Component analysis of multisubject multivariate longitudinal data

Timmerman, Marieke

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

Publication date:
2001

Link to publication in University of Groningen/UMCG research database

Citation for published version (APA):

Copyright
Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

Take-down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

Downloaded from the University of Groningen/UMCG research database (Pure): http://www.rug.nl/research/portal. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.

Download date: 30-12-2018
3. Occasion components as evaluations of latent curves: possibilities for constraints to the time mode

In Chapter 2, four component models that are potentially useful for modeling longitudinal three-way data were discussed. Although the models can be used in a completely exploratory way, it is often advantageous to use existing knowledge of the processes generating the data to constrain the model. Potential advantages of model constraints are better estimates of the model parameters, reduction of numerical problems and time required for computation, reduction of transformational freedom in a substantively sensible way, and a reduction of the chance to end in a local minimum, especially in ill-defined models. Possible constraints are unimodality and non-negativity of components (Bro & Sidiropoulos, 1998), equality, symmetry, and orthogonality of components (Bro, 1998), and certain patterns of zero values in one (or more) of the component matrices and the core (Kiers, 1991; Kiers, Ten Berge & Rocci, 1997; Kiers & Smilde, 1998). In this chapter, the rationale for imposing constraints on component models for longitudinal data is introduced.

As discussed in Section 1.2, the expected functional form of the process under study is an important factor in selecting the sampling time points in longitudinal research. The successive scores obtained by repeatedly collecting the score on a variable from the same subject are a kind of ‘functional data’ (Ramsay & Silverman, 1997). Functional data may appear in various ways, but they are all smooth to a larger or smaller extent, in the sense of being repeatedly differentiable (Ramsay & Silverman, 1997). The smoothness property of functional data does not imply that the observed data are evaluations of a smooth function, since observational error or noise may disturb smoothness. An approach to modeling univariate functional data is trying to filter out the error term by imposing smoothness restrictions. In this way, no rigid parametric assumptions about the dependence of scores and time are imposed. Several smoothing techniques are available (Wahba, 1990; Hastie & Tibshirani, 1990; Ramsay & Silverman, 1997), like kernel smoothing and polynomial smoothing. Another approach is to impose a certain functional relationship between scores and time, and to estimate parameters of the function from the observed data. The form of the functional relationship is ideally determined on the basis of the mechanism producing the data. If the latter is not possible, one could base the decision on the form of the observed data. Interpretability of the parameters of the functions should play a role. Smoothing techniques and imposing functional relationships are widely used in the univariate case.
The earliest approach for identifying the structure in two-way functional data was principal component analysis (Tucker, 1958, 1966b; Rao, 1958). Functional principal component analysis (PCA), as it is called by Ramsay and Silverman (1997), aims at describing the dominant modes of variation of a functional data set. This exploratory method is similar to PCA applied to a multisubject or multivariate longitudinal data matrix. A functional PCA reveals a number of component scores at the different time points, and subject or variable weights for these components. The component scores at the time points can be viewed as estimated evaluations of so-called ‘basis functions’. Hence, an observed score at a certain time point is a weighted sum of the basis functions that are evaluated in that particular time point.

Meredith and Tisak (1990) proposed a factor analysis counterpart of the functional PCA model, the latent curve model. They treated the individual weights for the latent curves as unobserved latent variables, and made distributional assumptions about the error term. The latent curve model differs from a standard factor analysis model in that the expected value of the observed variable is the average growth curve, and hence is (usually) non-zero. Also, the expected values of the weights for the latent curves are non-zero. The latent curves can be left unconstrained, but they can also be partly or fully prespecified.

Unconstrained latent curves, as well as components representing basis functions resulting from a functional PCA, have transformational indeterminacy. Common transformational procedures aiming at simple structure (e.g., Varimax; Kaiser, 1958) are generally not useful in transforming latent curves or basis functions. Tucker (1966b) defined some criteria for transforming the solution of his basis functions, like non-negativity of the component scores, and non-negativity of the slopes. However, in practice, empirical latent curve analyses with unconstrained latent curves seem to use only one latent curve (Meredith & Tisak, 1990; Jones & Meredith, 1996), possibly to avoid transformational indeterminacy, as Browne (1993) suggested. As in practice more than one latent curve could be needed to model empirical data satisfactorily, other approaches are needed. For example, transformational indeterminacy can be removed by fully fixing the basis functions or latent curves. Meredith and Tisak (1990) mentioned the use of orthogonal polynomials as a possibility. An alternative is to specify the basis functions or latent curves only partly, which reduces or even removes the transformational indeterminacy. In analyzing ‘growth data’, Browne and Du Toit (1991) and Browne (1993) used latent curves that summarize basis features of a certain (nonlinear) function that represents growth. The component models for longitudinal three-way data discussed in Chapter 2 can be viewed as three-way extensions of functional PCA. It can be useful to constrain the basis functions in the component models for longitudinal three-way data in similar ways.

The specification of the latent curves offers the opportunity to estimate evaluations of basis functions at unobserved time points. This is useful, for example, if the scores on variables are collected at different sets of time points. Standard procedures to estimate the component models for longitudinal three-way data require
all elements of the data array to be observed. The ‘missing elements’, which result from the unequal measurement occasions for the variables and/or subjects, could be estimated, and imputed in the three-way array. The thus obtained full three-way array can be analyzed using standard estimation procedures. To use this approach, it is necessary that the missing data can be considered to be missing completely at random (Little & Rubin, 1987). This implies that missingness is unrelated to the (observed or unobserved) scores themselves.

The possibilities for model constraints as discussed for two-way functional data can be used in the component models for longitudinal three-way data as well. The first one to explore further is the use of smoothness constraints on the occasion components. In Chapter 4, a method for imposing smoothness constraints in the Tucker3 and CP models is proposed. The smoothness constraints can be combined with monotonocity constraints. The method can also easily be applied to the Tucker2 and Tucker1 models, but this option is not treated explicitly.

A second approach is to impose a certain functional form on the occasion components. This is particularly attractive if substantive considerations point to a specific functional relationship. Browne and Du Toit (1991) and Browne (1993) elaborated structured latent curve models for learning data, in which the columns of the occasion component scores matrix are parameterized parsimoniously. This approach is extended to modeling longitudinal three-way data, and learning data in particular, in Chapter 5.