Analysis of Qualitative Variables in Structural Models with Unique Solutions

Giorgio Vittadini

Università degli Studi di Milano,
20126 – Viale Sarca 202
Milano, Italia, e-mail: vittadin@imiucca.csi.unimi.it

Abstract: A new method based on the Multidimensional Scaling and the Restricted Regression Component Decomposition is proposed in order to obtain solutions for structural models with mixed variables.

Keywords: Structural Models with Mixed Variables, Normality Hypothesis, Multidimensional Scaling, Alternating Least Squares, Restricted Regression Component Decomposition.

1. Structural Model with Qualitative Variables

The structural model for the study of causal relationship among latent variables is composed of one structural and two measurement equations:

$$H=HB+\Xi\Gamma+E=\Xi\Gamma(I-B)^{-1}+E(I-B)^{-1};\ Y=H\Lambda_y+U;\ X=\Xi\Lambda_x+U \qquad (1)$$

where $Y'=(y'_{(1)},\ldots,y'_{(t)})(n_y,t)$, $X'=(x'_{(1)},\ldots,x'_{(t)})(n_x,t)$ are the observed mixed variables; $\Xi'=(\xi'_{(1)},\ldots,\xi'_{(t)})(n_\xi,t)$, $H'=(\eta'_{(1)},\ldots,\eta'_{(t)})(n_\eta,t)$ are the latent variables; $E'=(\varepsilon'_{(1)},\ldots,\varepsilon'_{(t)})$ (n_ε,t) are the errors in equations; $\Delta'=(\delta'_{(1)},\ldots,\delta'_{(t)})$ $(n_\delta t)$, $U'=(u'_{(1)},\ldots,u'_{(t)})$ (n_u,t), are the errors in variables. It is assumed that: all the random variables have zero mean and finite variance, B is a low matrix with zero on the main diagonal, (Y, X, H) are identically distributed and (Ξ, E, Δ, U) are identically and independently distributed. The model is usually proposed with restrictions on parameters and on covariances.

The solutions are reached starting from the variance covariance matrix of the reduced model where the variables H in the measurement models are substituted by the value of $(\Xi\Gamma(I-B)^{-1}+(I-B)^{-1}E)$ obtained from the structural equations.

2. Methods for Obtaining Solutions

First of all given that a continuous bivariate normal variable (J_l^*, J_m^*) with distribution $\Phi\left(j_l^*, j_m^*, \rho_{(j_l^* j_m^*)}\right)$ underlies every pair of ordinal bivariate observed variables (J_l, J_m), the polychoric correlation between the two components of them is calculated. When there is one only ordinal variable, the polyserial correlation coefficient between an observed quantitative variable and the normal underlying variable is defined. The observed frequencies of the qualitative variables $n_{j_{l_q} j_{m_r}}$ $(q = 1,...n_l; r = 1,...,n_r)$ given, the polychoric coefficients are reached in different ways. Jöreskog (1994) hypothesizes that the marginal probability of the normal variables are equal to the marginal frequencies of the two-way table $(\pi_{j_{l_q}\bullet} = n_{j_{l_q}\bullet})$ $(\pi_{j_{m_r}\bullet} = n_{j_{m_r}\bullet})$ of the ordinal variables. Therefore, first of all, he reaches the thresholds, using the marginal distributions of the normal bivariate distribution; for given thresholds using such distribution, he obtains the correlation coefficient maximizing the log likelihood of the sample respect to $\rho_{(j_l^* j_m^*)}$. Lee et al. (1990) estimate the thresholds by means of Partition Maximum Likelihood (PML) which is simpler from the computational point of view.

Lee et al. (1995) estimate simultaneously the correlation coefficients and the thresholds concerning pairs of variables maximizing the log likelihood of the sample respect to $\rho_{(j_l^* j_m^*)}$ and respect to every threshold $j_{l_q}^*, j_{m_r}^*$, by means of PML (but Jöreskog (1994) observes that "different estimates of thresholds for one variable may be obtained from different pairs of variables").

By means of Full Maximum Likelihood, in one case Lee et al. (1992) simultaneously reach all the thresholds of the polychoric correlations; in another case (Lee et al. (1990)), they reach also the parameters, the variances and the covariances of the latent variables. Moreover these two last methods take up too much computer time (Lee et al. (1990) Lee et al. (1995)).

In the first three methods, the parameters and the covariances of the latent variables and errors are obtained from the polychoric or polyserial correlations. There is not a unique understanding about the method of obtaining the parameters and the latent variables. Jöreskog (1990), (1994), Rigdon and Ferguson (1991) propose the Weighted Least Squares method and criticise the Maximum Likelihood method because the standard error parameter estimates are asymptotically incorrect, but Lee et al. (1995) continue to prefer the General Least Squares method and criticize the proposal of Jöreskog because it requires sample sizes larger than 200 and more computer time.

3. Some Critical Observations

Comparing, in some Montecarlo studies, different correlation coefficients when the underlying bivariate distribution is normal, the polychoric correlation is shown to be "the best in the sense of being closest to the true correlation" (Quiroga (1992)) [even if in a Montecarlo study Babakus et al. (1987) show that the polychoric coefficient provides the best estimates of model parameters and the worst fit statistics]. However, Muthen (1984), Aish and Jöreskog (1990) and Quiroga (1992) say also that "the assumption of underlying bivariate normality is too strong for most ordinary variables used in social sciences". There are not many studies on the use of polychoric coefficient with underlying not normal distribution. Lee and Lam (1988) study the robustness of polychoric coefficient only when the underlying distribution is elliptical (containing multivariate normal, platicentric and leptocentric distributions). Lee et al. (1995) even though obtaining quite satisfactory results about such robustness with moderate size random samples, say that "to draw a non definite conclusion, a longer simulation is needed". Moreover in literature only Quiroga (1992) tries, in a systematic way, to extend the distributive hypothesis of the continuous variables underlying the qualitative variables. First of all, she says that when you leave the normality assumptions, such variables are surely not consistent and slightly biased (Quiroga (1992)). Moreover, in a Montecarlo study she verifies that, when the underlying distribution is a skew-normal bivariate, the polychoric coefficient is the best choice only for sample size of 200-400 and large number of categories (5-9) and underestimates the true correlation coefficient. Then, by means of a measure of not normality, she verifies that when the underlying distributions are generated by the Fleishman-Vale-Maurelli polynomial transformation (with departure from normality due to skewness and kurtosis) the polychoric coefficient is robust, but overestimates the true correlation. Finally she proposes an extended polychoric coefficient with distribution given by a mixture of a normal and univariate skew-normal density function but she does not give empirical verifications. Therefore, until now, there are neither theoretical demonstrations nor empirical simulations which give satisfactory and generally valid reasons for using polychoric coefficient with underlying not normal distribution.

Moreover the solutions based on polychoric coefficients: are not sensible to different scale of qualitative variables because they generally deal only with ordinal variables, reach solutions from variance covariance-covariance matrix of the reduced model different from the solutions obtained from the observed variables of the original measurement models and have the same problems of non identification of parameters and indeterminacy of latent variables of structural models with quantitative variables (Vittadini 1989).

4. An Alternative Proposal

In the quantitative case the problems of not uniqueness of the solutions are resolved by using linear combinations instead of causal latent variables of the observed variables (Wold (1982), Haagen and Vittadini (1991)). In this paper we propose to obtain the latent variables of the model as linear combinations of the observed mixed variables simultaneously quantified, by means of methods of multidimensional scaling using simultaneously optimal scaling and ordinary least squares method. In fact such methods resolve the problem of not normality of the variables because are distribution free (Young (1981)) and give unique solutions once chosen the method of multidimensional scaling. In order to avoid subjective choices about methodologies of multidimensional scaling (and therefore subjective solutions), we propose: to quantify the qualitative variable and to obtain the linear combinations of them by means of a unique objective function; to reach flexible solutions as regards to different kinds of linear combinations requested by the problems; to take into account the different scale of the qualitative variable. Therefore, among the variety of multidimensional scale methods we choose the family of Alsos method (Young (1981), and Keller and Wansbeek (1983)). These methods are based on alternating optimal scaling which quantifies qualitative variables and ordinary least squares which reach linear combinations of them in a iterative way. So they obtain solutions in a different way along the scale of the variables (ordinal-nominal, continuous-discrete), and the aim of analysis (e.g. Principal components, canonical correlation) giving answers to previous problems. Among the family of Alsos methods we avoid methods such as OSMOD (Saito and Otsu (1988)) or INDOMIX-CAMIX (Kiers (1991)) which obtain solutions in two stages. Instead we choose methods that simultaneously obtain quantifications of qualitative variables and their linear combinations (ADDALS (De Leeuw, Young, Takane (1976)) MORALS CORALS (Young, De Leeuw Takane, (1976)), PRINCALS (De Leeuw and Van Rijckevorsel (1980)) OVERALS (Van Der Burg and De Leeuw (1988)) respectively from the perspective of variance analysis, canonical correlation, principal components, multiple correspondence analysis. Moreover in order to obtain the latent variables from their real indicators as in Wold (1982), we apply the chosen Alsos methods to the subsets of mixed variables Y_β, X_δ characterized by submatrices $l_{y(\beta,\bullet)}, l_{x(\delta,\bullet)}$ with coefficients all different from zero. So we simultaneously obtain the quantification Y_β^*, X_δ^* of such mixed variables Y_β, X_δ and their linear transformations $\tilde{\eta}_\beta, \tilde{\xi}_\delta$ according to different aim of the analysis (e.g. canonical correlation, principal analysis etc.). Then in order to take into account the restrictions:

$$cov(\eta_\beta, \eta_\pi) = 0; cov(\xi_\delta, \xi_\gamma) = 0; cov(\delta_{\beta_k}, \delta_{\beta_s}); cov(u_{\delta_\lambda}, u_{\delta_\mu}) = 0;$$
$$b_{(\beta,\mu)} = 0; \gamma_{(\delta,\varphi)} = 0; \; l_{y(\beta,\alpha)} = 0; l_{x(\delta,\nu)} = 0; \qquad (2)$$

we obtain by means of the Restricted Regression Component Decomposition (RRCD) of quantified variables Y^*, X^* (Haagen and Vittadini (1998)) by means of an iterative process:

$$\eta_\beta^+ = Q_{\cdot\eta_\mu \cup \cdot y_\alpha} \eta_\beta^0 \; (\eta_\beta^0 = Q_{\tilde{\eta}_\pi} \tilde{\eta}_\beta \; (\beta \neq \pi); {}^*\eta_\mu = Q_{H^\circ_{(\beta,\mu)}} \eta_\mu^\circ \; ; {}^*y_\alpha = Q^\circ_{H_{(\beta)}} y_\alpha \;)$$
$$\xi_\delta^+ = Q_{\cdot\eta_\varphi \cup \cdot x_\nu} \xi_\delta^0 \; (\xi_\delta^0 = Q_{\tilde{\xi}_\gamma} \tilde{\xi}_\delta \; (\gamma \neq \delta); {}^*\eta_\varphi = Q_{H^\circ_{(\varphi)} \cup \Xi^\circ_{(\delta)}} \eta_\varphi^\circ; {}^*x_\nu = Q_{\Xi^\circ_{(\delta)}} x_\nu \;)$$
$$\tilde{\varepsilon}_j = Q_{\tilde{H}_{(j)} \cup \tilde{\Xi}} \tilde{\eta}_j \; ((j) = 1, \ldots j\text{-}1); \; \delta^0_{\beta_k} = Q_{\left(H^\circ \cup \Xi^\circ \cup y_{\beta_s}\right)} y_{\beta_k}; u^0_{\delta_\lambda} = Q_{\left(X \cup \Xi^\circ \cup x_{\delta_\mu}\right)} x_{\delta_\lambda} \quad (3)$$

where $H^\circ_{(\mu,\beta)}$ are the H° without $\eta^\circ_\mu, \eta^\circ_\beta$. $Q_{\tilde{\eta}_\pi}$ is the complement orthogonal to the orthogonal projector on the space generated by $\tilde{\eta}_\pi$ and the other symbols are defined in a similar way. So we have the following RRCD of $y_{\beta_k}, x_{\delta_\lambda}, \eta^\circ_\beta$:

$$Q_{\Xi^\circ \cup y^\circ_{\beta_s}/H^\circ} y_{\beta_k} = P_{H^\circ} y_{\beta_k} + Q_{H^\circ \cup \Xi^\circ \cup y_{\beta_s}} y_{\beta_k}; \; Q_{Y \cup x_{\delta_\mu}/\Xi^\circ} x_{\delta_\lambda} = P_{\Xi^\circ} x_{\delta_\lambda} + Q_{Y \cup x_{\delta_\mu} \cup \Xi^\circ} x_{\delta_\lambda};$$
$$\eta^\circ_\beta = P_{H^\circ_{(\beta)}} \eta^\circ_\beta + P_{\Xi^\circ / H^\circ_{(\beta)}} \eta^\circ_\beta + Q_{H^\circ_{(\beta)} \cup \Xi^\circ} \eta^\circ_\beta \qquad (4)$$

5. Numerical Example

The following variables are observed on a sample of 150 families casually chosen from 4103 american families that have been codified in the Federal Reserve Board research regarding National Income and Wealth of 1983.
Y_1 Job contract household (y_{11}), spouse (y_{12}); Occupation kind household (y_{13}), spouse (y_{14}); Occupation sector household (y_{15}), spouse (y_{16}). Y_2 Total health (y_{21}); Income (y_{22}); Debt (y_{23}). X_1 Age household (x_{11}), spouse (x_{12}); Sex household (x_{13}), spouse (x_{14}); Number of children (x_{15}); Race (x_{16}); Residence region (x_{17}); Civil Status (x_{18}); X_2 Educational Level household (x_{21}), spouse (x_{22}); Full time job years household (x_{23}), spouse (x_{24}) ; Part time job years household (x_{25}), spouse (x_{26}); Latent variables: labour force (η_1); Health and income (η_2), Civil status (ξ_1), Instruction grade (ξ_2).
In order to verify the causal dependence of the latent variables H from the latent variables Ξ we use the alternative proposal shown in paragraph 4 with the following restrictions: $l_{x_{(Y_1,\eta_2)}} = 0, l_{y_{(Y_2,\eta_1)}} = 0, l_{x_{(x_1,\xi_2)}} = 0, l_{x_{(x_2,\xi_1)}} = 0$, $cov(\Delta_1, \Delta_2) = 0$, $cov(U_1, U_2) = 0$. The qualitative variables are quantified and the latent variables

are obtained as principal components with Princals method, the restrictions are then taken into account by means of RRCD. The variance-covariance of the observed variables and the results are shown in table 1.

Table 1: *The alternative proposal for the sample of American families.*

$$S_t$$

$$\begin{bmatrix}
0.0393 & -0.0284 & -0.0085 & 0.0248 & -0.0049 & 0.0580 & 10.853 & -0.0165 & -0.0785 \\
-0.0284 & 8.2791 & -3.2858 & 0.0917 & 0.8073 & -0.3585 & -79.107 & 13.671 & 0.5563 \\
-0.0085 & -3.2858 & 17.0934 & -0.0313 & -0.6348 & 4.1994 & -325.60 & -17.426 & -1.1561 \\
0.0248 & 0.0917 & -0.0313 & 0.1490 & 0.0107 & -0.0858 & 80.249 & 8.5699 & -0.0821 \\
-0.0049 & 0.8073 & -0.6348 & 0.0107 & 2.5359 & -0.7303 & -162.35 & 16.437 & 0.4694 \\
0.0580 & -0.3585 & 4.1994 & -0.0858 & -0.7303 & 11.735 & -65.583 & -57.510 & 0.1070 \\
10.853 & -79.107 & -325.696 & 80.249 & -0.0162 & -65.583 & 320931 & 40682 & -397.51 \\
-0.0165 & 13.671 & -17.426 & 8.5699 & 16.437 & -57.510 & 40682.1 & 17114 & -21.173 \\
-0.0785 & 0.5563 & -1.1561 & -0.0821 & 0.4694 & 0.1070 & -397.51 & -21.173 & 7.4133
\end{bmatrix}$$

$$S_\varepsilon$$

$$\begin{bmatrix}
1003.4 & -16.291 & 5.9262 & 12.573 & -16.997 & -3.0333 & 37.233 & -11.617 & 9.0494 & -3.1692 & 1.8904 & -2.0307 & 34.131 & -69.770 \\
-16.291 & 109.04 & -5.5373 & 3.0773 & 0.7344 & -1.1737 & 7.0739 & 0.3456 & -2.5367 & -2.8791 & 197.29 & -2.1334 & 45.014 & 14.060 \\
5.9262 & -0.5537 & 9.3333 & -0.1390 & -0.0330 & -0.1111 & 1.2672 & -0.0332 & -0.2626 & -0.1623 & -6.5191 & 2.4767 & 0.6669 & -0.9171 \\
12.573 & 3.0773 & -1.3799 & 3.2374 & -0.1312 & -0.7070 & -4.0200 & -0.2396 & -3.5546 & -0.3074 & 7.7005 & -9.1676 & 3.3415 & 0.3945 \\
-16.997 & 0.7344 & -3.3033 & -0.1312 & 1.1152 & -0.5375 & -2.3432 & 0.3512 & 2.0414 & 0.0965 & -0.2343 & 6.1353 & -0.7796 & 1.2531 \\
-3.0333 & -1.1173 & -1.1111 & -0.0007 & -0.0057 & 1.0000 & 0.3016 & -0.0033 & 1.5454 & 0.0109 & -2.1712 & -1.6969 & -0.0370 & -0.0547 \\
37.233 & 7.0739 & 1.2672 & -4.0200 & -2.3432 & 3.0161 & 197.31 & -1.3076 & -2.5917 & -8.4634 & 19.037 & 4.1361 & 14.046 & 0.2234 \\
-11.617 & 0.3456 & -3.3332 & -0.2396 & 0.3512 & -2.9393 & -1.3076 & 0.2403 & -2.3333 & 0.0719 & 0.4641 & 1.9576 & -0.2342 & 0.3446 \\
9.0494 & -2.5367 & -0.2626 & -0.3564 & 0.2041 & 1.5454 & -2.5917 & -0.5033 & 4.0091 & 2.4420 & -5.7035 & 1.2327 & -0.6716 & -0.3408 \\
-3.1692 & -2.8791 & -1.6222 & 0.2074 & 0.0965 & 1.0909 & -8.4634 & 0.0719 & 2.4420 & 4.2763 & -4.3171 & 2.7717 & 0.4076 & 0.2460 \\
1.8904 & 197.29 & -6.5191 & 7.7005 & -0.2343 & -1.7121 & 19.037 & 0.4641 & -5.7035 & -4.3171 & 127.34 & -1.2349 & 51.225 & 13.152 \\
-2.0307 & -2.1334 & 2.4767 & -9.1676 & 6.1353 & -1.6969 & 4.1361 & 1.9767 & 1.2327 & 0.2771 & -1.2349 & 1.8704 & -2.8137 & 2.2771 \\
34.131 & 45.014 & 6.6969 & 3.3415 & -0.7796 & -3.7070 & 14.046 & -0.2341 & -6.7161 & 0.4076 & 51.225 & -2.8137 & 84.526 & -1.4733 \\
-69.770 & 14.060 & -9.7171 & 0.3945 & 1.2531 & -5.4747 & 0.2234 & 0.3446 & -3.4020 & 0.246 & 13.152 & 2.2771 & -1.4733 & 47.498
\end{bmatrix}$$

$$B \quad \begin{bmatrix} 0 & -0.0962 \\ -0.6957 & 0 \end{bmatrix}$$

$$\Gamma \quad \begin{bmatrix} 0.0171 & 0 \\ 0 & 0.0042 \end{bmatrix}$$

$$I_{y(Y_1,\eta_1)} \quad \begin{bmatrix} 0.0039 & -1.9313 & 2.2994 & -0.0765 & -0.5419 & 1.7726 \end{bmatrix}$$

$$I_{y(Y_2,\eta_2)} \quad \begin{bmatrix} 340.5859 & 93.4431 & -1.3412 \end{bmatrix}$$

$$I_{x(X_1,\xi_1)} \quad \begin{bmatrix} 22.6810 & -1.2115 & 0.0563 & 0.5079 & -0.7005 & 0.0106 & 3.7964 & -0.4775 \end{bmatrix}$$

$$I_{x(X_2,\xi_2)} \quad \begin{bmatrix} 1.7533 & 1.6531 & -3.5028 & 2.9609 & -3.9295 & 3.3142 \end{bmatrix}$$

$$S_H \quad \begin{bmatrix} 0.7220 & -0.0767 \\ -0.0767 & 0.4585 \end{bmatrix}$$

$$S_\Xi \quad \begin{bmatrix} 0.9393 & -0.0899 \\ -0.0899 & 0.9344 \end{bmatrix}$$

$$S_E \quad \begin{bmatrix} 0.9305 & 0.0872 \\ 0.0872 & 0.9253 \end{bmatrix}$$

$$S_{\Delta_1} \quad \begin{bmatrix}
0.0375 & -0.0053 & -0.0707 & 0.0297 & 0.0137 & 0.0483 \\
-0.0053 & 4.5337 & 1.1799 & -0.0355 & -0.3292 & 2.9146 \\
-0.0707 & 1.1799 & 11.0353 & 0.0189 & 0.8952 & 0.4736 \\
0.0297 & -0.0355 & 0.0189 & 0.1396 & 0.0123 & 0.0542 \\
0.0137 & -0.3292 & 0.8952 & 0.0123 & 2.1494 & 0.1236 \\
0.0483 & 2.9146 & 0.4736 & 0.0542 & 0.1236 & 10.7678
\end{bmatrix}$$

$$S_{\Delta_2} \quad \begin{bmatrix} 143417.7 & 6456.7 & 118.6 \\ 6456.7 & 7047.8 & 81.98 \\ 118.6 & 81.93 & 3.442 \end{bmatrix}$$

$$S_{U_1} \quad \begin{bmatrix}
450.7216 & -2.2195 & -0.4063 & 0.9215 & -2.2592 & -0.4451 & -14.7014 & -1.4334 \\
-2.2195 & 95.5067 & -0.4428 & 7.4965 & -0.2513 & -0.0554 & 9.7258 & 0.3422 \\
-0.4063 & -0.4428 & 0.0841 & -0.1741 & 0.0138 & -0.0256 & 0.7438 & -0.0034 \\
0.9215 & 7.4965 & -0.1741 & 2.7618 & 0.1564 & -0.0112 & -5.1128 & -0.0032 \\
-2.2592 & -0.2513 & 0.0138 & 0.1564 & 0.5211 & -0.0030 & -0.0254 & 0.0002 \\
-0.4451 & -0.0554 & -0.0036 & -0.0112 & -0.0030 & 0.0281 & 0.0583 & 0.0007 \\
-14.7014 & 9.7258 & 0.7438 & -5.1128 & -0.0254 & 0.0583 & 12.4576 & 0.0117 \\
-1.4334 & 0.3422 & -0.0034 & -0.0032 & 0.0002 & 0.0007 & 0.0117 & 0.0039
\end{bmatrix}$$

$$S_{U_2} \quad \begin{bmatrix}
1.0039 & 0.1470 & 0.9455 & -1.0903 & 1.7255 & -0.9186 \\
0.1470 & 1.3709 & 1.3589 & -1.5341 & 2.1256 & -0.5184 \\
0.9455 & 1.3589 & 10.5982 & -2.8086 & 1.1425 & 0.5277 \\
-1.0903 & -1.5341 & 2.8086 & 10.1653 & 2.9774 & -1.6561 \\
1.7255 & 2.1256 & 1.1425 & 2.9774 & 45.6884 & -10.1856 \\
-0.9186 & -0.5184 & 0.5277 & -1.6561 & -10.1856 & 23.5754
\end{bmatrix}$$

With the example we can verify that the alternative proposal respects all the properties of the structural model described in paragraph 1 and the restrictions indicated in this paragraph. But the alternative proposal obtains unique

solutions solving all the problems of non-identification of parameters and indeterminacy of latent variables of the structural models with qualitative variables.

References

Aish, A. M., Jöreskog, K. G. (1990). A panel model for political efficacy and responsiveness: an application of LISREL7 with weighted least squares, *Quality and Quantity*, 24, 405-426.

Babakus, E., Ferguson, C.E. Jr, Jöreskog K.G. (1987). The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions, *Journal of Marketing Research*, 24, 222-228.

De Leeuw, J., Van Rijckevorsel, J.(1980). *Homals & princals, some generalizations of principal components analysis*, In E. Diday et al. (Eds.), Data Analysis and Informatics, North-Holland Publishing Company, 231-241.

De Leeuw, J., Young, F.W., Takane Y. (1976). Additive structure in qualitative data: an alternating least squares method with optimal scaling features, *Psychometrika*, 41, 471-503.

Haagen, K., Vittadini, G. (1991). Regression Component Decomposition in Structural Analysis, *Communications in Statistics*, 20, 1153-1161.

Haagen, K., Vittadini, G. (1998). Regression Component Decomposition Restricted. Un'alternativa al Lisrel model, *Metron*, 56, 1-2, in corso di pubblicazione.

Jöreskog, K.G. (1990). New developments in Lisrel: analysis of ordinal variables using polychoric correlations and weighted least squares, *Quality and Quantity*, 24, 387-404.

Jöreskog, K.G. (1994). On the estimation of polychoric correlations and their asymptotic covariance matrix, *Psychometrika*, 59, 3, 381-389.

Kiers, H. A. L. (1991). Simple structure in component analysis techniques for mixtures of qualitative and quantitative variables, *Psychometrika 56, 2, 197-212.*

Keller, W. J., Wansbeek, T. (1983). Multivariate methods for quantitative and qualitative data, *Journal of Econometrics*, 22, 91-111.

Lee, S.Y, Lam, M. L.(1988). Estimation of polychoric correlation with elliptical latent variables. *Journal of statistic Computation and Simulation*, 30, 173-188.

Lee, S.Y., Poon, W.Y., Bentler, P.M. (1990). Full maximum likelihood analysis of structural equation models with polytomous variables, *Statistics and Probability Letters*, 9, 91-97.

Lee, S.Y., Poon, W.Y., Bentler, P.M. (1995). A two-stage estimation of structural equation models with continuous and polytomous variables, *British Journal of Mathematical and Statistical Psychology*, 48, 339-358.

Muthen, B. (1984). A general structural equation model with dichotomous, ordered categorical, and continuous latent variable indicators, *Psychometrika*, 49, 1, 115-132.

Quiroga, A.M. (1992). *Studies of the Polychoric Correlation and other Correlation Measures for Ordinal Variables*, PhD thesis, Uppsala University.

Rigdon, E.E., Ferguson, C.E. Jr. (1991) The performance of the polychoric correlation coefficient and selected fitting functions in confirmatory factor analysis with ordinal data, *Journal of Marketing Research*, 28, 491-497.

Saito, T., Otsu, T. (1988). A method of optimal scaling for multivariate ordinal data and its extensions, *Psychometrika*, 53, 1, 5-25.

Van Der Burg, E., De Leeuw, J.(1988). Homogeneity analysis with k sets of variables: an alternating least squares method with optimal scaling features, *Psychometrika*, 53,2, 177-197.

Vittadini, G. (1989). *Indeterminacy Problems in the LISREL Model*, in Multivariate Behavioral Research, Forth Worth (Texas), 24, 4, 397-414.

Wold, H. (1982). *Soft Modelling: the basic design and some extensions*, in Jöreskog K.G., Wold H., *Systems under indirect observation: casuality, structure, prediction*, North - Holland, Amsterdam, 2, 1-54.

Young, F.W. (1981). Quantitative analysis of qualitative data, *Psychometrika*, 46, 357-388.

Young, F.W., De Leeuw, J., Takane, Y. (1976). Regression with qualitative and quantitative variables: an alternating least squares method with optimal scaling features, *Psychometrika*, 41, 4, 505-529.