HENK A. L. KIERS

Comparison of “anglo-saxon” and “french” three-mode methods

Statistique et analyse des données, tome 13, nº 3 (1988), p. 14-32
<http://www.numdam.org/item?id=SAD_1988__13_3_14_0>

Statistique et Analyse des Données
1988 - Vol. 13 n° 3 p. 14-32

COMPARISON OF "ANGLO-SAXON" AND "FRENCH" THREE-MODE METHODS*

Henk A.L. KIERS

Department of Psychology
Grote Markt 31/32, 9712 HV
Groningen, The Netherlands

Résumé: *Sept méthodes pour l'analyse de données à trois indices sont décrites en termes de minimisation de fonction de perte . Basées sur cette description des comparaisons globales et spécifiques sont faites entre plusieurs des méthodes presentées ici. Une attention particulière est portée aux comparaisons entre méthodes françaises et anglo-saxonnes.*

Abstract: *Seven methods for the analysis of three-mode data are described in terms of minimizing loss functions. On the basis of this description global and specific comparisons are made between a number of the methods presented here. Special attention is paid to the comparison of french methods and anglo-saxon methods.*

Mots clés: *Analyse de données à trois indices, Analyse en Composantes Principales*

Indices de classification STMA: 06-010, 06-070

* An earlier version of this paper has been presented at the "evolution data" workshop during the Fifth International Symposium on Data Analysis and Informatics, September, 1987.
The author is obliged to Jos ten Berge and Richard Harshman for useful comments.

Manuscrit reçu le 9 juin 1988
Révisé le 24 novembre 1988

1 – INTRODUCTION

In the present paper a number of methods for the analysis of three-mode data is discussed. Before discussing these methods it seems useful to describe various types of three-mode data. Three-mode data are defined here as observations on elements that are classified according to the "categories" to which they belong of three different "modes". These modes can be a set of observation units (denoted as "objects"), a set of variables, a set of occasions, a set of judges, etc. In the sequel it will be assumed that the first mode refers to objects, the second mode refers to variables and the third mode refers to occasions. Three-mode data can be for instance longitudinal data (repeated observation of the same variables on the same set of objects).

In the present paper only two types of three-mode data will be considered. The first type of three-mode data is called "three-way data". Three-way data are defined here as a set of data consisting of observations of all objects on all variables at all occasions. As a result, the data can be pictured by means of a completely filled three-way array, as in Figure 1. The second type of data to be treated here is called "multiple sets data". Multiple sets data is defined here as observations of different sets of objects on the same set of variables. The different sets of objects are assumed here to be measured at different occasions, but this assumption is merely made for convenience. The methods presented here are in no way limited to this specific kind of multiple sets data. Because multiple sets data consist of observations on different sets of objects, it is not possible to picture multiple sets data by means of a three-way array. A useful way of picturing multiple sets data is to collect the observations on each set of objects in a data matrix of objects by variables, and to collect the resulting data matrices in a supermatrix containing the data matrices for all sets of objects below each other, as in Figure 1.

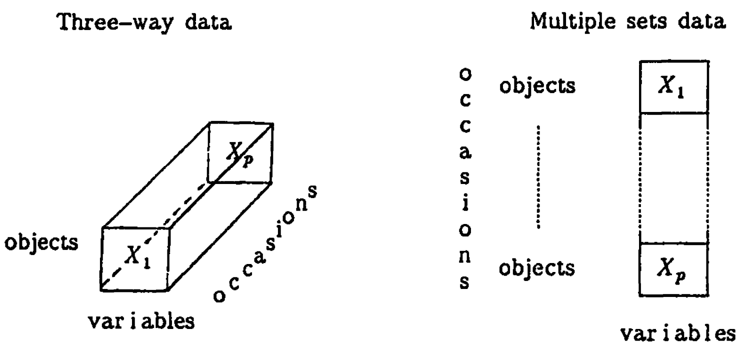

Figure 1. "Three-way data" and "Multiple sets data"

Before discussing methods for analyzing three-mode data, the notation that is to be used in the present paper will be described. For the case of three-way data the following notation is used. Let the elements x_{ijk} denote the elements of the three-way array, where $i = 1,..,n$ is the subscript for the objects, $j = 1,..,m$ is the subscript for the variables, and $k = 1,..,p$ is the subscript for the occasions. Matrix X_k is defined as the n by m matrix containing the elements of the k^{th} frontal slice of the three-way array. That is, X_k contains the observations of the n objects on the m variables at occasion k.

The notation for multiple sets data is chosen such that it optimally corresponds to the notation for three-way data. That is, matrix X_k again denotes a matrix of observations of a set of objects on the variables, at occasion k. However, because at different occasions different sets of objects, with possibly different numbers of objects, are observed, the matrices X_k do not necessarily have the same orders. Let n_k denote the number of objects observed at occasion k, then the order of matrix X_k is n_k by m.

For both three-way data and multiple sets data, the matrix of sums of cross-products among the variables for occasion k is defined as $C_k \equiv X_k{}'X_k$. Obviously, when X_k is centered column-wise, C_k is proportional to a covariance matrix and when the columns of X_k are standardized, C_k is proportional to a correlation matrix.

Above, only two types of three-way data have been described. Obviously, many other types of three-way data are conceivable. An example of such a type of three-mode data is data consisting of observations of one set of objects on different sets of variables. However, this case can be treated in a way equivalent to the case of multiple sets data, when in applying a three-mode method the roles of objects and variables are interchanged.

During the last decades many methods have been developed for the analysis of three-mode data. In order to analyze one's data, a data analyst is facing the problem of choosing from the many different three-mode methods available. Obviously, such a choice can only be made on the basis of comparisons of the different three-mode methods. Various three-mode methods described in the anglo-saxon literature have been compared with each other on large scale (Law, Snyder, Hattie & McDonald, 1984). However, the many three-mode methods developed in France have been mainly overlooked in anglo-saxon literature, and, conversely, these french methods seem to have been developed almost independently of their anglo-saxon counterparts. As a result, when confronted with choosing a three-mode method for analyzing one's data, a data analyst has hardly any basis for choosing between french and anglo-saxon methods.

In the present paper I will briefly describe a number of anglo-saxon three-mode methods, and a number of french three-mode methods. These methods will be described in such a way that a global comparison of all methods is immediately available. That is, I will describe all methods as methods for minimizing certain loss functions, which in itself yields a straightforward basis for comparing the methods. In addition, some more specific relations between certain methods will be treated. These comparisons are in no way exhaustive, but are meant especially for the purpose of comparing the french methods with the anglo-saxon methods.

The discussions of the methods will be primarily technical. That is, the aim of the descriptions is to show briefly on what mathematical criteria the methods are based. For more detail on the methods, their interpretations, and computational aspects, the reader is referred to the sources to be mentioned below. The primary aim of the present paper is to provide the reader with an overview of a number of french and anglo-saxon three-mode methods, and with some insight in the differences between these methods.

2 – DESCRIPTION OF SEVEN THREE-MODE METHODS

In the present section seven methods for the analysis of three-mode data will be described. Some of these are suitable only for analyzing three-way data, while others are especially suitable for analyzing multiple sets data. At this point it should be noted that techniques for the analysis of multiple sets data can always be applied also for analyzing three-way data, because three-way data can be seen as consisting of p different matrices, $X_1,..,X_p$, of objects by variables, where the information that in the case of three-way data the observations are actually made on the *same* objects at all occasions may be ignored. However, it might not be advisable to analyze three-way data as multiple sets data, because this does not use all the information of the three-way data that is available. Conversely, multiple sets of data cannot be analyzed by means of techniques for analyzing three-way data. For every method it will be made clear for which type of data it is suitable.

2.1 – TUCKALS-2

The first method to be described is what is called TUCKALS-2 by Kroonenberg and De Leeuw (1980). This method is used for analyzing three-way data, and hence it is not applicable to multiple sets data. The model underlying the TUCKALS-2 method, the socalled Tucker-2 model, is derived from Tucker's original three-mode factor analysis model (Tucker, 1966). The Tucker-2 model can be described as

$$\hat{X}_k = AH_kB' \qquad (1)$$

for $k = 1,..,p$. In this model matrix $\hat{X}_k$ denotes the model prediction of X_k, the k^{th} frontal slice of the observed three-way data, matrix A is an n by r ($r \leq n$) matrix with component scores of the objects on r object-components ("idealized objects"), matrix B is an m by s ($s \leq m$) matrix of variable loadings on the s variable-components ("idealized variables"), and matrix H_k is an r by s matrix relating the idealized objects to the idealized variables for occasion k.

Equivalently, the model can be described element-wise as

$$\hat{x}_{ijk} = \sum_{l=1}^{r} \sum_{l'=1}^{s} a_{il} b_{jl'} h_{ll'k} , \tag{2}$$

where a_{il} denotes the loading of the i^{th} object on the l^{th} idealized object, $b_{jl'}$ denotes the loading of the j^{th} variable on the l'^{th} idealized variable, and $h_{ll'k}$ denotes the term relating the l^{th} idealized object to the l'^{th} idealized variable, at occasion k.

The Tucker-2 model is fit to the data by minimizing the loss function

$$\sigma_1 (A,B,H_1,..,H_p) = \sum_{k=1}^{p} \| X_k - AH_kB' \|^2, \tag{3}$$

where $\|.\|^2$ denotes the squared Euclidean norm of the matrix concerned.

Kroonenberg and De Leeuw (1980) have also described the more general TUCKALS-3 method, which consists of fitting Tucker's original three-mode factor analysis model in the least squares sense. This method is not described here, because it differs essentially from the other methods considered in the present paper. That is, it does not only reduce the number of objects and the number of variables to a smaller number of idealized objects and idealized variables, but it also reduces the number of occasions to a number of "idealized occasions". In this respect it does not only differ from the TUCKALS-2, but also from the other methods to be described here. For a description of TUCKALS-3 I refer to Kroonenberg and De Leeuw (1980) and for comparisons of this method with other methods the reader is referred to Ten Berge, De Leeuw and Kroonenberg (1987), who compare TUCKALS-3 and PARAFAC, and to Kiers (1988), who compares TUCKALS-3 with a number of french and anglo-saxon three-way methods by considering TUCKALS-3 as part of a hierarchy of three-way methods.

2.2 - CANDECOMP/PARAFAC

Carroll and Chang (1970) and Harshman (1970) independently developed a model which decomposes a three-way array in a very simple way. Harshman called his model PARAFAC (PARAllel FACtor analysis), whereas Carroll and Chang christened their method CANDECOMP (CANonical DECOMPosition). The models are developed for three-way data, and are not suitable for multiple sets data. The CANDECOMP/PARAFAC model can be described as

$$\hat{X}_k = AD_kB', \tag{4}$$

for $k = 1,..,p$, where matrices A and B are matrices of order n by r, and m by r, respectively, and matrix D_k is a diagonal matrix of order r.

Clearly, the CANDECOMP/PARAFAC model is a special case of the TUCKALS-2 model. The important difference of the CANDECOMP/PARAFAC model with the Tucker-2 model is that in the CANDECOMP/PARAFAC model only one set of components is defined (instead of two, as for the TUCKALS-2 model). That is, whereas in the TUCKALS-2 model components are defined for both variables and objects, with H_k containing the relations between these components, in PARAFAC r components are defined simultaneously for variables and objects.

The CANDECOMP/PARAFAC model is based on a very simple rationale. The expression for one entry in the three-way data array is

$$\hat{x}_{ijk} = \sum_{l=1}^{r} a_{il}b_{jl}d_{kl} \ . \tag{5}$$

The elements a_{il}, b_{jl} and d_{kl} are component coordinates of the objects, variables, and occasions, respectively, on the l^{th} CANDECOMP/PARAFAC component. According to the model, there are only proportional differences between objects, variables and occasions with respect to each of the components, and these differences represent multiplicative effects.

The CANDECOMP/PARAFAC model is fit to the data by minimizing the loss function

$$\sigma_2\ (A,B,D_1,..,D_p) = \sum_{k=1}^{p} \| X_k - AD_kB' \|^2, \tag{6}$$

where D_k is a square diagonal matrix of order r, $k = 1,..,p$.

An important feature of the CANDECOMP/PARAFAC method is that it yields a solution with unique axes. That is, whereas, in general, factor analytic solutions are determined only up to a rotation of axes, this model does not allow for such rotation of axes. Differently rotated axes will not, in general, fit the data equally well. Hence, one has an empirical basis for determining the orientation of the axes. The usefulness of this so called "unique axes property" lies in the fact that components can be interpreted in only one undebatable way, namely by interpreting the axes in the orientation found here.

2.3 – Simultaneous Components Analysis

Simultaneous Components Analysis (SCA) is a method for analyzing multiple sets data. It has been proposed by Millsap and Meredith (1988) as "Components Analysis in cross-sectional and longitudinal data", but I will denote it by the name SCA, suggested by Kiers and Ten Berge (1988). The method is a straightforward generalization of Principal Components Analysis (PCA). In order to make this clear, PCA is described as the method that minimizes the loss function

$$\sigma_3(B,P) = \| X - XBP' \|^2 , \tag{7}$$

where the m by r matrix B contains the component weights for constructing the matrix of component scores $F = XB$, and P (m by r) contains the weights for optimally reconstructing the scores in X from F. Hence, XBP', the projection of X on the principal components space, contains the reconstruction of the original data, which is often called "the explained part" of X. Therefore, $(X-XBP')$ contains the unexplained part of the data, and minimizing σ_3 implies minimizing the unexplained inertia or equivalently, maximizing the proportion of explained inertia, which is a well-known interpretation of PCA. It can be verified that a solution for minimizing (7) consists of choosing B and P both equal to the matrix containing the first r eigenvectors of $X'X$.

SCA is a generalization of PCA such that in SCA also the proportion of explained inertia is maximized, while the same component weights are applied to the variables at every occasion. Therefore, there is only one matrix B for every X_k. Hence, the components have the same meaning at every occasion. Matrix P_k contains the pattern scores, or projection coordinates for optimally reconstructing the scores in X_k from $F_k = X_kB$. SCA consists of minimizing the loss function

$$\sigma_4\ (B,P_1,..,P_p) = \sum_{k=1}^{p} \|\ X_k - X_kBP_k'\|^2, \tag{8}$$

which implies minimizing the sum of the amounts of unexplained inertia over p occasions. In this way the method yields one set of component weights that explain most inertia at all occasions simultaneously.

2.4 – LEVIN/TUCKER/JAFFRENNOU

Like SCA, the method for simultaneous factor analysis proposed by Levin (1966) is typically meant for the analysis of multiple sets data. His method can be described as PCA of the super matrix $Y = (X_1'\,\vdots\,..\,\vdots\,X_p')'$. This is equivalent to one of the stages in Tucker's three-mode Principal Components Analysis (Tucker, 1966). As stated by Jaffrennou (1978), the latter in turn is equivalent to one of the stages of Jaffrennou's method for analyzing a three-mode array. Therefore, the method will be denoted as the Levin/Tucker/Jaffrennou method (L/T/J).

Tucker and Jaffrennou both note that PCA of matrix Y does not yield the least squares solution to fitting the Tucker-2 model. Instead, it yields the least squares solution for the problem of fitting $\sum_{k=1}^{p} X_k'X_k$ to the model prediction $\sum_{k=1}^{p} \hat{X}_k'\hat{X}_k$, where $\hat{X}_k = AH_kB'$, according to (1). Therefore, it has been considered to be a reasonable approximation of fitting the Tucker-2 model in the least squares sense.

In order to align the L/T/J method with SCA described above, the L/T/J method is interpreted as PCA on the supermatrix Y, containing $X_1,..,X_p$, below each other. Using the description of PCA as given by (7), the L/T/J method can be described as the method minimizing the loss function

$$\sigma_5(B,P) = \|\ Y - YBP'\|^2 = \sum_{k=1}^{p} \|\ X_k - X_kBP'\|^2, \tag{9}$$

where B denotes a component weights matrix and P a component pattern (or loading) matrix containing weights for reconstructing the variables from the components, both matrices of order m by r. It follows from the description of

PCA as minimizing loss function σ_3, that the minimum of σ_5 is attained when both matrix B and matrix P are chosen to be the principal r eigenvectors of $\sum_{k=1}^{p} X_k'X_k$.

2.5 – STATIS

STATIS, developed by L'Hermier des Plantes (1976), is a method suitable for the analysis of multiple sets data, but can usefully be applied for the analysis of other types of three-mode data as well. In the variant of STATIS that will be described here, data sets are represented by their cross-product matrices, that is, $C_1,..,C_p$. For the purpose of simplification I will assume that the variable and individual metric matrices that may be used in STATIS are all equal to identity matrices. This does not reduce generality, however, because these metric matrices may be assumed to have been built in into the data matrices X_k.

STATIS consists of a three-step procedure. The first step is performing PCA on the set of C_k matrices, considered as variables. This step is essentially based on the correlation measure, the RV–coefficient, proposed by Escoufier (1973) for describing the association between two data sets. STATIS starts by computing the "correlations" between all pairs of cross-product matrices C_k, $k = 1,..,p$. On the resulting correlation matrix a PCA is performed. This PCA yields weights for the p cross-product matrices on the components. The second step is defining the compromise matrix C as the first principal component of the C_k matrices. That is, assuming that α_k gives the first component weight for matrix C_k, $k = 1,..,p$, then the compromise C is given by $C = \sum_{k=1}^{p} \alpha_k C_k = \sum_{k=1}^{p} \alpha_k X_k'X_k$.

The third step, and this is the only step I will consider here, is PCA on the compromise matrix. In order to make this step better comparable to the methods described above, I will describe this method in terms of minimizing a loss function. It is readily verified that PCA on matrix C is equivalent to minimizing

$$\sigma_6(B,P) = \left\| \begin{bmatrix} \sqrt{\alpha_1}X_1 \\ \vdots \\ \sqrt{\alpha_p}X_p \end{bmatrix} - \begin{bmatrix} \sqrt{\alpha_1}X_1 \\ \vdots \\ \sqrt{\alpha_p}X_p \end{bmatrix} BP' \right\|^2 = \sum_{k=1}^{p} \alpha_k \, \| X_k - X_kBP' \|^2, \quad (10)$$

over matrices B and P, both of order m by r, where r is the number of principal components maintained in the solution. Although not used explicitly, the matrices B and P are implicitly used in STATIS as well. In the solution matrix B can again be chosen equal to P, the matrix of component loadings of the compromise variables, and matrix X_kB defines component scores of objects of occasion k. Finally, matrix $X_k'X_kB\Lambda^{-1/2}$, where Λ is the diagonal matrix with the eigenvalues corresponding to the principal components, gives coordinates for the variables for each of the occasions (cf. Lavit, 1985).

2.6 – Analyse Factorielle Multiple

Escofier and Pages (1983,1984) developed Analyse Factorielle Multiple (AFM) for the simultaneous analysis of a number of data sets with the same objects and different variables as an alternative to Generalized Canonical Analysis (GCA; Carroll, 1968), among others. However, AFM can just as well be used for analyzing multiple sets data (Escofier, personal communication). I will treat AFM only for the latter case, even though it does not seem to be the most usual way to describe the method. It should be noted, however, that, as has been mentioned in the introduction, the same technique can be used when the roles of objects and variables are interchanged.

AFM consists of two steps. In the first step the data sets $X_1,..,X_p$ are normalized such that their first principal components all explain the same amount of inertia. This comes down to using $\sqrt{\beta_k}X_k$, instead of X_k, where β_k is the inverse of the largest eigenvalue of C_k.

The second step consists of a (two-way) PCA on the total of all sets of objects, considered as one set of objects with scores on the set of variables. It is readily verified that this PCA, and therefore also AFM, comes down to minimizing the loss function

$$\sigma_7(B,P) = \left\| \begin{bmatrix} \sqrt{\beta_1}X_1 \\ \vdots \\ \sqrt{\beta_p}X_p \end{bmatrix} - \begin{bmatrix} \sqrt{\beta_1}X_1 \\ \vdots \\ \sqrt{\beta_p}X_p \end{bmatrix} BP' \right\|^2 = \sum_{k=1}^{p} \beta_k \, \| X_k - X_kBP' \|^2, \quad (11)$$

over B and P (both of order m by r).

2.7 – LONGI

Recently, Pernin (1987) proposed the method LONGI for the analysis of longitudinal data. Obviously, longitudinal data are a kind of three-way data. It should be noted that the method actually takes into account the "three-way nature" of the data, and that it cannot be applied to multiple sets data.

One of the purposes of LONGI is to find linear combinations of the variables (called "indices de situations") that account maximally for the differences between the objects, while varying minimally within the objects over different occasions. Like STATIS and AFM, LONGI can be applied in many different ways. I will only describe a simplified case, in which there are no missing data. This simplified case can be described as follows.

The matrices X_k are gathered in a supermatrix Y such that each column-supervector corresponds to one variable and contains the scores of all object-occasion combinations for that variable. (These variables are subsequently centered by substracting the means per occasion, but we assume that this transformation has already been carried out). Next, a discriminant analysis is performed such that discriminant functions (linear combinations of the variables) are found that maximally discriminate between the objects. This comes down to performing a canonical correlation analysis on the set of variables in Y and the set of indicator variables indicating the object to which each object-occasion combination refers. When Y contains all matrices X_k below each other, this indicator matrix N contains p times the identity matrix below each other. This canonical analysis can be shown to minimize the loss function

$$\sigma_8(B,C) = \| YB - NC \|^2 = \sum_{k=1}^{p} \|X_kB - C\|^2, \tag{12}$$

subject to $C'C = I$, where B contains weights for the variables for constructing the different discriminant functions and C contains the weights for the objects.

3 – GLOBAL COMPARISON OF THREE-MODE METHODS

Above, seven three-mode methods have been described in terms of fitting least squares loss functions. As has been announced above, this way of presenting the methods facilitates comparison of the methods. A summary of the differences between the methods is given by Table 1, which provides all loss functions, and in addition denotes what type of data the methods can handle.

Table 1. The loss functions for seven three-mode methods

method	loss function	data type
TUCKALS-2	$\sigma_1(A,B,H_1,..,H_p) = \sum_{k=1}^{p} \Vert X_k - AH_kB'\Vert^2$	three-way
CANDECOMP/ PARAFAC	$\sigma_2(A,B,D_1,..,D_p) = \sum_{k=1}^{p} \Vert X_k - AD_kB'\Vert^2$	three-way
SCA	$\sigma_4(B,P_1,..,P_p) = \sum_{k=1}^{p} \Vert X_k - X_kBP_k'\Vert^2$	multiple sets
L/T/J	$\sigma_5(B,P) = \sum_{k=1}^{p} \Vert X_k - X_kBP'\Vert^2$	multiple sets
STATIS	$\sigma_6(B,P) = \sum_{k=1}^{p} \alpha_k \Vert X_k - X_kBP'\Vert^2$	multiple sets
AFM	$\sigma_7(B,P) = \sum_{k=1}^{p} \beta_k \Vert X_k - X_kBP'\Vert^2$	multiple sets
LONGI	$\sigma_8(B,C) = \sum_{k=1}^{p} \Vert X_kB - C\Vert^2$	three-way

In the sequel, some more specific relations will be shown to exist among certain three-mode methods. Before describing these, however, some references will be given of papers that treat the comparison of certain of these methods. Table 2 shows what has been compared where, and which comparisons are to be made in the present paper, in the next section.

Table 2. References on comparisons of three-mode methods

	CANDECOMP PARAFAC	L/T/J	SCA	STATIS/AFM	LONGI
TUCKALS-2	Kroonen-berg 1983 Harshman & Lundy 1984	Kroonen-berg 1983	?	Kroonen-berg 1985	?
CANDECOMP/ PARAFAC		?	?	*this paper*	?
L/T/J			Kiers & Ten Berge 1988	trivial	?
SCA				*this paper*	*this paper*
STATIS/AFM					Pernin 1987

4 - SOME SPECIFIC COMPARISONS BETWEEN THREE-MODE METHODS

4.1 - Comparison of CANDECOMP/PARAFAC and STATIS

The first comparison that will be made is that between CANDECOMP/ PARAFAC and STATIS, when used for analyzing three-way data. This comparison is very similar to the comparison Kroonenberg (1985) made between TUCKALS-2 and STATIS. Suppose the CANDECOMP/ PARAFAC coordinate matrices are constrained to be orthonormal. When the data are perfectly fit by the model we have $X_k = AD_kB'$ for $k = 1,..,p$. Hence, $\sum_{k=1}^{p} C_k = \sum_{k=1}^{p} X_k'X_k = \sum_{k=1}^{p} BD_k^2B'$. In other words, matrix B contains the eigenvectors of matrix $\sum_{k=1}^{p} C_k$. In case the α-weights in STATIS are all taken equal, matrix B also contains the variable loadings from PCA on the compromise matrix in STATIS.

Obviously, this is only a limiting case. Usually the data will not perfectly fit the CANDECOMP/PARAFAC model, especially not when orthonormality constraints are imposed. However, it can be conjectured that the CANDECOMP/PARAFAC and STATIS solutions will not differ very much when the data approximately fit the CANDECOMP/PARAFAC model with orthonormality constraints.

4.2 – Comparison of SCA and STATIS/AFM

The methods SCA, and STATIS/AFM can easily be compared by considering the loss functions that are minimized by the methods:

SCA $$\sigma_4\ (B,P_1,..,P_p) = \sum_{k=1}^{p} \| X_k - X_k B P_k' \|^2, \qquad (8)$$

STATIS $$\sigma_6(B,P) = \sum_{k=1}^{p} \alpha_k \| X_k - X_k B P' \|^2, \qquad (10)$$

AFM $$\sigma_7(B,P) = \sum_{k=1}^{p} \beta_k \| X_k - X_k B P' \|^2. \qquad (11)$$

STATIS and AFM only differ with respect to the weights used in the loss functions. Therefore, comparison of AFM with SCA will yield similar results as comparison of STATIS with SCA. For this reason, only the latter comparison will be treated here.

The loss functions of SCA and STATIS differ in two respects: 1. the STATIS loss function is weighted by the weights α_k; 2. the SCA loss function contains different matrices P_k for each occasion, while in STATIS the same matrix P is used for each occasion (or population). The first difference can in fact easily be undone by assuming that the matrices X_k are the original matrices multiplied by $\sqrt{\alpha_k}$. (Then, in fact, STATIS is equivalent to the L/T/J method). The second difference cannot be undone. In fact, this difference is the same as that between the L/T/J method and SCA, which is discussed by Kiers and Ten Berge (1988). In fact, SCA has been developed precisely in order to find components for the matrices X_k that optimally explain each of the matrices X_k, instead of finding components that optimally explain the variables as if they have been measured in one large population containing all subpopulations. The SCA method considers the sets as consisting of separate observations that are to be explained optimally within each of the separate sets.

4.3 – Comparisons of LONGI with other methods

Finally, a few comparative remarks will be given pertaining to LONGI. These are based on the specific description of LONGI as it has been given here. From the LONGI loss function

$$\sigma_8(B,C) = \sum_{k=1}^{p} \|X_k B - C\|^2, \tag{12}$$

it is clear that, as in SCA, components $F_k = X_k B$ can be computed that are based on the same component weights at every occasion. However, these component weights are found in such a way that the components optimally resemble a postulated overall component matrix C, which is orthonormal. In this way LONGI fits a model completely different from the SCA model.

LONGI is clearly related to generalized canonical analysis (GCA) (Carroll, 1968; cf. Carlier, Lavit, Pagès, Pernin & Turlot, 1988). In GCA the function

$$\sigma_9(B_k,C) = \sum_{k=1}^{p} \|X_k B_k - C\|^2 \tag{13}$$

is minimized subject to $C'C = I$. The only difference between LONGI and GCA is that in LONGI the sets of variables all contain the same variables and the "canonical variates" are computed from these variables by using the same set of weights at each occasion (Carlier et al., 1988). Hence, LONGI can be considered as a fixed weights version of generalized canonical analysis.

Finally, I would like to remark that the solution for matrix C in GCA is given by the first r eigenvectors of $\sum_{k=1}^{p} X_k(X_k'X_k)^{-1}X_k'$, which is equivalent to the solution of STATIS when STATIS is applied to projection operators, as proposed by Glaçon (1981, p.17), and the α_k weights are taken equal. Hence, the relation between LONGI and STATIS applied to projection operators is practically the same as that between LONGI and GCA.

5. – REFERENCES

CARLIER, A., LAVIT, Ch., PAGÈS, M., PERNIN, M.O. & TURLOT, J.C., "Analysis of data tables indexed by time: a comparitive review", Paper presented at the meeting "MULTIWAY '88", March 28–30, 1988, Rome.

CARROLL, J.D., "Generalization of canonical correlation analysis to three or more sets of variables", Proceedings of the 76th Convention of the American Psychological Association, 1968, Vol 3, pp. 227–228.

CARROLL, J.D. & CHANG, J.J., "Analysis of individual differences in multidimensional scaling via an n–way generalization of "Eckart–Young" decomposition", Psychometrika, 1970, Vol. 35, pp. 283–319.

ESCOFIER, B. & PAGES, J., "Méthode pour l'analyse de plusieurs groupes de variables – Application a la caractérisation de vins rouges du Val de Loire", Revue de Statistique Appliquée, 1983, Vol. 31, pp. 43–59.

ESCOFIER, B. & PAGES J., "L'analyse factorielle multiple", Cahiers du bureau universitaire de recherche opérationnelle, Université Pierre et Marie Curie, Paris, 1984, n°42.

ESCOUFIER, Y. (1973) "Le traitement des variables vectorielles", Biometrics, Vol. 29, pp.751–760.

GLAÇON, F., "Analyse conjointe de plusieurs matrices de données", Thèse, Université de Grenoble, 1981.

HARSHMAN, R.A., "Foundations of the PARAFAC procedure: models and conditions for an "explanatory" multi–mode factor analysis", UCLA Working Papers in Phonetics, 1970, Vol. 16, pp. 1–84.

HARSHMAN, R.A. & LUNDY, M.E., "The PARAFAC model for Three–Way Factor Analysis and Multidimensional Scaling", In H.G. Law, C.W. Snyder, J.A. Hattie, and R.P. McDonald (Eds.), Research methods for multimode data analysis (pp. 122–215), New York: Praeger, 1984.

JAFFRENNOU, P.A., "Sur l'analyse des familles finies de variables vectorielles", Thèse, Université Lyon I, 1978.

KIERS, H.A.L. "Hierarchical relations between three-way methods", Paper presented at the 30^{th} meeting of the ASU, Grenoble, May 30 – June 2, 1988.

KIERS, H.A.L. & TEN BERGE, J.M.F., "Simultaneous Components Analysis with equal component weight matrices in two or more populations", Manuscript.

KROONENBERG, P.M., Three mode principal component analysis: Theory and applications, Leiden: DSWO press, 1983.

KROONENBERG, P.M., "Young girls' differential morphological development", Paper at Symposium "The Ins and Outs of Solving Real Problems", June 1985, Brussels.

KROONENBERG P.M. & DE LEEUW, J., "Principal component analysis of three-mode data by means of alternating least squares algorithms", Psychometrika, 1980, Vol. 45, pp. 69–97.

LAVIT, C., "Application de la méthode STATIS", Statistique et Analyse de Données, 1985, Vol. 10, pp. 103–116.

LAW, H.G., SNYDER Jr.,C.W., HATTIE, J.A., and MCDONALD, R.P.(Eds.) Research Methods for Multimode Data Analysis, New York: Praeger, 1984.

LEVIN, J., "Simultaneous factor analysis of several gramian matrices", Psychometrika, 1966, Vol. 31, pp. 413–419.

L'HERMIER DES PLANTES, H., "Structuration des tableaux à trois indices de la statistique", Thèse, Université Montpellier II, 1976.

MILLSAP, R.E. & MEREDITH, W., "Component Analysis in Cross-sectional and

Longitudinal Data", Psychometrika, in press, 1988.

PERNIN, M.O., "Contribution à la méthodologie d'analyse de données longitudinale", Thèse, Université Lyon I, 1987.

TEN BERGE, J.M.F., DE LEEUW, J. & KROONENBERG, P.M. "Some additional results on Principal Components Analysis of three-mode data by means of Alternating Least Squares algorithms", Psychometrika, 1987, Vol. 52, p.183–192.

TUCKER, L.R., "Some mathematical notes on three-mode factor analysis", Psychometrika, 1966, Vol. 31, pp. 279–311.