	5.0	0.00	0.00
21	+	-	+	-	-	-	4.6	4.6	4.7	4.7	-0.06	-0.04
22	+	-	+	-	+	+	4.7	4.7	4.7	4.7	-0.02	0.00
23	+	-	+	+	-	+	6.5	6.3	6.6	6.3	-0.06	-0.01
24	+	-	+	+	+	-	6.5	6.3	6.6	6.3	-0.07	-0.02
25	+	+	-	-	-	-	4.7	4.5	4.9	4.7	-0.11	-0.07
26	+	+	-	-	+	+	4.8	4.6	4.9	4.7	-0.04	-0.01
27	+	+	-	+	-	+	6.4	6.2	6.8	6.3	-0.20	-0.02
28	+	+	-	+	+	-	6.3	6.2	6.8	6.3	-0.23	-0.03
29	+	+	+	-	-	+	5.9	6.0	6.1	6.0	-0.08	-0.02
30	+	+	+	-	+	-	6.0	6.0	6.1	6.0	-0.06	-0.03
31	+	+	+	+	-	-	7.6	7.5	8.1	7.6	-0.27	-0.08
32	+	+	+	+	+	+	7.7	7.6	8.1	7.6	-0.16	-0.01
							0.12		0.11		0.08	
							0.18		0.18		0.13	

4.3. Results

For an illustration of the estimates from BLLS and SVD with their errors, see Figs. 5–7. These estimates represent the design point $(- + - - + -)$, that is they were generated with moderately overlapping profiles and spectra, two specimens, 20 chromatographic times and 20 wavelengths, the more informative calibration set and the larger σ , i.e., overall about 30% noise added to the theoretical responses.

Our simulation results for constituent No. 1 are listed in Tables 1 and 2, constituent No. 2 gave similar results. Given in the tables are the mean values of $12 +$ the logarithm of the precision indices m , see Section 4.2, for α_1 , β_1 and γ_{11} , both empirical (as observed in the simulations) and asymptotic (theoretical values). For the empirical values we give the range of the standard error at the bottom of the column. For comparisons between the two methods we also give half the difference between the logarithms of the precision indices m for BLLS and SVD. The main effects of the factorial experiment are given in Table 3. All interactions were negligible. The first row of Tables 1 and 2 is the intuitively most precise point, i.e., the smaller σ , four specimens, 20

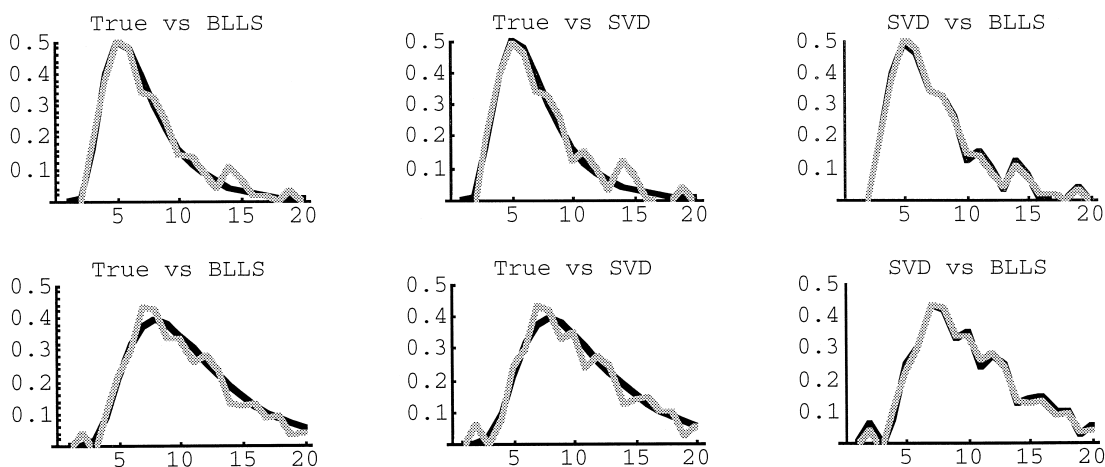


Fig. 5. Estimation of scaled elution profiles; two constituents.

points, more informative design and moderately overlapping profiles and spectra. This point is coded by only minus signs, the remaining points of the 2^{6-1} factorial experiment follow in standard order, see Table 2.

A first, overall impression from the tables is that there is good agreement between the observed simulation results and the asymptotic theory, not least for the SVD estimator. This confirms that the approximate variance formulas well represent the true precision for both BLLS and SVD. Surprisingly neither BLLS nor SVD lose much precision when we let the elution profiles or the spectra be much overlapping.

From the given tables, we see that the SVD estimator is usually not much less efficient than BLLS for uncorrelated homoscedastic noise. SVD works significantly worse than BLLS only under the less informative design, as on rows 9–16 and 25–32 in Table 1, so that in the case of a really bad calibration design the SVD estimator should perhaps not be advocated. The efficiency of SVD compared to BLLS is seen more clearly from Table 3, where we list theoretical and observed main effects from the factorial experiment. All interactions were negligible. The tables also show a general similarity between BLLS and SVD. However, they do not behave precisely the same, and we now focus on some of the differences.

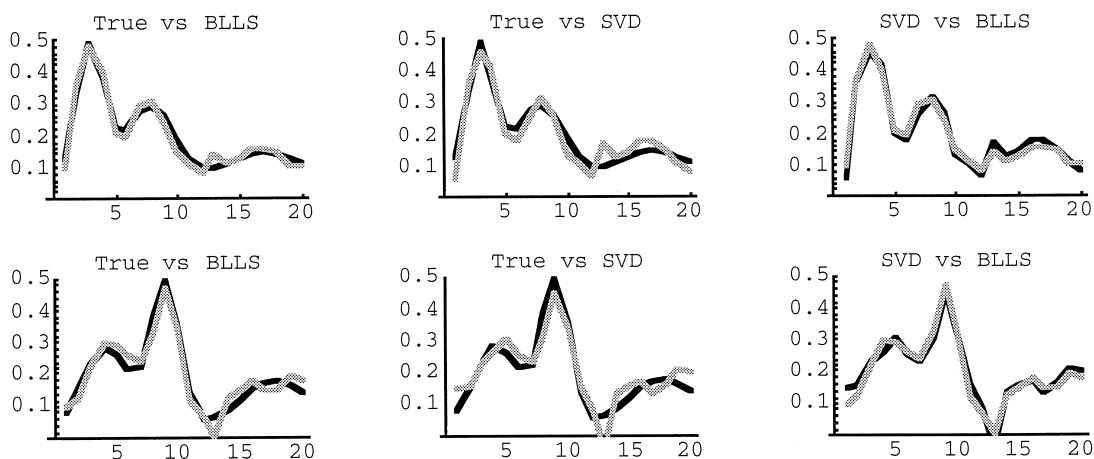


Fig. 6. Estimation of scaled spectra; two constituents.

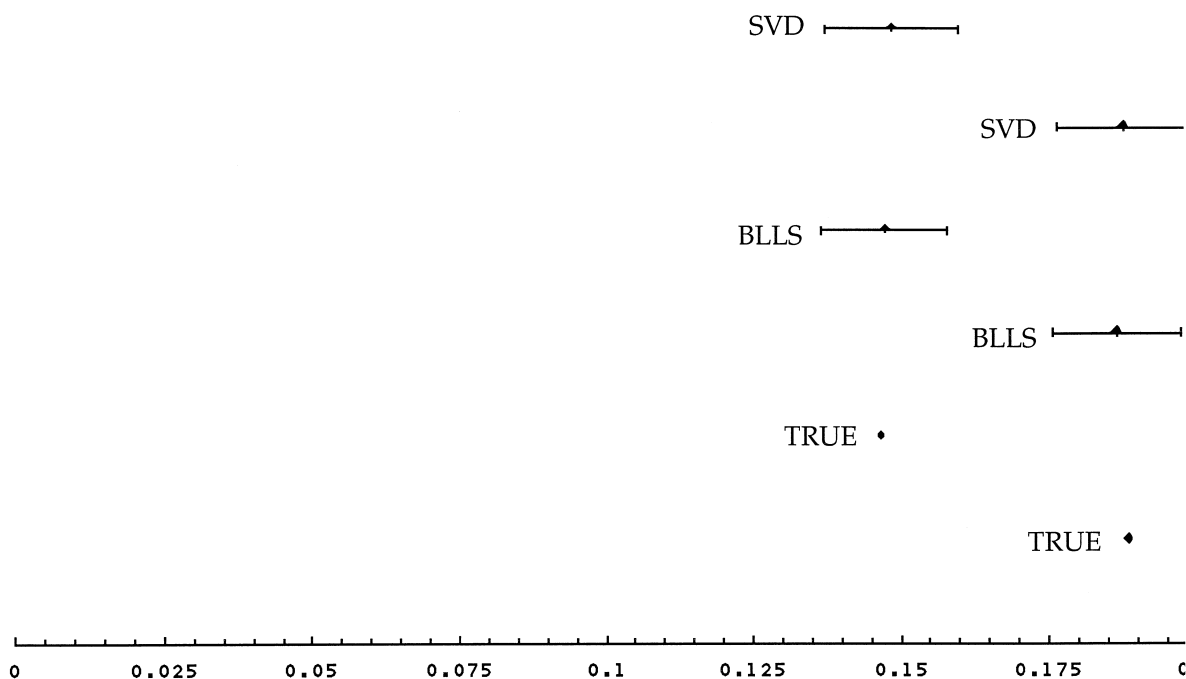


Fig. 7. Estimation of scale factors (γ values); two constituents with ± 2 S.E.M., true values, BLS and SVD, respectively.

4.3.1. Influence from design

BLS is much less sensitive to the choice of calibration design than SVD for the estimation of α and β (but not γ). The fact that SVD is much worse than BLS for the less informative design may be related to the need for inverting the matrix \mathbf{D} , which depends on the calibration set \mathbf{C} . For the less informative calibration set the matrix \mathbf{D} is closer to singular than for the more informative calibration set.

4.3.2. Influence from form

In particular the SVD method is remarkably insensitive to the form factors. The moderately overlapping and the much overlapping levels of these factors are formed from each other by moving only the profile or spectrum for constituent No. 2, so we expect no significant changes in the estimates for constituent No. 1. We use the log-normal distribution for the elution profiles, see Fig. 2. Some small negative main effects appearing for the factor 'form 1' can be explained by a scale change induced by the log-normal distribution.

5. Real data

We have also tried the SVD and BLS methods on real data, which need not satisfy our model assumptions. Data were generated from nine specimens with different mixtures of four polyaromatic hydrocarbons (PAH): anthracene, fluoranthene, chrysene and perylene. The instrumentation was LC–UV, measuring chromatographic times 0–2.10 min with step about 0.36 s and wavelengths 209.5–401.5 nm with step 2 nm. We used the chromatographic time interval 1.30–2.10 and the whole wavelength region in our analyses. With the particular LC-column the PAHs were expected to overlap to some extent, and did so, see Fig. 8.

Table 3

Main effects from the 2^{6-1} experiment; observed and theoretical effects for the logarithmic responses

Response (method)	σ		No. of specimen		No. of points		Design		Form 1		Form 2	
	Observed	Theoretical	Observed	Theoretical	Observed	Theoretical	Observed	Theoretical	Observed	Theoretical	Observed	Theoretical
α_1 (BLLS)	1.2	1.1	0.8	0.8	0.6	0.6	0.4	0.4	0.0	-0.1	0.2	0.2
α_1 (SVD)	1.1	1.1	0.8	0.8	0.7	0.7	0.7	0.7	0.0	0.0	0.0	0.0
α_2 (BLLS)	1.2	1.1	0.8	0.8	0.6	0.6	0.4	0.3	-0.1	-0.1	0.3	0.3
α_2 (SVD)	1.1	1.1	0.8	0.8	0.7	0.7	0.7	0.7	-0.1	-0.1	0.0	0.0
β_1 (BLLS)	1.2	1.1	0.9	0.8	0.7	0.7	0.3	0.2	0.3	0.2	0.0	0.0
β_1 (SVD)	1.1	1.1	0.8	0.8	0.6	0.7	0.7	0.7	0.0	0.0	0.0	0.0
β_2 (BLLS)	1.2	1.1	0.9	0.8	0.7	0.7	0.3	0.2	0.1	0.2	0.0	0.0
β_2 (SVD)	1.1	1.1	0.8	0.8	0.6	0.7	0.7	0.7	-0.1	-0.1	0.0	0.0
γ_{11} (BLLS)	1.1	1.1	0.9	0.8	0.6	0.7	0.6	0.6	0.0	0.0	0.0	0.0
γ_{11} (SVD)	1.2	1.1	0.9	0.8	0.7	0.7	0.7	0.7	0.0	0.0	0.0	0.0
γ_{22} (BLLS)	1.1	1.1	0.9	0.8	0.6	0.7	0.6	0.6	-0.1	-0.1	0.1	0.1
γ_{22} (SVD)	1.2	1.1	0.9	0.8	0.6	0.7	0.7	0.7	-0.1	-0.1	0.0	0.0

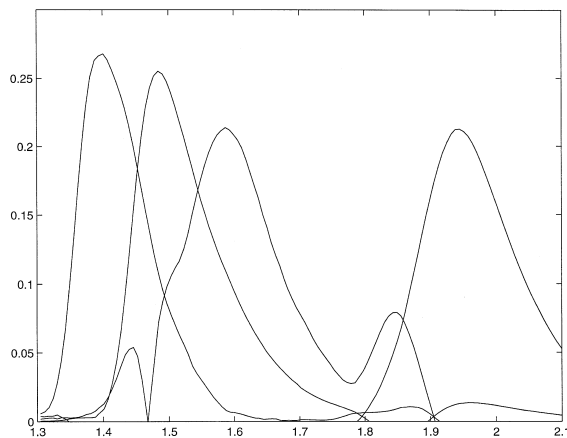


Fig. 8. SVD estimate of the elution profiles.

The purely random additive noise seemed to be very small, but the uncertainty in data was far from negligible. The LC turned out to be affected by drift, so there were problems with centering the peaks on the correct retention times. We tried to set the retention times so that the maximums of the first peak were at the same retention time for all specimens. This ad hoc solution helped to some extent. The pre-processing also included some necessary data cleaning. The time intervals were still not of exactly constant length within and between specimens. Such problems are outside the scope of the present paper, however.

The resulting estimates of the elution profiles from SVD and BLS are mostly quite good, but with some extra 'ghost-peaks'. The result for SVD can be seen in Fig. 8, the results for BLS looks almost the same. We think that the ghost-peaks are due to the drift in the LC. From inspection of data the retention times seem to be about 1.40, 1.48, 1.58 and 1.86 min, respectively for anthracene, fluoranthene, chrysene and perylene. BLS yields the estimates 1.40, 1.48, 1.59 and 1.93 min, and SVD differs only in the estimate of the retention time for perylene, where the estimate is 1.95 min.

The estimates of the spectra also differ very little between the SVD and BLS method. The estimates of the spectra for chrysene can be seen in Fig. 9. The bad resolution (ghost-peaks) in the retention times probably yields some mixing of the spectra. The estimated spectra for chrysene is the one that differs most between the two

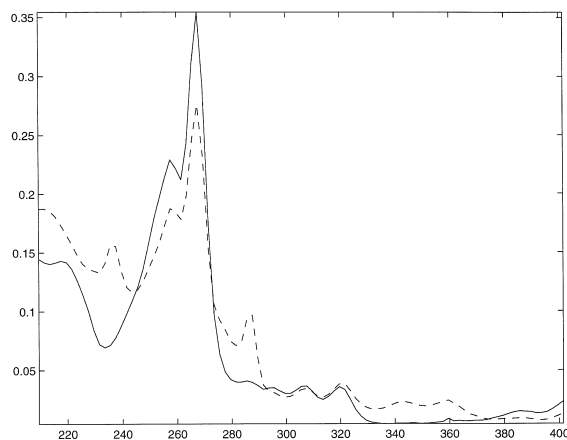


Fig. 9. Spectra for chrysene; BLS (—) and SVD (---).

methods. The differences in the other spectrum cannot be seen in pictures like Fig. 9, because they are too similar.

6. Conclusions and further research

Our tests on real and simulated data confirm that the BLLS and SVD estimators behave quite similarly. The SVD method is typically slightly less efficient than BLLS, or in a badly designed calibration clearly less efficient, under uncorrelated, homoscedastic noise. A compensating advantage of the SVD estimator is that it does not require iterations. Under unfavourable conditions the alternating BLLS algorithm described in Section 2.1 shows a very slow convergence. Furthermore the SVD estimator has explicit asymptotic variance formulas, in contrast to the BLLS estimator, and the approximate standard errors work well even with few samples, bad calibration designs and very large noise-levels. We therefore advocate the use of the SVD estimator for bilinear regression as a reasonable alternative to bilinear least squares.

We are next carrying out more extensive method tests and method comparisons, on real and simulated data, including TLD and multilinear PLS type methods. We will also study how well the calibration model works for predicting the constituents of a new sample. Other important aspects are diagnostics for detecting the presence of unexpected interfering constituents and the treatment of other deviations from the bilinear model which may appear in practice, like in the example of Section 5.

Acknowledgements

The second author is grateful to the Swedish Natural Science Research Council for financial support. We are also grateful to Ludvig Moberg, Department of Analytical Chemistry, Stockholm University, who generated the PAH data and to Göran Robertsson, Scotia LipidTeknik, who found a way to translate those data from an old Hewlett-Packard system into Matlab.

Appendix A. Linearization and variance–covariance matrix for BLLS

In this appendix we linearize the bilinear model with its nonlinear constraints to fit it into a linear framework and use this to find the approximate variance–covariance matrix of Result 3.1. The set of parameters can be expressed as a column vector $\boldsymbol{\theta} = (\text{Vec}(\mathbf{A}); \text{Vec}(\mathbf{B}); \text{Diag}(\boldsymbol{\Gamma}))$. The Vec-operator transforms a matrix to a vector by putting its column vectors on top of each other, i.e., $\text{Vec}(\mathbf{M}) = (m_{11}, \dots, m_{nn})^T$. We also use the notation $\text{Diag}(\mathbf{M})$ to get the vector $(m_{11}, \dots, m_{nn})^T$ of diagonal elements from a $n \times n$ quadratic matrix \mathbf{M} . We now make a first-order Taylor expansion of $\text{Vec}(\mathbf{A}\mathbf{C}^{(k)}\boldsymbol{\Gamma}\mathbf{B}^T)$ with respect to the parameter vector $\boldsymbol{\theta}$. The necessary derivatives can be derived by tensor algebra (see Graham [13]) and are given by

$$\begin{aligned}\Delta_{\mathbf{A}}^{(k)} &= \frac{\partial \text{Vec}(\mathbf{A}\mathbf{C}^{(k)}\boldsymbol{\Gamma}\mathbf{B}^T)}{\partial \text{Vec}(\mathbf{A})} = \mathbf{C}^{(k)}\boldsymbol{\Gamma}\mathbf{B}^T \otimes \mathbf{I}_{p \times p} \\ \Delta_{\mathbf{B}}^{(k)} &= \frac{\partial \text{Vec}(\mathbf{A}\mathbf{C}^{(k)}\boldsymbol{\Gamma}\mathbf{B}^T)}{\partial \text{Vec}(\mathbf{B})} = \mathbf{U}^T (\mathbf{I}_{q \times q} \otimes \boldsymbol{\Gamma}\mathbf{C}^{(k)}\mathbf{A}^T) \\ \Delta^{(k)}\boldsymbol{\Gamma} &= \frac{\partial \text{Vec}(\mathbf{A}\mathbf{C}^{(k)}\boldsymbol{\Gamma}\mathbf{B}^T)}{\partial \text{Diag}(\boldsymbol{\Gamma})} = \mathbf{P}(\mathbf{B}^T \otimes \mathbf{C}^{(k)}\mathbf{A}^T).\end{aligned}$$

Here \mathbf{U} is the permutation matrix associating $\text{Vec}(\mathbf{X})$ with $\text{Vec}(\mathbf{X}^T)$. The matrix \mathbf{P} picks out the rows numbered $r + (r - 1)R$ for $r = 1, \dots, R$, where R equals the number of constituents. Vectorizing all observation matrices $\mathbf{Z}^{(k)}$ jointly, this leads to a linearization of the whole three-dimensional array

$$\begin{pmatrix} \text{Vec}(\mathbf{Z}^{(1)}) \\ \vdots \\ \text{Vec}(\mathbf{Z}^{(K)}) \end{pmatrix} - \zeta = \begin{pmatrix} \Delta_{\mathbf{A}}^{(1)} & \Delta_{\mathbf{B}}^{(1)} & \Delta^{(1)}\Gamma \\ \vdots & \vdots & \vdots \\ \Delta_{\mathbf{A}}^{(K)} & \Delta_{\mathbf{B}}^{(K)} & \Delta^{(K)}\Gamma \end{pmatrix} \begin{pmatrix} \text{Vec}(\mathbf{A}) \\ \text{Vec}(\mathbf{B}) \\ \text{Diag}(\Gamma) \end{pmatrix} + \begin{pmatrix} \text{Vec}(\mathbf{E}^{(1)}) \\ \vdots \\ \text{Vec}(\mathbf{E}^{(K)}) \end{pmatrix},$$

where ζ is a constant, depending only on the central point of the linearization. We also have to linearize the constraints, which for \mathbf{A} can be written as $\text{Diag}(\mathbf{A}^T\mathbf{A}) = \mathbf{1}$ for \mathbf{A} . Here the required derivative is

$$\left(\frac{\partial \text{Diag}(\mathbf{A}^T\mathbf{A})}{\partial \text{Vec}(\mathbf{A})} \right)^T = 2 \begin{pmatrix} \alpha_1^T & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \ddots & & \vdots \\ \vdots & & \ddots & \mathbf{0} \\ \mathbf{0} & \cdots & \mathbf{0} & \alpha_R^T \end{pmatrix} = 2\delta_{\mathbf{A}}.$$

The calculation for \mathbf{B} is analogous and the linearized constraints can be written jointly as

$$\begin{pmatrix} \delta_{\mathbf{A}} & \mathbf{0} \\ \mathbf{0} & \delta_{\mathbf{B}} \end{pmatrix} \begin{pmatrix} \text{Vec}(\mathbf{A}) \\ \text{Vec}(\mathbf{B}) \end{pmatrix} = \begin{pmatrix} \mathbf{1} \\ \vdots \\ \mathbf{1} \end{pmatrix}.$$

We have thus transformed the original bilinear model with its nonlinear constraints into the linear framework $\mathbf{Y} = \text{Vec}(\mathbf{Z}) - \zeta = \mathbf{X}\boldsymbol{\theta} + \boldsymbol{\epsilon}$ with linear constraints $\mathbf{H}^T\boldsymbol{\theta} = \boldsymbol{\eta}$, say.

By the Lagrange multiplier method, with multiplier $\boldsymbol{\lambda}$, the solution to this linear problem is found as root to the equation system

$$\begin{pmatrix} \mathbf{X}^T\mathbf{X} & \mathbf{H} \\ \mathbf{H}^T & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\theta} \\ \boldsymbol{\lambda} \end{pmatrix} = \begin{pmatrix} \mathbf{X}^T\mathbf{Y} \\ \boldsymbol{\eta} \end{pmatrix},$$

that is

$$\begin{pmatrix} \hat{\boldsymbol{\theta}} \\ \hat{\boldsymbol{\lambda}} \end{pmatrix} = \begin{pmatrix} \mathbf{X}^T\mathbf{X} & \mathbf{H} \\ \mathbf{H}^T & \mathbf{0} \end{pmatrix}^{-1} \begin{pmatrix} \mathbf{X}^T\mathbf{Y} \\ \boldsymbol{\eta} \end{pmatrix} = \begin{pmatrix} \mathbf{G}_{11} & \mathbf{G}_{12} \\ \mathbf{G}_{21} & \mathbf{G}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{X}^T\mathbf{Y} \\ \boldsymbol{\eta} \end{pmatrix}.$$

The variance for the estimator $\hat{\boldsymbol{\theta}}$ is finally obtained as

$$\text{Var}(\hat{\boldsymbol{\theta}}) = \text{Var}(\mathbf{G}_{11}\mathbf{X}^T\mathbf{Y} + \mathbf{G}_{12}\boldsymbol{\eta}) = \text{Var}(\mathbf{G}_{11}\mathbf{X}^T\mathbf{Y}) = \sigma^2\mathbf{G}_{11}\mathbf{X}^T\mathbf{X}\mathbf{G}_{11} = \sigma^2\mathbf{G}_{11},$$

where we use the symmetry and the particular form of the matrix inverted. A slightly more general matrix result is found in exercise 15, page 69 of Rao and Mitra [14].

Appendix B. Linearization and variance–covariance matrix for SVD

In this appendix we first (approximately) solve the eigenvector problem induced by the singular value decomposition. We then linearize the expressions for the SVD estimator found in this way and conclude by finding the asymptotic variance–covariance matrix.

We first concentrate on $\hat{\boldsymbol{\alpha}}_r$ from Eq. (7), since an analogous reasoning works for $\hat{\boldsymbol{\beta}}_r$ from Eq. (8). In Eq. (7) the first three terms are rank one. The first two of them have $\boldsymbol{\alpha}_r$ as eigenvector and the third has eigenvector $\boldsymbol{\xi}^{(r)}\boldsymbol{\beta}_r$, with

$$\boldsymbol{\xi}^{(r)} = \sum_{k=1}^K \sum_{r_1=1}^R (\mathbf{D}^{-1})_{rr_1} c_{r_1 r_1}^{(k)} \mathbf{E}^{(k)}. \quad (11)$$

We linearize by neglecting the second-order error term $\boldsymbol{\xi}^{(r)}(\boldsymbol{\xi}^{(r)})^T$ and conclude that the linearized form of Eq. (7) has two non-zero eigenvalues, and that the corresponding eigenvectors are of the form

$$\mathbf{u} = k\boldsymbol{\alpha}_r + m\boldsymbol{\xi}^{(r)}\boldsymbol{\beta}_r, \quad (12)$$

for some coefficients k and m . Inserting Eq. (12) in the linearized Eq. (7), letting $s_r = \boldsymbol{\alpha}_r^T \boldsymbol{\xi}^{(r)} \boldsymbol{\beta}_r$ and $t_r = \boldsymbol{\beta}_r^T (\boldsymbol{\xi}^{(r)})^T \boldsymbol{\beta}_r$ and identifying terms, we get the equation system

$$k\gamma_{rr}^2 + k\gamma_{rr}s_r + m\gamma_{rr}^2s_r + m\gamma_{rr}t_r = k\lambda$$

$$k\gamma_{rr} + m\gamma_{rr}s_r = m\lambda$$

$$k^2 + 2kms_r + m^2t_r = 1.$$

The last equation is a constraint to unit length of \mathbf{v} . Solving this equation system for λ , k and m we find

$$\lambda = \gamma_{rr}^2/2 + \gamma_{rr}s_r \pm \sqrt{\gamma_{rr}^4/4 + s_r\gamma_{rr}^3 + \gamma_{rr}^2t_r} \approx \gamma_{rr}^2/2 + \gamma_{rr}s_r \pm (\gamma_{rr}^2/2 + \gamma_{rr}s_r)$$

$$k = \pm (\lambda/\gamma_{rr} - s_r) / \sqrt{(\lambda/\gamma_{rr})^2 + t_r - s_r^2}$$

$$m = \pm 1 / \sqrt{(\lambda/\gamma_{rr})^2 + t_r - s_r^2},$$

with k and m of the same sign. The root approximation in λ is based on the fact that t_r and s_r^2 should both be small relative to s_r . This yields one solution

$$\lambda_1 \approx \gamma_{rr}^2 + 2\gamma_{rr}s_r$$

$$k_1 \approx \pm (\gamma_{rr} + s_r) / \sqrt{(\gamma_{rr} + 2s_r)^2 + t_r - s_r^2} \approx \pm (1 - s_r/\gamma_{rr})$$

$$m_1 \approx \pm 1 / \sqrt{(\gamma_{rr} + 2s_r)^2 + t_r - s_r^2} \approx \pm 1 / (\gamma_{rr} + 2s_r),$$

and another with eigenvalue $\lambda_2 \approx 0$. Now $\hat{\gamma}_{rr} \approx \sqrt{\lambda_1} \approx \sqrt{\gamma_{rr}^2 + 2\gamma_{rr}s_r} \approx \gamma_{rr} + s_r$. The estimate $\hat{\boldsymbol{\alpha}}_r$ is the eigenvector of Eq. (7) corresponding to the largest eigenvalue, i.e., $\hat{\boldsymbol{\alpha}}_r \approx k_1\boldsymbol{\alpha}_r + m_1\boldsymbol{\xi}^{(r)}\boldsymbol{\beta}_r$. The analogous calculation for $\hat{\boldsymbol{\beta}}_r$ will be left out. We conclude that the SVD estimates can be expressed as

$$\begin{aligned} \hat{\gamma}_{rr} &\approx \gamma_{rr} + \boldsymbol{\alpha}_r^T \boldsymbol{\xi}^{(r)} \boldsymbol{\beta}_r \\ \hat{\boldsymbol{\alpha}}_r &\approx \boldsymbol{\alpha}_r (1 - \boldsymbol{\alpha}_r^T \boldsymbol{\xi}^{(r)} \boldsymbol{\beta}_r / \gamma_{rr}) + \boldsymbol{\xi}^{(r)} \boldsymbol{\beta}_r / \gamma_{rr} \\ \hat{\boldsymbol{\beta}}_r &\approx \boldsymbol{\beta}_r (1 - \boldsymbol{\alpha}_r^T \boldsymbol{\xi}^{(r)} \boldsymbol{\beta}_r / \gamma_{rr}) + (\boldsymbol{\xi}^{(r)})^T \boldsymbol{\alpha}_r / \gamma_{rr}. \end{aligned} \quad (13)$$

We now use the linearizations (Eq. (13)) of the SVD estimates, in combination with the assumption of uncorrelated homoscedastic noise to find approximative covariances. The random parts of these expressions are $\boldsymbol{\alpha}_r^T \boldsymbol{\xi}^{(r)} \boldsymbol{\beta}_r$, $(\mathbf{I} - \boldsymbol{\alpha}_r \boldsymbol{\alpha}_r^T) \boldsymbol{\xi}^{(r)} \boldsymbol{\beta}_r / \gamma_{rr}$ and $(\mathbf{I} - \boldsymbol{\beta}_r \boldsymbol{\beta}_r^T) (\boldsymbol{\xi}^{(r)})^T \boldsymbol{\alpha}_r / \gamma_{rr}$, respectively, so we need only the variances

and covariances of the vectors $\xi^{(r)}\beta_r$ and $(\xi^{(r)})^T\alpha_r$. From the definition of $\xi^{(r)}$ (Eq. (11)), we see that the covariances needed can be traced back to

$$\begin{aligned}\text{Cov}\left[\mathbf{E}^{(k_1)}\beta_{r_1}, \mathbf{E}^{(k_2)}\beta_{r_2}\right] &= \sigma^2\delta_{k_1k_2}\left(\beta_{r_1}^T\beta_{r_2}\right)\mathbf{I} \\ \text{Cov}\left[\left(\mathbf{E}^{(k_1)}\right)^T\alpha_{r_1}, \left(\mathbf{E}^{(k_2)}\right)^T\alpha_{r_2}\right] &= \sigma^2\delta_{k_1k_2}\left(\alpha_{r_1}^T\alpha_{r_2}\right)\mathbf{I} \\ \text{Cov}\left[\mathbf{E}^{(k_1)}\beta_{r_1}, \left(\mathbf{E}^{(k_2)}\right)^T\alpha_{r_2}\right] &= \sigma^2\delta_{k_1k_2}\alpha_{r_2}\beta_{r_1}^T,\end{aligned}$$

where δ_{ij} denotes Kronecker's delta. The covariances for $\xi^{(r)}\beta_r$ and $(\xi^{(r)})^T\alpha_r$ are now found in this manner:

$$\begin{aligned}\text{Cov}\left[\xi^{(r_1)}\beta_{r_1}, \xi^{(r_2)}\beta_{r_2}\right] &= \text{Cov}\left[\sum_{k_1} \sum_{r_3} (\mathbf{D}^{-1})_{r_1r_3} c_{r_3r_3}^{(k_1)} \mathbf{E}^{(k_1)}\beta_{r_1}, \sum_{k_2} \sum_{r_4} (\mathbf{D}^{-1})_{r_2r_4} c_{r_4r_4}^{(k_2)} \mathbf{E}^{(k_2)}\beta_{r_2}\right] \\ &= \sum_{k_1} \sum_{k_2} \sum_{r_3} \sum_{r_4} (\mathbf{D}^{-1})_{r_1r_3} c_{r_3r_3}^{(k_1)} (\mathbf{D}^{-1})_{r_2r_4} c_{r_4r_4}^{(k_2)} \sigma^2 \delta_{k_1k_2} \left(\beta_{r_1}^T\beta_{r_2}\right)\mathbf{I} \\ &= \sum_{r_3} \sum_{r_4} \sum_k \left(c_{r_3r_3}^{(k)} c_{r_4r_4}^{(k)}\right) (\mathbf{D}^{-1})_{r_1r_3} (\mathbf{D}^{-1})_{r_2r_4} \sigma^2 \left(\beta_{r_1}^T\beta_{r_2}\right)\mathbf{I} \\ &= \sum_{r_3} \sum_{r_4} (\mathbf{D}^{-1})_{r_1r_3} d_{r_3r_4} (\mathbf{D}^{-1})_{r_4r_2} \sigma^2 \left(\beta_{r_1}^T\beta_{r_2}\right)\mathbf{I} \\ &= (\mathbf{D}^{-1}\mathbf{D}\mathbf{D}^{-1})_{r_1r_2} \sigma^2 \left(\beta_{r_1}^T\beta_{r_2}\right)\mathbf{I} = \sigma^2 (\mathbf{D}^{-1})_{r_1r_2} \left(\beta_{r_1}^T\beta_{r_2}\right)\mathbf{I}.\end{aligned}$$

Analogously,

$$\text{Cov}\left[\left(\xi^{(r_1)}\right)^T\alpha_{r_1}, \left(\xi^{(r_2)}\right)^T\alpha_{r_2}\right] = \sigma^2 (\mathbf{D}^{-1})_{r_1r_2} \left(\alpha_{r_1}^T\alpha_{r_2}\right)\mathbf{I}$$

and

$$\text{Cov}\left[\xi^{(r_1)}\beta_{r_1}, \left(\xi^{(r_2)}\right)^T\alpha_{r_2}\right] = \sigma^2 (\mathbf{D}^{-1})_{r_1r_2} \alpha_{r_2} \beta_{r_1}^T.$$

Applied on expressions (13) we obtain the covariance formulas of Result 3.2.

References

- [1] Sanchez, Kowalski, Generalized rank annihilation factor analysis, *Anal. Chem.* 58 (1986) 496–499.
- [2] Bro, PARAFAC. Tutorial and applications, *Chemometr. Intelligent Lab. Syst.* 38 (1997) 149–171.
- [3] Sanchez, Kowalski, Tensorial resolution: a direct trilinear decomposition, *J. Chemometr.* 4 (1990) 29–45.
- [4] Burdick, Tu, McGown, Millican, Resolution of multicomponent fluorescent mixtures by analysis of the excitation-emission-frequency array, *J. Chemometr.* 4 (1990) 15–28.
- [5] Wold, Geladi, Esbensen, Öhman, Multi-way principal components-and PLS-analysis, *J. Chemometr.* 1 (1987) 41–56.
- [6] Bro, Multiway calibration, multilinear PLS, *J. Chemometr.* 10 (1996) 47–62.
- [7] Smilde, Comments on multilinear PLS, *J. Chemometr.* 11 (1997) 367–377.
- [8] de Jong, Regression coefficients in multilinear PLS, *J. Chemometr.* 12 (1998) 77–81.
- [9] Geladi, Analysis of multi-way (multi-mode) data, *Chemometr. Intelligent Lab. Syst.* 7 (1989) 11–30.
- [10] Leurgans, Ross, Multilinear models: applications in spectroscopy, *Stat. Sci.* 7 (1992) 289–319.
- [11] Kruskal, Rank, Decomposition and uniqueness for 3-way and N -way arrays, in: R. Coppi, S. Bolasco (Eds.), *Multiway Data Analysis*, Amsterdam, 1989.
- [12] Rao, *Linear Statistical Inference and its Applications*, 2nd edn., Wiley, USA, 1973.
- [13] Graham, *Kronecker Products and Matrix Calculus with Applications*, Ellis Horwood, Chichester, 1981.
- [14] Rao, Mitra, *Generalized Inverse of Matrices and Its Applications*, Wiley, USA, 1971.