DOCUMENT RESUME

ED 320 949 TM 015 220

AUTHOR Millsap, Roger E.

TITLE Component vs Factor Analytic Approaches to Longitudinal Data.

PUB DATE Jun 86

NOTE 13p.; Paper presented at the Annual Meeting of the Psychometric Society (Toronto, Ontario, Canada, June 21-23, 1986).

PUB TYPE Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.

DESCRIPTORS Comparative Analysis; Computer Software; Cross Sectional Studies; Error of Measurement; *Factor Analysis; *Individual Differences; *Longitudinal Studies; *Mathematical Models; Multivariate Analysis; *Regression (Statistics)

ABSTRACT

A component analytic method for analyzing multivariate longitudinal data is presented that does not make strong assumptions about the structure of the data. Central to the method are the facts that components are derived as linear composites of the observed or manifest variables and that the components must provide an adequate representation of the observed variables. Specifically, the components are derived so as to minimize the sum of squared errors in the linear regression of observed component variables. Although no structural assumptions are required, the method derives components under a variety of stationary constraints. An advantage of the method for researchers studying change at the individual level is that scores on the derived components are uniquely calculable and have clear properties. The method might, therefore, be usefully applied as a precursor to the application of strong models for change in the component scores at the individual level. The method can easily be generalized to encompass data measured longitudinally in multiple groups. A program has been written in FORTRAN to perform the method's analysis, although the program is not "exportable" at this time. An example of the application of the component analysis method is presented. (TJH)

Component vs Factor Analytic Approaches to Longitudinal Data

Roger E. Millsap

Baruch College
The City University of New York

Paper presented at the annual meeting of the Psychometric Society, June 21-24, 1986 at the Ontario Institute for Studies in Education, Toronto, Ontario, Canada.

A number of researchers have proposed extensions of the common factor model to treat longitudinal multivariate data, Cattell (1963), Corballis and Traub (1970), Hakstian (1973), Joreskog and Sorbom (1977), McDonald (1984), Swaninathan (1984), and Tisak (1984) being a few examples. A different factor analytic tradition that is also applicable to such data is represented by the work in three-mode factor analysis begun by Tucker (1963, 1964, 1966) and continued by Kroonenburg and De Leeuw (1980), Sands and Young (1980), and Harshman and Berenbaum (1981), among others. Finally, we have more recent developments, whose roots lie in earlier work by Tucker (1958) and Rao (1958), that combine factor analytic methods with models of individual growth curves. Meredith and Tisak (1984) and McArdle (1986) are actively developing these methods, and Molenaar's (1985) work in dynamic factor analysis is similar in spirit.

If these diverse methods have anything in common, it is that they begin by assuming rather strong models, at least with respect to the covariance structure in the data. For example, the common factor approach assumes the common factor model at each occasion, typically with an identical number of factors at each occasion. Further assumptions are made concerning the covariance structure of the unique factors over occasions. With additional distributional assumptions, large sample tests of fit are available. If the basic model provides an adequate fit, we can proceed to test more restrictive models in a nested fashion. These further restrictions usually involve stationarity constraints on portions of the model, such as the factor pattern matrices. The ultimate goal is usually to discover which aspects of the model can be taken as stationary over occasions, leading to conclusions about stability in factor structure.

In what follows, I will present a component analytic method for analysing multivariate longitudinal data that does not make strong assumptions about the structure of the data. Although no structural assumptions are required, the method derives components under a variety of stationarity constraints as explained below. An advantage of the method for researchers studying change at the individual level is that scores on the derived components are uniquely calculable and have clear properties. The method might therefore be usefully applied as a precursor to the application of strong models for change in the component scores at the individual level. The method can easily be generalized to encompass data measured longitudinally in multiple groups, although I will not explore this extension in this paper.

The "assumption free" nature of the component method to be presented can be a virtue or a vice, depending upon our timidity in making assumptions about the data at hand. I share the sentiments expressed by Rogosa and others (Rogosa, Brandt, & Zimowski, 1982; Rogosa & Willett, 1985) concerning the need for strong models of change at the individual level. When we are unable or unwilling to formulate such models however, the method to be presented is a useful data-analytic strategy.

The Component Method

The view of component analysis which underlies the method proposed here was given in a recent paper by Meredith & Millsap (1985). A key idea is that components are derived as linear composites of the observed or manifest variables, and that the components must provide an adequate representation of the observed variables. Specifically, the components are derived so as to minimize the sum of squared errors in the linear regression of observed on component variables. Let X be an nx1 vector of observable random variables and define a mx1 vector of component score variables $Z = W'X$, where W is an nxm matrix of compositing weights. Assume $\epsilon(X) = 0$ and $\epsilon(XX') = \Sigma$. The paper by Meredith and Millsap demonstrated that all forms of component analysis, weighted or unweighted, can be formulated in terms of finding the compositing matrix W which maximizes the function

$$F(W;\Sigma,G) = \mathrm{tr}\{W'\Sigma G\Sigma W(W'\Sigma W)^{-1}\}. \qquad (1)$$

The matrix G is an nxn nonsingular weighting matrix that can be used to differentially weight the elements of X in deriving the components, as discussed for example in Mulaik (1972). The matrix G might be used to adjust the metric of X, or to weight the elements of X in proportion to their reliability. There are many other possibilities.

<u>The multiple-occasion case</u>. Consider a more general function similar to F in (1)

$$H(U;A,B) = \mathrm{tr}\{U'AU(U'BU)^{-1}\} \qquad (2)$$

with A and B both qxq matrices, and U a qxs matrix. Let V be a qsx1 supervector containing the s columns of U in concatenated form, beginning with the first column. Thus if $U = [u_1, u_2, \ldots, u_s]$, then $V' = [u_1', u_2', \ldots, u_s']$. Now let V be subject to r linear constraints of the form

$$C'V \geqslant M \qquad (3)$$

with C an qsxr matrix and M an rx1 vector, both chosen a priori. Some of the r equations in (3) may be equalities. In what follows, we will assume that we have an algorithm for finding a matrix U that maximizes H in (2), subject to the constraints in (3), for given matrices A and B. Clearly, the function F in (1) is a special case of H in (2) in which U is unconstrained, $A = \Sigma G \Sigma$, and $B = \Sigma$. The constraints in (3) could be used to set upper and/or lower bounds on the elements of U, to fix various elements of U to selected values, or to equate elements of U. More complex linear constraints are possible as well.

To return to the multiple-occasion case now let X be a supervector composed of n random variables observable on each of p occasions: $X' = [X_1', X_2', \ldots, X_p']$. Let $\epsilon(X) = 0$ and $\epsilon(XX') = \Sigma$, a supermatrix of the form

$$\Sigma = \begin{vmatrix} \Sigma_{11} & \Sigma_{12} & \cdots & \Sigma_{1p} \\ \Sigma_{21} & \Sigma_{22} & \cdots & \Sigma_{2p} \\ & & & \\ \Sigma_{p1} & \Sigma_{p2} & \cdots & \Sigma_{pp} \end{vmatrix} \tag{4}$$

We can define an mx1 vector of component scores at the ith occasion as $Z_i = W_i'X_i$, with W_i an nxm compositing weight matrix to be applied to the observable variables X_i. Let D_W be an npxmp block diagonal matrix whose diagonal submatrices are the matrices W_i

$$D_W = \begin{vmatrix} W_1 & 0 & 0 & \ldots 0 \\ 0 & W_2 & 0 & 0 \ldots 0 \\ & & & \\ 0 & 0 & 0 & W_p \end{vmatrix} \tag{5}$$

We can now write the supervector of component scores Z as $Z = [Z_1', Z_2', \ldots, Z_p']$, with $Z = D_W'X$. It also follows that $\epsilon(ZZ') = D_W'\Sigma D_W$ and $\epsilon(XZ') = \Sigma D_W$.

As in the single occasion case, we want to choose the compositing weights D_W so that the resulting components provide an adequate representation of the observed variables. Consider the linear regression of observed on component variables

$$\epsilon(X|Z) = PZ = \hat{X} \tag{6}$$

with P an npxmp component pattern matrix. At this point we must make a choice. We can regress each X_i on all elements of Z, or we can restrict the regression to be "within occasion", regressing X_i on Z_i. I will term the first alternative the "full information" case, since we use all component information in the regression. I will term the second alternative the "limited information" case, because only "within-occasion" information is utilized. This choice is significant because the solutions for D_W in the two cases need not be identical.

__The full information case__. In this case we take P to be a full matrix, and we can consider the matrix P which will minimize the weighted risk function

$$R_F = \epsilon\{(X - \hat{X})' G(X - \hat{X})\} = \epsilon[tr\{G(X - \hat{X})(X - \hat{X})'\}] \tag{7}$$

In (7), G is a matrix of weights to be applied in the regression of X on Z, and has a role analogous to G in the single occasion case in (1). The matrix P which minimizes R_F is

$$P = \Sigma D_W(D_W'\Sigma D_W)^{-1} \tag{8}$$

Clearly P is a function of D_W, and we can choose D_W to minimize R_F or to maximize

$$H(D_W; \Sigma G\Sigma, \Sigma) = tr\{D_W'\Sigma G\Sigma D_W(D_W'\Sigma D_W)^{-1}\} \tag{9}$$

The algorithm for solving H in (2) can utilize the linear constraints in (3) to force U to have a block diagonal form D_W. Further constraints can be used to identify D_W and to impose stationarity constraints over occasions. In the extreme case, we might require $W_i = W$ for all i, resulting in stationary compositing weights. Less restrictive forms of stationarity can be required, equating selected elements of the compositing matrices W_i across occasions. We can also place lower bounds on the compositing weights, forcing all weights to be nonnegative, for example.

In all cases, the effect of imposing additional constraints can be judged by comparing the resulting value of R_F in (7) to it's value under a "just-identified" model. If the increase in R_F produced by imposing stationarity constraints is acceptably small, we can proceed with stationary components.

The limited information case. The limited information case can also be treated in terms of a function of the form in (2). We begin by taking the pattern matrix in (6) to be block diagonal, with nxm block diagonal matrices P_i being the regression weights in the ith occasion

$$P = \begin{vmatrix} P_1 & 0 & 0 & \ldots & 0 \\ 0 & P_2 & 0 & 0 & ..0 \\ & & & & \\ 0 & 0 & 0 & & P_p \end{vmatrix} \qquad (10)$$

The risk function simplifies in this case (restricting G to be block diagonal) to

$$R_L = \sum_i tr\{G_i(X_i-\hat{X}_i)(X_i-\hat{X}_i)'\} \qquad (11)$$

The pattern matrices P_i which minimize R_L are of the form

$$P_i = \Sigma_{ii}W_i(W_i'\Sigma_{ii}W_i)^{-1} \qquad (12)$$

and again we can choose the W_i to minimize R_L. Without further constraints, the solution here is simply the usual "within-occasion" principal component solution.

We can formulate the problem in terms of a function H by defining D_G to be the block diagonal weighting matrix, and the matrix D_S to be a block diagonal matrix

$$D_S = \begin{vmatrix} \Sigma_{11} & 0 & 0 & \dots & 0 \\ 0 & \Sigma_{22} & 0 & 0 & ..0 \\ & & & & \\ 0 & 0 & \dots & & \Sigma_{pp} \end{vmatrix} \qquad (12)$$

Then R_L is minimized by choosing D_W to maximize

$$H(D_W; D_S D_G D_S, D_W) = \mathrm{tr}\{D_W' D_S D_G D_S D_W (D_W' D_S D_W)^{-1}\} \qquad (13)$$

We can again use our algorithm to impose constraints on D_W. All of the constraints considered in the full information case are applicable here. The effect of imposing various constraints can again be evaluated by comparing the resulting value of R_L to the value given by the "just-identified" principal component solution.

Discussion

In general, the full information solution will provide a better approximation to X in that $R_L > R_F$. However, the limited information solution is closer in spirit to the traditional principal component solution. Furthermore, a solution may not exist in the full information case if the components are highly correlated across occasions. Both solutions yield component scores which are correlated, both within and across occasions.

The foregoing development has focused upon the compositing weight matrix D_W, the weights used to derive the component scores from the observed variables. But we need not restrict our attention to the compositing weights, as either (9) or (13) can be written in terms of the component pattern or component structure matrices. In the full information case, the pattern matrix P is given in (8), and the structure matrix is $R = \Sigma D_W$. We can write H in (9) in terms of either P or R

$$H(P;G,\Sigma^{-1}) = tr\{P'GP(P'\Sigma^{-1}P)^{-1}\} \quad (14)$$

$$H(R;G,\Sigma^{-1}) = tr\{R'GR(R'\Sigma^{-1}R)^{-1}\} \quad (15)$$

Similarly, in the limited information case, we can express the block diagonal pattern and structure matrices as $D_P = D_S D_W (D_W' D_S D_W)^{-1}$ and $D_R = D_S D_W$. The function H in (13) can be written in either of the following ways

$$H(D_P;D_G,D_S^{-1}) = tr\{D_P' D_G D_P (D_P' D_S^{-1} D_P)^{-1}\} \quad (16)$$

$$H(D_R;D_G,D_S^{-1}) = tr\{D_R' D_G D_R (D_R' D_S^{-1} D_R)^{-1}\} \quad (17)$$

The point of this development is that we may apply constraints of the form in (3) to the component pattern or structure matrices, using our algorithm to maximize H. For example, we can impose stationarity constraints on the pattern or structure matrices. Naturally, we might employ such constraints to achieve simple structure as well.

To conclude, the component analysis method presented in this paper provides a flexible alternative to currently available approaches for the analysis of multivariate longitudinal data. I have written a program in FORTRAN to perform the analysis, although the program is not "exportable" at this time. An example of an application of the component method in the limited information case will appear this year in Psychology and Aging (Haan, Millsap, & Hartka, 1986).

REFERENCES

Cattell, R. B. (1963). The structuring of change by P-technique and incremental R-technique. In C. W. Harris (Ed.) Problems in measuring change. Madison: University of Wisconsin Press.

Corballis, M. C. & Traub, R. E. (1970). Longitudinal factor analysis. Psychometrika, 35, 79-98.

Haan, N., Millsap, R. E. & Hartka, E. (1986). As time goes by: Change and stability in personality over fifty years. Psychology and Aging, in press.

Hakstian, A. R. (1973). Procedures for the factor analytic treatment of measures obtained on different occasions. British Journal of Mathematical and Statistical Psychology, 26, 219-239.

Harshman, R. A. & Berenbaum, S. A. (1981). Basic concepts underlying the PARAFAC-CANDECOMP three-way factor analysis model and its application to longitudinal data. In D. H. Einhorn, J. A. Clausen, N. Haan, M. P. Honzik & P. H. Mussen (Eds.) Present and past in middle life. New York: Academic Press.

Joreskog, K. G. & Sorbom, D. (1977). Statistical Models and methods for analysis of longitudinal data. In D. J. Aigner & A. S. Goldberger (Eds.) Latent variables in socio-economic models. Amsterdam: North-Holland.

Kroonenburg, P. M. & De Leeuw, J. (1980). Principal component analysis of three-mode data by means of alternating least squares algorithms. Psychometrika, 45, 69-97.

McArdle, J. J. (1986). Dynamic bu. structural equation modeling with repeated measures data. In J. R. Nesselroade & R. B. Cattell (Eds.) Handbook of mulivariate experimental psychology Vol II, New York: Plenum.

McDonald, R. P. (1984). The invariant factors model. In H. G. Law, C. W. Snyder Jr., J. A. Hattie & R. P. McDonald (Eds.) Research methods for multimode data matrices, New York: Praeger.

Meredith, W. & Millsap, R. E. (1985). On component analysis. Psychometrika, 50, 495-507.

Meredith, W. & Tisak, J. (1984). Tuckerizing curves. Paper presented at the annual meeting of the Psychometric Society, June 21, at the University of California, Santa Barbara, CA.

Molenaar, P. C. M. (1985). A dynamic factor model for the analysis of multivariate time series. Psychometrika, 50, 181-202.

Mulaik, S. (1972). The foundations of factor analysis. New York: McGraw-Hill.

Rogosa, D. R., Brandt, D. & Zimowski, M. (1982). A growth curve approach to the measurement of change. Psychological Bulletin, 90, 726-748.

Rogosa, D. R. & Willett, J. B. (1985). Understanding correlates of change by modeling individual differences in growth. Psychometrika, 50, 203-228.

Sands, R. & Young, F. W. (1980). Component models for three-way data: An alternating least squares algorithm with optimal scaling features. Psychometrika, 45, 39-67.

Swaminathan, H. (1984). Factor analysis of longitudinal data. In H. G. Law, C. W. Snyder Jr., J. A. Hattie & R. P. McDonald (Eds) Research methods for multimode data matrices. New York: Praeger.

Tisak, J. (1984). Exploratory longitudinal factor analysis in multiple populations with applications to growth in intelligence. Unpublished doctoral dissertation, University of California-Berkeley.

Tucker, L. R. (1963). Implications of factor analysis of three-way matrices for measurement of change. In C. W. Harris (Ed.) Problems in measuring change. Madison: University of Wisconsin Press.

Tucker, L. R. (1964). The extension of factor analysis to three-dimensional matrices. In H. Gulliksen & N. Fredericksen (Eds.) Contributions to mathematical psychology. New York: Holt, Rinehart, & Winston.

Tucker, L. R. (1966). Some mathematical notes on three-mode factor analysis. Psychometrika, 31, 279-311.