NORTH-HOLLAND

A Singular Value Decomposition of a *k*-Way Array for a Principal Component Analysis of Multiway Data, PTA-*k*

Didier Leibovici*
School of Computing and Mathematical Sciences
University of Greenwich, Woolwich Campus
Wellington Street
London SE18 6PF, U.K.

and

Robert Sabatier
Faculté de Pharmacie
Avenue Charles Flahaut
34000 Montpellier, France

Submitted by Richard A. Brualdi

ABSTRACT

Employing a tensorial approach to describe a k-way array, the singular value decomposition of this type of multiarray is established. The algorithm given to attain a singular value, based on a generalization of the transition formulae, has a Gauss-Seidel form. A recursive algorithm leads to the decomposition termed SVD-k. A generalization of the Eckart-Young theorem is introduced by consideration of new rank concepts: the orthogonal rank and the free orthogonal rank. The application of this generalization in data analysis is illustrated by a principal component analysis (PCA) over k modes, termed PTA-k, which conserves most of the properties of a PCA. © 1998 Elsevier Science Inc.

*E-mail: d.Leibovici@greenwich.ac.uk.

LINEAR ALGEBRA AND ITS APPLICATIONS 269:307–329 (1998)

655 Avenue of the Americas, New York, NY 10010 0024-3795/98/$19.00
PII S0024-3795(97)00229-2

1. INTRODUCTION

While in factorial data analysis the duality scheme introduced by Cailliez and Pagès (1976) has permitted adequate comprehension in algebraic terms of such exploratory methods as principal component analysis (PCA) and simple or multiple correspondence analysis (CA or MCA), its capacity to explain a multimode analysis is limited. Thus, it is most appropriate for two-way arrays. For a three-way array, three duality schemes can be drawn, but each entry does not play the same role, as for example in the case of the statis and prestatis methods (i.e. statis on the arrays); see Lavit (1988). These duality schemes are described in Leibovici (1993), along with other models linked to this composite design.

Algebraists and statisticians, attempting to generalize from models such as PCA or singular value decomposition (SVD), have developed models for three-way arrays and k-way arrays to extend existing models. The main problem lies, however, in the manner in which the data may be represented, and thereafter, for what optimization. It is in this context that the Tucker models (Tucker, 1966), developed in PCA over three modes (PCA-3) by Kroonenberg and De Leeuw (1980), have been introduced. The latter authors have systematically used the Kronecker product. Focusing on SVD have been the Parafac model (Harshman, 1970; Kruskal, 1977) and Candecomp by Caroll and Chang (1970), which are actually the same model. Yoshisawa (1987) has described a model combining both the orthogonal Parafac and the Tucker approach. Apart from the Kronecker product, which can be seen as the tensor product operating between matrices for a fixed representation, an extension of the algebraic framework was also required.

Preliminary work was conducted in this area by Kaptein et al. (1986), and Franc (1992) in his thesis included a new algebraic approach, the tensor algebra. A k-way array is seen as a tensor of order k, an element of the tensor product of k vector spaces. This new approach enabled Franc to describe algebraically and analytically Candecomp, Parafac, and PCA-3, and also to extend them to k modes without difficulty.

The purpose of this presentation is to base an extension of PCA to a PCA with k modes, by deriving singular values and the SVD, using the tensorial approach in order to obtain a theorem similar to that of Eckart and Young (1936). In the second section, simple theoretical elements of the tensor product are described. Two further sections are devoted to the explanation of the SVD for a tensor of order 2, 3, or k. An algorithm to obtain the SVD-k will be shown in Section 5, and a generalization of the Eckart-Young theorem for a tensor of order k in Section 6. This last part leads to the elaboration of a method termed principal tensor analysis over k modes (PTA-k), which can be used as a standard method for multiway multidimensional analysis, as PCA is for multidimensional analysis.

2. TENSOR PRODUCT AND MULTIWAY ARRAYS

Firstly, it is essential to recall some definitions for a simple construction of the tensor product and some of the main properties of the calculus from which subsequent methodologies will be derived. These points are given in greater detail in Chambadal and Ovaert (1968), Schwartz (1975), Charles and Allouch (1984), and Lang (1984).

DEFINITION 1.

(i) Let $E_1, \dots, E_k$ be k Euclidean vector spaces of finite dimensions, with metrics $D_1, \dots, D_k$. With a k-tuple $(a_1, \dots, a_k)$ of vectors in these spaces, let the element denoted $a_1 \otimes a_2 \otimes \cdots \otimes a_k$ be a k-linear map on $E_1 \times E_2 \times \cdots \times E_k$ defined by

$$a_1 \otimes a_2 \otimes \cdots \otimes a_k(x_1, \dots, x_k) = \langle a_1, x_1 \rangle_{E_1} \langle a_2, x_2 \rangle_{E_2} \cdots \langle a_k, x_k \rangle_{E_k}, \quad (1)$$

where $\langle\ ,\ \rangle_{E_1}$ indicates the inner product in E_1. This element is termed a decomposed tensor.

(ii) The space generated by all the decomposed tensors is termed the tensor product of the k spaces $E_1, \dots, E_k$, and is denoted $E_1 \otimes E_2 \otimes \cdots \otimes E_k$. Its dimension is the product of the dimensions.

(iii) The inner product in $E_1 \otimes E_2 \otimes \cdots \otimes E_k$ is defined as

$$\begin{aligned} &\langle a_1 \otimes a_2 \otimes \cdots \otimes a_k, x_1 \otimes x_2 \otimes \cdots \otimes x_k \rangle_{E_1 \otimes E_2 \otimes \cdots \otimes E_k} \\ &\quad = \langle a_1, x_1 \rangle_{E_1} \langle a_2, x_2 \rangle_{E_2} \cdots \langle a_k, x_k \rangle_{E_k} \end{aligned}$$

for decomposed tensors. Let $\{ej_1, \dots, ej_{d_j}\}$ be a basis of E_j, and X and A be two tensors:

$$\begin{aligned} &\Big\langle \sum_{i_1 i_2 \dots i_k} A_{i_1 i_2 \dots i_k} e1_{i_1} \otimes e2_{i_2} \otimes \cdots \otimes ek_{i_k}, \\ &\qquad \sum_{i_1 i_2 \dots i_k} X_{i_1 i_2 \dots i_k} e1_{i_1} \otimes e2_{i_2} \otimes \cdots \otimes ek_{i_k} \Big\rangle_{E_1 \otimes E_2 \otimes \cdots \otimes E_k} \end{aligned} \quad (2)$$

$$= \sum_{i_1 i_2 \ldots i_k} X_{i_1 i_2 \ldots i_k} A_{i_1 i_2 \ldots i_k} \langle el_{i_1}, el_{i_1} \rangle_{E_1} \cdots \langle ek_{l_k}, e_{kl_k} \rangle E_k$$

$$= {}^t\overleftrightarrow{X}\left(D_1 \otimes^k D_2 \cdots \otimes^k D_k\right)\overleftrightarrow{A}$$

$$= \langle X, A \rangle_{E_1 \otimes E_2 \otimes \cdots \otimes E_k}$$

where $A_{i_1 i_2 \ldots i_k}, X_{i_1 i_2 \ldots i_k} \in \mathbb{R}$, $\otimes^k$ means the Kronecker product, and $\overleftrightarrow{X}$ is the vectorialization of the tensor X, i.e., its representation as a vector of length $\dim(E_1 \otimes E_2 \otimes \cdots \otimes E_k)$. This definition leads to the expression

$$(E_1 \otimes E_2 \otimes \cdots \otimes E_k)^* = E_1^* \otimes E_2^* \otimes \cdots \otimes E_k^*, \tag{3}$$

where * means the dual space.

(iv) Chambadal and Ovaert (1968) generalize the assertion (3) defining the tensor product of two linear applications: Let $A_1 : E_1 \to F_1$ and $A_2 : E_2 \to F_2$; then let A: $E_1 \otimes E_2 \to F_1 \otimes F_2$ such that $A(x_1 \otimes x_2) = A_1(x_1) \otimes A_2(x_2)$; this unique linear application is expressed as $A = A_1 \otimes A_2$.

(v) A useful operation is proposed by Schwartz (1975), generalizing the image of a vector by a linear application as the contracted product of a vector by a tensor, here denoted .. (no notation having been given by the author). It consists of tensor multiplication of the tensor and the vector followed by contraction on the space to which the vector belongs. For example, let A be a tensor of $E \otimes F \otimes G$, and let $\{e_i\}_{1,n}$, $\{f_j\}_{1,q}$, and $\{g_k\}_{1,p}$ be bases of E, F, and G:

$$A = \sum_{ijk} A_{ijk} e_i \otimes f_j \otimes g_k.$$

Consider a vector $z^* \in G^*$. Then

$$A .. z^* = \sum_{ijk} A_{ijk} e_i \otimes f_j \langle g_k, z^* \rangle$$

$$= \sum_{ijk} A_{ijk} e_i \otimes f_j \left\langle g_k, \sum_m z_m g_m^* \right\rangle = \sum_{ijk} A_{ijk} z_k e_i \otimes f_j. \tag{4}$$

$A .. z^*$ is an element of $E \otimes F$. With z an element of G, $A .. z$ will often be expressed in the same way, explaining a contraction as an inner product. In (4), $\langle g_k, z^* \rangle$ is then changed to $\langle g_k, z \rangle_G$. Thus the inner product of two tensors can be seen as the contracted product between them, and so the metric may be expressed [see (iv)] as

$$\langle A, X \rangle_{E_1 \otimes E_2 \otimes \cdots \otimes E_k} = A \,..\, X = A \,..\, (D_1 \otimes D_2 \otimes \cdots \otimes D_k) \,..\, X. \quad (5)$$

REMARK 1.

(1) Note that (4) can be obtained by transforming A to a matrix with $\dim(E \otimes F)$ rows and $\dim(G)$ columns expressed as

$$\overrightarrow{A}^{G} \quad (6)$$

with qn rows and p columns. Computing the image of z by this matrix leads to

$$\overleftrightarrow{A} \,..\, z^* = \overrightarrow{A}^{G} z^*. \quad (7)$$

If complete vectorialization is put into bijection, for example $E \otimes F \otimes G$ and $\boldsymbol{L}(\mathbb{R}; E \otimes F \otimes G)$, then the indexed vectorialization as in (6) identifies $E \otimes F \otimes G$ as $\boldsymbol{L}(G^*; E \otimes F)$.

(2) The fundamental difference between $\otimes^k$ and $\otimes$ is that the Kronecker product operates with a specific and fixed choice of base (lexicographic order of indices), i.e., $\otimes$ is algebraic, whereas $\otimes^k$ is arithmetic (on coordinates). The advantage of the tensor product is the flexibility of its representations. They depend on the operation applied.

(3) Using the contracted product with the inner product enables one to have an underlying use of metrics.

There are several important properties of the tensor product which may be considered fundamental to factorial data analysis.

PROPERTY 1.

(*a*) *Definition* 1(*iii*) *describes the universal property of the tensor product, which is generally taken for the definition and construction of the tensor* *product. For any bilinear map S the space tensor product implies the commutative diagram*

$$\begin{array}{ccc} E \times F & \xrightarrow{S\ \text{(bilinear)}} & \mathbb{R} \\ \downarrow & \nearrow_{\tilde{S}\ \text{(linear)}} & \\ E \otimes F & & \end{array} \tag{8}$$

(*b*) *The tensor product of two subspaces of E and F is a subspace of* $E \otimes F$.

(*c*) $E \otimes F^*$ *is isomorphic to* $\mathbf{L}(F; E)$, *the space of linear maps from F to E, and* $E^* \otimes F$ *is isomorphic to* $\mathbf{L}(E; F)$. *Even if* $E \otimes F \neq F \otimes E$, *they are isomorphic.*

(*d*) *The operation* $\otimes$ *is associative.*

By Property 1(c) a matrix is identified with a linear map and with a tensor of order two: $E \otimes F \sim \mathbf{L}(F^*; E) \sim \mathbf{M}(n; q; \mathbb{R})$. The factorial analysis methods can thus be described by tensor calculus. This approach can be generalized to an array with k ways, by consideration of the latter as a tensor of order k, i.e., an element of a tensor product of k vector spaces. In practice, and in our presentation, those spaces will be $\mathbb{R}^{m_t}$, where m_t is the number of cells in way t.

3. SINGULAR VALUES FOR TWO MODES

Let $S_X : E^* \times F^* \to \mathbb{R}$ be the bilinear map defined by $S_X(e_i^*, f_j^*) = X_{ij}$ with $\{e_i^*\}_{1,n}\{f_j^*\}_{1,q}$ the canonical bases of the spaces. The universal property of the tensor product implies

$$\begin{array}{ccc} E^* \times F^* & \xrightarrow{S_X} & \mathbb{R} \\ \downarrow & \nearrow_{\tilde{S}_X} & \\ E^* \otimes F^* & & \end{array} \tag{9}$$

Then for all ψ^* and φ^* in E^* and F^*,

$$S_X(\psi^*, \varphi^*) = S_X\Big(\sum_i \psi_i e_i^*, \sum_j \varphi_j f_j^*\Big) = \sum_j \sum_i \psi_i \varphi_j X_{ij} = {}^t\psi X_\varphi$$

$$= {}^t(\psi \otimes^k \varphi)\overleftrightarrow{X} = \tilde{S}_X(\psi^* \otimes \varphi^*) = X..(\psi^* \otimes \varphi^*). \tag{10}$$

PROPERTY 2.

(*i*) *The first singular value can be expressed by different maximizations*:

$$\sigma_1 = \max_{\substack{\|\psi^*\|_{E^*}=1\\ \|\varphi^*\|_{F^*}=1}} \tilde{S}_X(\psi^* \otimes \varphi^*) = \max_{\substack{\|\psi^*\|_{E^*}=1\\ \|\varphi^*\|_{F^*}=1}} \langle \psi^* \otimes \varphi^*, X\rangle$$

$$= \max_{\substack{\|\psi\|_{E}=1\\ \|\varphi\|_{F}=1}} \langle \psi \otimes \varphi, X\rangle_{E\otimes F} = \max_{\substack{\|\psi\|_{E}=1\\ \|\varphi\|_{F}=1}} X ..(\psi \otimes \varphi)$$

$$= X ..(\psi_1 \otimes \varphi_1) \tag{11}$$

(*ii*) *The tensor solution in* (11) *is unique up to an orthogonal transformation leaving X invariant.*

Proof. It is a maximization of a continuous linear map over

$$\{\psi \otimes \varphi | \|\psi\|_E = 1, \|\varphi\|_F = 1\} \subset \{z | \|z\|_{E\otimes F} = 1\}, \tag{12}$$

which is closed in a compact set (the unit sphere); thus it is itself compact. This implies the existence of σ_1. The uniqueness is because of the linear map. ■

In expressing the Lagrange problem associated with this maximization, the classical transition formulae which lead to the eigenequations of the well-known operators are found. In matrix form these are

$$\left.\begin{aligned} X\varphi &= \sigma\psi \\ {}^tX\psi &= \sigma\varphi \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} X\,{}^tX\psi &= \sigma^2\psi \\ {}^tXX\varphi &= \sigma^2\varphi \end{aligned}\right. \tag{13}$$

If there are metrics D and Q on E and F respectively, (13) becomes

$$\left.\begin{aligned} XQ\varphi &= \sigma\psi \\ {}^tXD\psi &= \sigma\varphi \end{aligned}\right\} \Rightarrow \left\{\begin{aligned} XQ\,{}^tXD\psi &= \sigma^2\psi \\ {}^tXDXQ\varphi &= \sigma^2\varphi \end{aligned}\right. \tag{14}$$

In a tensorial form the transition formulae are

$$\begin{aligned} X ..\varphi &= \sigma\psi, \\ X ..\psi &= \sigma\varphi, \end{aligned} \tag{15}$$

where X is the tensor equivalent to the matrix X.

The other singular values can be obtained by consideration of orthogonality constraints in the optimization. Lemma 1 given in Section 4 (for the generalization) enables us to affirm the existence and uniqueness. *A priori* there is a choice with regard to orthogonality in E or F or both. Let E_1 and F_1 be the subspaces generated by the first solution, so that

$$\begin{aligned} E \otimes F &= \left(E_1 \overset{\perp}{\oplus} E_1^{\perp}\right) \otimes \left(F_1 \overset{\perp}{\oplus} F_1^{\perp}\right) \\ &= (E_1 \otimes F_1) \overset{\perp}{\oplus} (E_1^{\perp} \otimes F_1^{\perp}) \overset{\perp}{\oplus} (E_1 \otimes F_1^{\perp}) \overset{\perp}{\oplus} (E_1^{\perp} \otimes F_1) \\ &= (E_1 \otimes F_1) \overset{\perp}{\oplus} (E_1 \otimes F_1)^{\perp} ; \end{aligned} \tag{16}$$

note that $E_1^{\perp} \otimes V_1^{\perp} \subset (E_1 \otimes F_1)^{\perp}$.

Given the duality [(13) or (15)] in the first solution, the maximization solution with constraint in $(E_1 \otimes F_1)^{\perp}$ is in fact in $E_1^{\perp} \otimes F_1^{\perp}$, i.e., the projections of X on the other subspaces of the orthogonal space lead to the null tensor. Thereafter, this space is termed the *orthogonal tensorial* space of the subspaces E_1 and F_1.

REMARK 2. The well-known core matrix in the PCA-3 of Kroonenberg and De Leeuw (1980) derive from this observation and from the lack of duality in these solutions. That is to say, for three modes the solution in the orthogonal space of the first solution is not always in the orthogonal-tensorial space of the preceding solution.

After reiterating the process of solution for singular values or, in this case, after diagonalization (13), an orthogonal decomposition of the tensor, the singular values decomposition SVD-2, may be expressed as

$$X = \sum_{s=1}^{\text{rank } X} \sigma_s \psi_s \otimes \varphi_s , \tag{17}$$

or in matrix form,

$$X = \sum_{s=1}^{\text{rank } X} \sigma_s \psi_s {}^t\varphi_s = \Psi \Lambda^{1/2}\, {}^t\Phi . \tag{18}$$

The well-known matrix approximation theorem may thus be formulated, it permits the performance of a PCA:

THEOREM 1 (Eckart and Young, 1936). *The best rank r ($r < q$)* approximation of a rank q matrix X, according to the norm coming from the inner product in $E \otimes F$, is given by the matrix built with the first r tensors of the *SVD*:

$$X_r = \sum_{s=1}^{r} \sigma_s \psi_s \otimes \varphi_s, \tag{19}$$

the squared distance being

$$\min_{\substack{Z \\ \operatorname{rank} Z = r < q = \operatorname{rank} X}} \|X - Z\|^2 = \|X - X_r\|^2 = \sum_{s=r+1}^{q} \sigma_s^2. \tag{20}$$

Proof. The SVD of X provides (complementing it) an orthogonal basis for the whole space $E \otimes F$, and also here for each of the spaces E and F. Let g_σ be the unitary transformation effecting the change of base. Thus for all matrices Z of rank r,

$$\begin{aligned}\|X - Z\|^2 &= \|g_\sigma(X - Z)\|^2 = \|g_\sigma(X) - g_\sigma(Z)\|^2 \\ &= \sum_{s=1}^{q} (\sigma_s - g_\sigma(Z)_s)^2 + \sum_{s=q+1}^{\dim(E \otimes F)} g_\sigma(Z)_s^2 \\ &\geqslant \sum_{s=r+1}^{q} \sigma_s^2.\end{aligned} \tag{21}$$

The inequality is due to the fact that rank Z = rank $g_\sigma(Z)$; to have rank r, r values must be chosen for $g_\sigma(Z)_s$ and the others set equal to zero. The best values are the first r singular values:

$$X_r = g_\sigma^* g_\sigma(X_r) = \sum_{s=1}^{r} \sigma_s \psi_s \otimes \varphi_s. \tag{22}$$

REMARK 3.

(1) A more appealing proof of (19), available for any unitarily invariant norm, based on the Poincaré separation theorem, can be found in Rao (1979).

(2) It should be noted at this point that saying the rank of a matrix is q corresponds to stating the existence, for the tensor form, of a decomposition into q decomposed tensors. But these tensors of rank one have the intrinsic property of being orthogonal and even tensorially orthogonal (according to the orthogonal-tensorial space). The generalizations proposed in Section 6 will be based on this fact in introducing new concepts of tensor rank.

4. THE DERIVATION OF SINGULAR VALUES FOR THREE OR k MODES

A presentation with three modes is given here; the extension to k modes follows immediately from this. As in Section 3, the universal property of the tensor product is also used to obtain the singular values of multiway data X.

Let $S_X : E^* \times F^* \times G^* \to \mathbb{R}$ be the trilinear map defined by $S_X(e_i^*, f_j^*, g_k^*) = X_{ijk}$ with $\{e_i^*\}_{1\ldots n}(f_j^*)_{1.q}\{g_k^*\}_{1\ldots p}$ being the dual canonical bases of the three spaces:

$$\begin{array}{ccc} E^* \times F^* \times G^* & \xrightarrow{S_X} & \mathbb{R} \\ \downarrow & \nearrow_{\tilde{S}_X} & \\ E^* \otimes F^* \otimes G^* & & \end{array} \tag{23}$$

Then for all $\alpha^*, \beta^*, \gamma^*$ of E^*, F^*, G^*,

$$S_X(\alpha^*, \beta^*, \gamma^*) = S_X\left(\sum_i \alpha_i e_i^*, \sum_j \beta_j f_j^*, \sum_k \gamma_k g_k^*\right) = \sum_{ijk} \alpha_i \beta_j \gamma_k X_{ijk}$$

$$= {}^t(\alpha \otimes^k \beta \otimes^k \gamma)\overleftrightarrow{X}$$

$$= \tilde{S}_X(\alpha^* \otimes \beta^* \otimes \gamma^*) = X ..(\alpha^* \otimes \beta^* \otimes \gamma^*). \tag{24}$$

$\overleftrightarrow{X}$ is the vectorialization of the data X (a three-way array) and is thus seen as a tensor of $E \otimes F \otimes G$.

PROPERTY 3.

(*i*) *The derivation of the first singular value may be expressed as*

$$\sigma_1 = \sigma = \max_{\substack{\|\alpha^*\|_{E^*}=1\\ \|\beta^*\|_{F^*}=1\\ \|\gamma^*\|_{G^*}=1}} \tilde{S}_X(\alpha^* \otimes \beta^* \otimes \gamma^*) = \max_{\substack{\|\alpha^*\|_{E^*}=1\\ \|\beta^*\|_{F^*}=1\\ \|\gamma^*\|_{G^*}=1}} \langle \alpha^* \otimes \beta^* \otimes \gamma^*, X \rangle$$

$$= \max_{\substack{\|\alpha\|_E=1\\ \|\beta\|_F=1\\ \|\gamma\|_G=1}} \langle \alpha \otimes \beta \otimes \gamma, X \rangle_{E\otimes F\otimes G}$$

$$= \max_{\substack{\|\alpha\|_E=1\\ \|\beta\|_F=1\\ \|\gamma\|_G=1}} [X .. (\alpha \otimes \beta \otimes \gamma)] = X .. (\psi \otimes \varphi \otimes \phi). \tag{25}$$

(*ii*) *The tensor solution is unique up to an orthogonal transformation leaving X invariant.*

Proof. The proof is basically the same as for Property 2, as a continuous linear map is maximized on

$$\{\alpha \otimes \beta \otimes \gamma \,|\, \|\alpha\|_E = 1, \|\beta\|_F = 1, \|\gamma\|_G = 1\} \subset \{z \,|\, \|z\|_{E\otimes F\otimes G} = 1\}. \tag{26}$$

which is a closed set within a compact set and therefore itself compact, yielding the existence and uniqueness of the solution with the same restriction as for two modes expressed in Property 2(ii). ∎

For the second and further solutions, it can be written:

PROPERTY 4. *The problem for the second singular value can be written*

$$\sigma_2 = \max_{\substack{\|\alpha\|_E=1\\ \|\beta\|_F=1\\ \|\gamma\|_G=1\\ \alpha\otimes\beta\otimes\gamma \perp \psi\otimes\varphi\otimes\phi}} X .. (\alpha \otimes \beta \otimes \gamma) \tag{27}$$

and has a solution.

Proof. A continuous linear map is maximized over the intersection between a compact set and a closed subspace, which is therefore a compact set. Thus the solution for the second singular value exists.

In order to totally control the orthogonality constraint, the following decomposition of the whole space is used:

$$\begin{aligned} E \otimes F \otimes G &= \left(\psi \overset{\perp}{\oplus} \psi^{\perp}\right) \otimes \left(\varphi \overset{\perp}{\oplus} \varphi^{\perp}\right) \otimes \left(\phi \overset{\perp}{\oplus} \phi^{\perp}\right) \\ &= (\psi \otimes \varphi \otimes \phi) \overset{\perp}{\oplus} (\psi^{\perp} \otimes \varphi^{\perp} \otimes \phi^{\perp}) \\ &\quad \overset{\perp}{\oplus} (\psi \otimes \varphi^{\perp} \otimes \phi^{\perp}) \overset{\perp}{\oplus} (\psi^{\perp} \otimes \varphi \otimes \phi^{\perp}) \overset{\perp}{\oplus} (\psi^{\perp} \otimes \varphi^{\perp} \otimes \phi) \\ &\quad \overset{\perp}{\oplus} (\psi \otimes \varphi \otimes \phi^{\perp}) \overset{\perp}{\oplus} (\psi \otimes \varphi^{\perp} \otimes \phi) \overset{\perp}{\oplus} (\psi^{\perp} \otimes \varphi \otimes \phi), \end{aligned} \tag{28}$$

so the projection of X on $E \otimes F \otimes G$ is

$$\begin{aligned} P_{E \otimes F \otimes G}(X) &= \sigma(\psi \otimes \varphi \otimes \phi) + P_{\psi^{\perp} \otimes \varphi^{\perp} \otimes \phi^{\perp}}(X) \\ &\quad + P_{(\psi \otimes \varphi^{\perp} \otimes \phi^{\perp}) \overset{\perp}{\oplus} (\psi^{\perp} \otimes \varphi \otimes \phi^{\perp}) \overset{\perp}{\oplus} (\psi^{\perp} \otimes \varphi^{\perp} \otimes \phi)}(X) \\ &= X. \end{aligned} \tag{29}$$

The equality (29) is due to the fact that the contracted product of two elements of the first solution gives the third, normalized to σ. The projection onto any of the three subspaces in (28) built with two elements of the first solution and the orthogonal space of the third is then null.

LEMMA 1. *Let T be a subspace of $E \otimes F \otimes G$. Then*

$$\max_{\substack{\|\alpha\|_E = 1 \\ \|\beta\|_F = 1 \\ \|\gamma\|_G = 1 \\ \alpha \otimes \beta \otimes \gamma \in T}} X .. (\alpha \otimes \beta \otimes \gamma) = \max_{\substack{\|\alpha\|_E = 1 \\ \|\beta\|_F = 1 \\ \|\gamma\|_G = 1}} P_T X .. (\alpha \otimes \beta \otimes \gamma). \tag{30}$$

Proof. The proof is immediate, based on the fact that the projection onto T is a self-adjoint operator. ∎

According to the preceding lemma and Property 4, a possible second solution may be obtained by adding the orthogonality constraint of the orthogonal-tensorial space of the first solution, i.e. $T = \psi^{\perp} \otimes \varphi^{\perp} \otimes \phi^{\perp}$. This type of tensor solution will be called a *k-mode solution* (here 3-mode).

Other possible solutions may be obtained by projecting onto one of the three remaining subspaces. These solutions linked to the first (a k-mode solution) are more easily obtained by deriving the singular values over $k-1$ modes after a contracted multiplication by one of the elements of the first triplet solution. This type of solution has been termed the ψ (or φ or ϕ) *associated solution* to the first k-mode solution.

To seek out all the singular values, the two types of possible solutions are thus sought:

(1) the *k-mode solutions*, obtained by projection onto an orthogonal-tensorial space (the first is the whole space) and maximization of $\tilde{S}$ by the algorithm described below;

(2) the *associated solutions of each k-mode solution*, obtained after a contracted product by an element of a k-mode solution and derivation of singular values of an order $k-1$ tensor.

PROPERTY 5. *The second singular value (and the others in decreasing order) will be the scalar product of X with:*

(i) *either a possible k-mode solution or an associated solution, or*

(ii) *a unitary combination of some such possible solutions (this will rarely happen because it must be a rank one tensor combination of orthogonal tensors, which occurs only when a factorization may be carried out—i.e., two elements of the two or more possible solutions are equal).*

Proof. The argument is left to the reader. ■

REMARK 4. The singular values obtained with k-mode solutions are in decreasing order, but, for example, it might happen that a singular value obtained with an associated solution of the first k-mode solution is bigger than the singular value obtained with the second k-mode solution. That is the reason why the previous property employs the word "possible."

5. ALGORITHM FOR THREE-MODE OR k-MODE SINGULAR VALUE DECOMPOSITION

For three modes, the research algorithm RPVSCC for the first solution is as follows:

Initialization: $\alpha_0, \beta_0, \gamma_0,$

Iteration $n+1$ has three steps:

$$\begin{aligned} &1. \quad (X .. \beta_n) .. \gamma_n = {}^{\alpha}\sigma_{n+1}\alpha_{n+1},\\ &2. \quad (X .. \gamma_n) .. \alpha_{n+1} = {}^{\beta}\sigma_{n+1}\beta_{n+1},\\ &3. \quad (X .. \alpha_{n+1}) .. \beta_{n+1} = {}^{\gamma}\sigma_{n+1}\gamma_{n+1}. \end{aligned} \tag{31}$$

For k modes, after initialization the jth step of iteration $n+1$ is

$$j. \quad X .. \left(s_{n+1}^1 \otimes \cdots \otimes s_{n+1}^{j-1} \otimes s_n^{j+1} \cdots \otimes s_n^k\right) = {}^{j}\sigma_{n+1}s_{n+1}^j. \tag{32}$$

The initialization chosen, which leads to quick convergence, is the Tucker1 solution (for definition of the Tucker1 model, see Remark 7). At convergence ${}^{\alpha}\sigma_l = {}^{\beta}\sigma_l = {}^{\gamma}\sigma_l$, and a criterion to stop the algorithm might be based on this fact, or simply on the distances for each component between successive steps. For three modes, the RPVSCC is in fact equivalent to the TUCKALS3 algorithm of Kroonenberg and De Leeuw (1980) with a single component in each mode. Thus, it has the properties of convergence as established by those authors, and the same considerations can be made for k modes.

PROPERTY 6.

(i) *The steps of the* RPVSCC *algorithm come from the Lagrange problem associated to the function* $\tilde{S}_X$:

$$L(\alpha, \beta, \gamma, {}^{\alpha}\sigma, {}^{\beta}\sigma, {}^{\gamma}\sigma) = X .. (\alpha \otimes \beta \otimes \gamma) - \tfrac{1}{2}\,{}^{\alpha}\sigma(\|\alpha\|_E^2 - 1) - \tfrac{1}{2}\,{}^{\beta}\sigma(\|\beta\|_F^2 - 1) - \tfrac{1}{2}\,{}^{\gamma}\sigma(\|\gamma\|_G^2 - 1), \tag{33}$$

where ${}^{\alpha}\sigma$, ${}^{\beta}\sigma$, *and* ${}^{\gamma}\sigma$ *are the Lagrange multipliers, with an analogous expression for* k *modes.*

(ii) *The Hessian of the normal equation for* (*i*) *is, for each parameter, a negative diagonal matrix, proportional to the identity.*

(iii) *Properties* (i) *and* (ii) *are valid for* k *modes.*

Proof. For example, by simple calculus the first derivative with respect to α can be written

$$\frac{\partial}{\partial \alpha} L(\alpha, \beta, \gamma, {}^{\alpha}\sigma, {}^{\beta}\sigma, {}^{\gamma}\sigma) = X .. (\beta \otimes \gamma) - {}^{\alpha}\sigma\alpha;$$

equating to zero gives (i). The proofs of the other assertions, as easy as for (i), are left to the reader. ■

We thus obtain a SVD in two major steps by the recursive SVD-k algorithm below:

A. *Find the k-mode solutions* in two iteration steps:

1. RPVSCC;
2. Projection of the tensor in step 1 (X or a projection of X) on the orthogonal-tensorial space of the solution calculated in step 1.

B. *Find solution associated with the k-mode solutions*:
for each k-mode solution,
for each component of the solution,

1. contracted product with this component;
2. SVD-$(k-1)$.[1]

The SVD of an order k tensor can be written as follows:

$k = 3$: $X = \sum_s \sigma_{sss} x_s \otimes y_s \otimes z_s + \sum_s \sum_{t=s+1} \sigma_{stt} x_s \otimes y_t^{x_s} \otimes z_t^{x_s}$
$+ \sum_s \sum_{t=s+1} \sigma_{tst} x_t^{y_s} \otimes y_s \otimes z_t^{y_s} + \sum_s \sum_{t=s+1} \sigma_{tts} x_t^{z_s} \otimes y_t^{z_s} \otimes z_s$,

$k > 3$: $X = \sum_s \sigma_s \varphi_s^1 \otimes \cdots \otimes \varphi_s^k$, where, given $r < t$, φ_r^i and φ_t^i are such that:

$\varphi_r^i = \varphi_t^i$ for the tth tensor the solution will occur no more than $k-2$ times if the solution r is a k-mode solution (so t is said to be associated to r), no more than $k-1$ times (before factorization) otherwise;

$\varphi_r^i \perp \varphi_t^i$ (occurs at least once): if it occurs k times and r is a k-mode solution, then t is another;

$\varphi_r^i \angle \varphi_t^i$, then either solutions t and r are solutions associated to the same k-mode solution, and this will happen no more than $k-2$ times, or they are associated to different k-mode solutions, no more than $k-1$ times.

[1]All the solutions are not to be retained, because of the duplicates generated by this procedure. Because of processing time problems, their removal may be effected in B.1: form the contracted product with each tensor of order $1, 2, \ldots, k-2$, coming from the k-mode solution, and then project on the orthogonal-tensorial space of the rest of this solution.

REMARK 5.

(1) Yoshisawa (1987) has described research on the k-mode solutions, but combining the elements of these solutions to build something like associated solutions, this author thus develops a model like a Tucker model for k modes.

(2) Denis and Dhorne (1989) have shown, under the name "rocket form," the SVD for three modes, but they try to describe it as a Tucker model, referring to the core matrix, thus forgetting that this decomposition gives an optimal basis for the whole tensor space but not for each of the three spaces. The decomposition is in fact somewhere between the Parafac-Candecomp model and the Tucker model.

6. APPROXIMATION OF A TENSOR: PTA-k

The PCA-3 of Kroonenberg via the TUCKALS3 algorithm does not permit either the definition of the SVD or nested solutions in the models. Here a generalization of the Eckart-Young (1936) theorem is proposed through a new tensor rank conception which is more appropriate for geometric description.

The Definition of the tensor rank adopted by Kruskal (1977) and other authors and described by Franc (1992) is:

DEFINITION 2. The rank of a tensor is the minimum number of decomposed tensors of which the sum is this tensor. The decomposed tensors may be built with unit vectors multiplied by coefficients.

Franc (1992) referred to difficulties in the calculation of the tensor rank and provided some majorizations. The SVD-k introduces new definitions of tensor rank.

DEFINITION 3. The orthogonal rank of a tensor is the minimum number of decomposed tensors which are two by two orthogonal, the sum giving the tensor. The decomposed tensors can be built with unit vectors multiplied by coefficients.

DEFINITION 4. The free orthogonal rank of a tensor is the minimum number of decomposed tensors (the sum giving the tensor) which are two by two orthogonal and where two vectors of the same space from two different decomposed tensors are either orthogonal or equal.

Thus for a given space, all involved vectors (without ties) constitute a free orthogonal system. For example, with $E \otimes E \otimes E$, let a and b be two unitary orthogonal vectors of E; then let

$$\begin{aligned} Y &= \sigma_1 a \otimes a \otimes a + \sigma_2 a \otimes b \otimes b + \sigma_3 b \otimes b \otimes b \\ &= \sigma_1 a \otimes a \otimes a + \sqrt{\sigma_2^2 + \sigma_3^2}\, \frac{\sigma_2 a + \sigma_3 b}{\sqrt{\sigma_2^2 + \sigma_3^2}} \otimes b \otimes b. \end{aligned} \tag{34}$$

The first expression for Y gives its free orthogonal rank 3. The second expression gives its orthogonal rank 2.

By construction the SVD-k will give the minimal decomposition according to the orthogonal rank. If not, some tensorial factorizations will provide it (Property 5). In rare cases, as in the previous example, the SVD-k will give the free orthogonal rank.

PROPERTY 7.

(i) *For a tensor of order two all the ranks are equal, and in general,*

$$\operatorname{rank} X \leqslant \operatorname{rank}_{\perp} X \leqslant \operatorname{rank}_{\perp L} X. \tag{35}$$

(ii) *If in the SVD-k there are only k-mode solutions, then all ranks are equal.*

(iii)(a) *If in a SVD-k of orthogonal rank q, u decomposed tensors are subtracted an orthogonal rank $q - u$ is obtained.*

(iii)(b) *If in a decomposition of free orthogonal rank q, u decomposed tensors are subtracted, a free orthogonal rank $q - u$ is obtained.*

Proof. The proof is left to the reader. ∎

Now it is possible to generalize Theorem 1 to the case of a tensor of order $k > 2$.

THEOREM 2 (Eckart and Young generalized by the orthogonal rank). *The best orthogonal rank r ($r < q$) approximation of an orthogonal rank q tensor X of order k is given by the sum of the r first tensors in SVD-k:*

$$\min_{\substack{Z \\ \operatorname{rank}_{\perp} Z = r}} \|X - Z\|^2 = \sum_{s \in I_{>r}(X)} \sigma_s^2, \tag{36}$$

where $\|\cdot\|$ *is the norm defined by the inner product,* $\operatorname{rank}_\perp$ *is the orthogonal rank, and* $I_{>r}(X)$ *is the subset of multiindices not corresponding to the* r *highest singular values; the minimum is*

$$Z = X_r = \sum_{s \in I_{\leqslant r}(X)} \sigma_s \varphi_s^1 \otimes \varphi_s^2 \otimes \cdots \otimes \varphi_s^k. \tag{37}$$

Proof. The SVD-k decomposition gives a basis (a free system which can be completed) of the whole space $E_1 \otimes E_2 \otimes \cdots \otimes E_k$. Let g_σ be the unitary transformation leading to the change of basis; then for all Z of orthogonal rank r,

$$\begin{aligned}\|X - Z\|^2 &= \|g_\sigma(X - Z)\|^2 = \|g_\sigma(X) - g_\sigma(Z)\|^2 \\ &= \sum_{s \in I_{\neq 0}(X)} [\sigma_s - g_\sigma(Z)_s]^2 + \sum_{s \in I_{=0}(X)} g_\sigma(Z)_s^2 \\ &\geqslant \sum_{s \in I_{>r}(X)} \sigma_s^2.\end{aligned} \tag{38}$$

The inequality is obtained from $\operatorname{rank}_\perp Z = \operatorname{rank}_\perp g_\sigma(Z)$, and (iii) of Property 7 implies the choice of r values of $g_\sigma(X)_s$, the best are the r first singular values of X, and so

$$Z = X_r = g_\sigma^* g_\sigma(X_r) = \sum_{s \in I_{\leqslant r}(X)} \sigma_s \varphi_s^1 \otimes \varphi_s^2 \otimes \cdots \otimes \varphi_s^k. \tag{39}$$

REMARK 6.

(1) An equivalence between this theorem and the SVD-k algorithm, which is a step by step search, can be seen from the lemma given in the appendix.

(2) Note that the higher singular values are not always the first k-mode solutions, because the singular value of an associated solution to the sth k-mode solution can be greater than the singular value of the $(s+1)$th k-mode solution.

(3) It is possible to give a theorem for the free orthogonal rank, but there is no proof of existence of the free orthogonal rank decomposition and no method to find it, should it exist.

DEFINITION 5. The orthogonal rank r approximation supplied by Theorem 2 is called PCA-k of order r by SVD-k, or principal tensor analysis on k modes of order r (PTA-k of order r), or just PTA-k, because of the existence of nested solutions.

This name can be justified by the following properties, which gives some comparisons between certain three-mode methods.

PROPERTY 8. *To compare the first solution given by Tucker* 1, TUCKALS3, *and PTA-*3 *let the three criteria be*

criterion CP (principal component): scalar squared maximum for $X .. \psi$, $X .. \varphi$, *and* $X .. \phi$;
criterion TDP (principal decomposed tensor); scalar squared maximum for $X .. (\psi \otimes \varphi)$, $X .. (\psi \otimes \phi)$, *and* $X .. (\varphi \otimes \phi)$;
criterion VS (singular value): maximum for $X .. (\psi \otimes \varphi \otimes \phi)$.

Then:

(i) *The Tucker*1 *solution only verifies CP.*
(ii) *PTA-*3 *solution verifies TDP, VS, and CP with the supplementary constraint being of orthogonal rank* 1.
(iii) *Tuckals*3 *solution (where* $s, t, u > 1$ *are given) does not verify any criterion and could be a compromise among the three.*

PROPERTY 9.

(i) *The traces of the variance operators (i.e. the sums of the eigenvalues) on each mode are equal.*
(ii) *If there are only k-mode solutions in the SVD-*3, *then we have also equality of the eigenvalues, and Tucker*1, *PCA-*3, *and PTA-*3 *are equivalent.*

Proof. The proofs are straightforward. ■

REMARK 7.

(1) Note that Tucker 1 solutions are built with the diagonalization of the variance operators on each mode (i.e. associated to three PCAs of rearranged arrays); see Kroonenberg and De Leeuw (1980).
(2) If the above methods do not verify the same criteria (CP, TDP, or VS) they are still not very different, as seen in Property 9, where (i) was noticed earlier by Jaffrenou (1978).
(3) Note that the first solutions of the PTA-3 and Candecomp are the same, and this is true for k modes as well.

7. GENERATING MULTIWAY DATA ANALYSIS BY PTA-k

As with PCA, in PTA-k the series of squared singular values gives an additive decomposition of the inertia or total variance (if the tensor is centered):

$$X \,..\, X = {}^t\overleftrightarrow{X}\overleftrightarrow{X} = \sum_s (\sigma_s)^2. \tag{40}$$

This series can be explained as a percentage of inertia. Examining and choosing some of the highest principal tensors by means of these percentages, we have a good summary of the data which enables graphic representations and modelization.

With metrics $D_1, \ldots, D_k$ on the spaces, by the tensorial structure it is easy to see that, to use a program with identity metrics, it can be operated as follows:

(1) transform X by

$$\left(D^{1/2} \otimes \cdots \otimes D_{k-1}^{1/2} \otimes D_k^{1/2}\right) .. X, \tag{41}$$

(2) perform the SVD-k, and perform the inverse transformation on each solution.

Usually, before computing PCA, the triplet to be analyzed should be chosen: the matrix analyzed (which is the data or a transformation of them), the metric on rows, and the metric on columns. With PTA-k a $(k + 1)$-tuple should be chosen, (the tensor analyzed being the data or a transformation of them) and the k metrics on all of the data entries.

Thus, PTA-k is like PCA, a standard method to produce other multiway multidimensional methods. It enables the generalization of CA, MCA, or PCAIV (PCA according to instrumental variables), or other PCAs under linear constraints. Some of these generalizations are explained in Leibovici (1993), and a program in SAS/IML has been written to carry out PTA-k with metrics. Some examples for the analysis of a longitudinal epidemiological study are also given, with a brief comparison of results between PCA-3 and PTA-3.

Graphical representations might be read jointly in view of relations within solutions. These relations are the k steps of the RPVSCC algorithm, and they constitute a generalization of transition formulae. Care should, however, be taken in dealing with representations. For example, in a tensor solution, if two vectors are changed to their opposites, the tensor remains the same. Some supplementary criteria on solutions and data are recommended to fix the solutions.

8. DISCUSSION AND CONCLUSIONS

PTA-k is similar to a generalization of PCA, retaining its best properties, as a result of a generalization of singular value decomposition to a k-way array. Tensorial formalism has been very efficient in obtaining these results. The orthogonal rank, defined here, seems to be well adapted for a descriptive statistical approach, as interest is focused on decomposition. The classical concept of tensor rank could be more suitable for modeling or computing, as for example the calculation of a matrix product with a minimum of multiplication. Unfortunately, evaluation of this rank is a difficult problem still open to debate; Franc (1992) has retraced historical results concerning majorizations of tensor rank, in establishing new ones.

Following each definition of tensor rank, a different generalization of the SVD might be obtained. One which could incorporate all the two-mode properties is the free orthogonal rank, but that decomposition does not seem to exist in general. It may be thought that this rank concept is too focused on decomposition and not enough on modeling. Nonetheless, the PCA-3 of Kroonenberg and De Leeuw (1980) is of this type, as well as the Kaptein model (Kaptein et al., 1986; Wansbeek and Verhees, 1990) as a strict generalization of PCA-3 to k modes. In them there is not necessarily an optimum for the free orthogonal rank, but there is a particular "construction" of it, i.e. in fixing approximations in each space. With the classical rank, the Parafac or Candecomp model is more focused on modeling: the classical rank is smaller than the others. Then, between the two latter ranks, there is the SVD-k developed in this paper, which considers the orthogonal rank. This decomposition leads to the PTA-k as the SVD does to the PCA. Note that if you keep only the k-mode solutions in SVD-k, you have the orthogonal Parafac model, but in its complete form. Here another rank could be defined: the orthogonal tensorial rank (i.e., the decomposed tensors are orthogonal according to the orthogonal tensorial space). It is important to note that many tensor ranks may be elaborated, but they do not always exist; however, approximations according to them can be made. Here it is proved that the orthogonal rank exists, and as consequence the orthogonal tensorial rank does not always exist.

APPENDIX

LEMMA (Additive optimization). *In a Hilbert space, the squared norm approximation of an element by an additive map of parameters is equivalent to optimization step by step, if an orthogonality constraint is imposed on the parameter.*

Proof. Without loss of generality, summation is chosen as the additive map:

$$\begin{aligned}
&\|X - (t_1 + t_2)\|^2 \\
&\quad = \|(X - t_1) - t_2)\|^2 \\
&\quad = \|X - t_1\|^2 - 2\langle (X - t_1), t_2\rangle_H + \|t_2\|^2 \\
&\quad = \|X - t_1\|^2 + \|X\|^2 - 2\langle X, t_2\rangle_H + 2\langle t_1, t_2\rangle_H + \|t_2\|^2 - \|X\|^2 \\
&\quad = \|X - t_1\|^2 + \|X - t_2\|^2 - \|X\|^2 + 2\langle t_1, t_2\rangle_H .
\end{aligned}$$

Therefore, minimizing the left-hand side, under orthogonality of the two parameters t_1 and t_2, is equivalent to minimizing with respect to each parameter with this orthogonality constraint. It is then possible to optimize step by step. ■

The authors thank A. Franc for his constructive criticism of this paper. This research was supported by the ANRS (Agence Nationale de Recherche sur le SIDA).

REFERENCES

Cailliez, F. and Pagès, J.-P. 1976. *Introduction à l'Analyse des Données*, SMASH, Paris.

Caroll, J. D. and Chang, J. J. 1970. Analysis of individual differences in multidimensional scaling via *n*-way generalization of "Eckart and Young" decomposition, *Psychometrika* 35:283–319.

Chambadal, L. and Ovaert, J.-L. 1968. *Algèbre Linéaire et Algèbre Tensorielle*, Dunod, Paris.

Charles, C. and Allouch, D. 1984. *Algèbre Générale*, PUF, Paris.

Denis, J. B. and Dhorne, T. 1989. Orthogonal tensor decomposition of 3-way tables, in *Multiway Data Analysis* (R. Coppi and S. Bolasco, Eds.), North-Holland, Elsevier Science.

Eckart, C. and Young, G. 1936. The approximation of one matrix by another of lower rank. *Psychometrika* 1:211–218.

Franc, A. 1992. Etude Algébrique des Multitableaux: Apports de l'Algèbre Tensorielle, Thèse de Doctorat, Spécialité Statistiques, Univ. de Montpellier II, Montpellier.

Harshman, R. A. 1970. Foundations of the PARAFAC procedure: Models and condition for "an explanatory" multi-mode factor analysis, *UCLA Working Papers Phonetics* 16:1–84.

Jaffrenou, P. A. 1978. Sur l'Analyse des Familles Finies de Variables Vectorielles: Bases Algébriques et applications à la Description Statistique, Thèse 3ème Cycle, Univ. de Lyon I, Lyon.

Kaptein, A., Neudecker, H., and Wansbeek, T. 1986. An approach to n-mode component analysis, *Psychometrika* 51(2):269–275.

Kroonenberg, P. M. and De Leeuw, J. 1980. Principal componant analysis of three-mode data by means of alternating least square algorithms, *Psychometrika* 45:69–97.

Kruskap, J. B. 1977. Three way arrays: Rank and uniqueness of trilinear decomposition with applications to arithmetic complexity and statistics, *Linear Algebra Appl.* 18:95–138.

Lang, S. 1984. *Algebra*, 2nd ed., Addison-Wesley, Reading, Mass.

Lavit, C. 1988. *Analyse Conjointe de Tableaux Quantitatifs*, Masson, Paris.

Leibovici, D. 1993. Facteurs à Mesures Répétées et Analyses Factorielles: Applications à un Suivi Épidémiologique, Thèse de Doctorat en Mathématiques et Applications, Spécialité Biostatistiques, Univ. de Montepplier II, Montpellier.

Rao, C. R. 1979. Separation theorems for singular values of matrices and their applications in multivariate analysis, *J. Multivariate Anal.* 9:362–377.

Schwartz, L. 1975. *Les Tenseurs*, Herman, Paris.

Tucker, L. R. 1966. Some mathematical notes on three mode factor analysis, *Psychometrika* 31:279–311.

Wansbeek, T. and Verhees, J. 1990. The algebra of multimode factor analysis, *Linear Algebra Appl.* 127:631–639.

Yoshisawa, T. 1987. Singular Value Decomposition of Multi-array Data and Its Applications, Tech. Rep. CS-87-038, Dept. of Computer Science, Yamanashi Univ. Takeda, Kofu 400.

Received 24 July 1996; final manuscript accepted 3 March 1997