A Causal Model with Latent Variables and Unique Solutions

Un Modello Causale Con Variabili Latenti a Soluzioni Uniche

Giorgio Vittadini
Milan Bicocca University – Department of Statistics
giorgio.vittadini@unimib.it

Klaus Haagen
Trento University – Department of Statistics
khaagen@cs.unitn.it

Riassunto: Per ovviare al problema della non unicità delle variabili latenti nei modelli strutturali in passato è stato proposto di ricavarle come combinazioni lineari delle variabili osservate. In questo lavoro si ottengono, mediante il metodo delle "osservazioni replicate" proposto da Haagen e Oberhofer (1999), le variabili latenti esogene come cause dei rispettivi indicatori osservati in modo univoco. Successivamente, nell'ipotesi che sia impossibile ricavare le variabili latenti endogene e gli errori delle equazioni mediante "osservazioni replicate" li si ottengono come cause uniche delle variabili osservabili utilizzando le ipotesi sottostanti il modello strutturale espresso in forma ridotta.

Keywords: Structural Model, Identificability, Indeterminacy, RCFM, Path diagram, PLS, RCD.

1. Introduction

It has been already demonstrated that the parameters of the structural model (Jöreskog, 1973, 1981) are not identifiable and the latent variables are indeterminate under general conditions (Bentler, 1982; Vittadini, 1988 and 1989). New methods for obtaining unique solutions have been proposed. The first proposal is the Partial Least Squares (PLS), valid for both the factorial model and the structural model (Nooan and Wold, 1982). Another one is the extension of the Regression Component Decomposition (RCD) (Schönemann and Steiger, 1976), to the structural models (Haagen and Vittadini, 1991; Haagen and Vittadini, for the Restricted Model, 1998). However the RCD and the PLS cannot be considered as structural models, because they are linear transformation methods (Vittadini, 1992).

An alternative method was proposed by Vittadini e Lovaglio (2001). Starting from the Path diagram and from the properties of the Partial Least Squares, the "exogenous latent variables" are obtained by means of a linear transformation of the observed variables, whereas the endogenous variables are still considered as causes of the respective indicators.

Haagen and Oberhofer (1999) have proposed the Replicated Common Factor Model (RCFM) which solves the indeterminacy problem of the common factors asymptotically. When the assumptions of the RCFM model are only satisfied for the measurement model with the endogenous variables it is possible to get unique solutions as causes for the exogenous variable and the errors in equation starting from the assumptions of the reduced model. All the proprieties of the structural model are satisfied. An example shows the validity of the theoretical results.

2. The structural model

The structural model is composed of one structural and two measurement equations:

$$\boldsymbol{\lambda}_{(i)}=\mathbf{A}\,\boldsymbol{\lambda}_{(i)}+\mathbf{P}\boldsymbol{\tau}_{(i)}+\boldsymbol{\gamma}_{(i)} \qquad (1)$$
$$\boldsymbol{x}_{(i)}=\mathbf{L}_X\boldsymbol{\lambda}_{(i)}+\boldsymbol{\delta}_{(i)} \qquad (2)$$
$$\boldsymbol{y}_{(i)}=\mathbf{L}_Y\boldsymbol{\tau}_{(i)}+\boldsymbol{\varepsilon}_{(i)} \qquad i=1,\ldots..n \qquad (3)$$

where $\boldsymbol{x}_{(i)}, \boldsymbol{y}_{(i)}$ are the i-eth observations on the variables $\boldsymbol{x}$ and $\boldsymbol{y}$ respectively.
It is assumed that the all random variables have zero mean and finite variance, and that $\boldsymbol{A}$ is a triangular matrix with zero on the main diagonal. $\boldsymbol{y}_{(i)}$, $\boldsymbol{x}_{(i)}$, $\boldsymbol{\lambda}_{(i)}$, $\boldsymbol{\tau}_{(i)}$ are vectors respectively with p, q, m, k components identically distributed, $\boldsymbol{\tau}_{(i)}, \boldsymbol{\gamma}_{(i)}, \boldsymbol{\delta}_{(i)}, \boldsymbol{\varepsilon}_{(i)}$ (i=1, …n) are identically and independently distributed.
Therefore, given $\boldsymbol{T}=(\boldsymbol{\tau}_{(1)},\ldots\ldots,\boldsymbol{\tau}_{(n)})$, $\Gamma=(\gamma_{(1)},\ldots\ldots,\gamma_{(n)})$, $\boldsymbol{\Delta}=(\boldsymbol{\delta}_{(1)},\ldots\ldots,\boldsymbol{\delta}_{(n)})$, $\boldsymbol{E}=(\varepsilon_{(1)},\ldots\ldots,\varepsilon_{(n)})$ it is hypothesized that:

(α) ($\boldsymbol{E}, \boldsymbol{\Delta}, \boldsymbol{\Gamma}$) and $\boldsymbol{T}$ independent;

(β) ($\boldsymbol{E}, \boldsymbol{\Delta}$) and $\boldsymbol{\Gamma}$ independent;

(γ) $\boldsymbol{E}$ and $\boldsymbol{\Delta}$ independent.

At this point if $\boldsymbol{X}=(\boldsymbol{x}_{(1)}, \ldots, \boldsymbol{x}_{(n)})$, $\boldsymbol{Y} = (\boldsymbol{y}_{(1)}, \ldots, \boldsymbol{y}_{(n)})$ are the n observations on the observed variables and $\Lambda = (\boldsymbol{\lambda}_{(1)}, \ldots \boldsymbol{\lambda}_{(n)})$ the (1) – (3) becomes:

$$\Lambda = \mathbf{A}\Lambda + \mathbf{P}\boldsymbol{T} + \Gamma = (\mathbf{I}\text{-}\mathbf{A})^{-1}\,\mathbf{P}\boldsymbol{T} + (\mathbf{I}\text{-}\mathbf{A})^{-1}\,\Gamma \qquad (4)$$

$$\boldsymbol{X}=\mathbf{L}_X\Lambda+\Delta \qquad (5)$$

$$\boldsymbol{Y}=\mathbf{L}_y\boldsymbol{T}+\boldsymbol{E} \qquad (6)$$

By means of (4) the model (5) – (6) can be expressed in the reduced form:

$$\begin{bmatrix}\boldsymbol{X}\\ \boldsymbol{Y}\end{bmatrix}=\begin{bmatrix}\mathbf{L}_x(\mathbf{I}-\mathbf{A})^{-1}\mathbf{P}; & \mathbf{L}_x(\mathbf{I}-\mathbf{A})^{-1}\\ \mathbf{L}_y & \mathbf{0}\end{bmatrix}\begin{bmatrix}\boldsymbol{T}\\ \Gamma\end{bmatrix}+\begin{bmatrix}\boldsymbol{0}\\ \boldsymbol{I}\end{bmatrix}\boldsymbol{E}+\begin{bmatrix}\boldsymbol{I}\\ \boldsymbol{0}\end{bmatrix}\Delta \qquad (7)$$

For (α), (β), the matrices $\mathbf{L}_x$, $\mathbf{L}_y$, **P** are respectively the regression coefficients matrices of Λ on $\boldsymbol{X}$, $\boldsymbol{T}$ on $\boldsymbol{Y}$, $\boldsymbol{T}$ on Λ. Every row $\mathbf{a}_\eta$ of **A** is the regression coefficients vector of Λ_η° on λ_η [$\Lambda_\eta^\circ = (\lambda_1,...,\lambda_{\eta-1})$].
Therefore from (7) we have:

$$\sum_{Y\Gamma} = \boldsymbol{0}; \tag{8}$$

and from (5):

$$\sum_{\Lambda\Delta} = \boldsymbol{0}; \tag{9}$$

We observe that the exogenous latent variables $\boldsymbol{T}$ are the causes of the observed indicators $\boldsymbol{Y}$ in the measurement model (6), and of the endogenous latent variables $\boldsymbol{\Lambda}$ in the structural model (4); the errors in equations Γ are the causes of the latent variables Λ in the structural model (4). Therefore the $\boldsymbol{T}$ and Γ are also the causes of the observed indicators $\boldsymbol{X}$ by means of Λ in the reduced model (7).
The problems of identification and indeterminacy for the structural models are the same to those discussed for the usual common factor model (Vittadini, 1988, 1989).

3. The replicated common factor model

In order to resolve the problem of identification and of indeterminacy of the factor model Haagen and Oberhofer (1992, 1999) have proposed the RFCM.
It is assumed that in the measurement model (6) every object (i) can be observed s-times. The vectors of errors ${}_t\varepsilon_{(i)}$ at every time t (t = 1, …, s) have zero mean and finite variance and are independent from every other vector, ${}_v\varepsilon_{(i)}$ concerning a different time v (v, t = 1, …, s). The common factors $\boldsymbol{T}$ are independent from the time t and from every ${}_t\varepsilon_{(i)}$ (t = 1, …, s).
The measurement model (6) in the RCFM becomes:

$$_s\boldsymbol{Y} = \mathbf{L}_Y \hat{\boldsymbol{T}} + {}_s\boldsymbol{E} \tag{10}$$

$$_s\boldsymbol{Y} = ({}_{(1)}\boldsymbol{Y},\ {}_{(2)}\boldsymbol{Y},\ldots,{}_{(s)}\boldsymbol{Y})\ (p,\ ns),\quad {}_s\boldsymbol{E} = ({}_{(1)}\boldsymbol{E},\ {}_{(2)}\boldsymbol{E},\ \ldots,\ {}_{(s)}\boldsymbol{E})\ (p,\ ns)\quad \hat{\boldsymbol{T}} = (\boldsymbol{T},\ldots,\boldsymbol{T})(\mathrm{m,ns})$$

Under general conditions and restrictions there are consistent maximum likelihood estimators for ($\mathbf{L}_y$, ${}_s\boldsymbol{E}$) by means of the RFCM (Oberhofer and Haagen 1992).
We know that (Guttman 1955) the solution for the measurement model (3) can be represented by:

$$\boldsymbol{T} = \mathbf{L}_{\mathbf{Y}}{}'\boldsymbol{S}_Y^{-1}\boldsymbol{Y}+\boldsymbol{Z} \tag{11}$$

with **Z** arbitrary random vector.
Oberhofer and Haagen (1992) demonstrate that for given $\mathbf{L}_Y$, $\boldsymbol{E}_S$ the asymptotic linear solution (when s→∞) is not indeterminate because:

$$T^* = \mathbf{L}_{Y'} \, S_{sY}^{-1} \, {}_sY \frac{1}{s}' = \mathbf{L}_{Y'} \, S_{sY}^{-1} \, \overline{Y}_S \tag{12}$$

where S_{sY} is the variance-covariance matrix of ${}_sY$ while $\overline{Y}_s$ are the correspondent means (with 1 (sq, 1)).
Therefore the measurement model (6) can be rewritten in the way along a RCD:

$$\overline{Y}_S = \overline{Y}_S \mathbf{P}_{T^*} + \overline{Y}_S Q_{T^*} \tag{13}$$

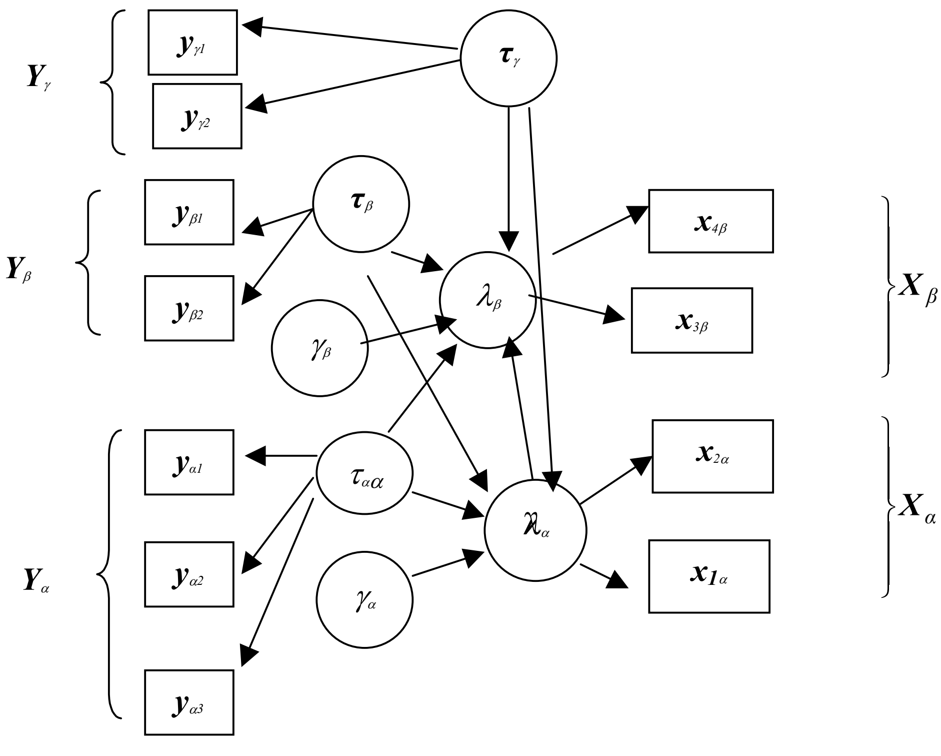
with:

$$E^* = \overline{Y}_S Q_{T^*} \tag{14}$$

4. Unique solutions for the structural models

If it is possible to obtain s replicated observations also on $\boldsymbol{X}$ we can obtain unique solutions by means of the RCFM. Otherwise from the latent variables $\boldsymbol{T^*}$ and from the reduced model (7) we can obtain also the latent variables Λ and the errors Γ as unique "causal" solutions.
In order to better understand the method for obtaining unique solutions, given the latent exogenous variables $\boldsymbol{T}^*$, we propose the following path diagram (Figure 1):

Figure 1



The specifications of paths suggest considering τ_α, τ_β, τ_γ the causes of $\boldsymbol{Y}_\alpha$, $\boldsymbol{Y}_\beta$, $\boldsymbol{Y}_\gamma$ respectively; λ_α, λ_β, the causes of $\boldsymbol{X}_\alpha$, $\boldsymbol{X}_\beta$ respectively; τ_α, τ_β, τ_γ, γ_α as the causes of λ_α and therefore by means of λ_α, of $\boldsymbol{X}_\alpha$; τ_α, τ_β, τ_γ, λ_α, γ_β are the causes of λ_β and therefore by means of λ_β of $\boldsymbol{X}_\beta$. Starting from the assumptions of the reduced model (7), the scores of Λ, Γ, E, Δ are performed.

Given $\boldsymbol{T}^*$, the first equation of the structural model in (4) and the first subset of observed variables $\boldsymbol{X}$ concerning λ_α in (5) are:

$$\lambda_\alpha = \mathbf{p}_\alpha \boldsymbol{T}^* + \gamma_\alpha \tag{15}$$

$$\mathbf{X}_\alpha = \boldsymbol{L}_{X_\alpha} \mathbf{p}_\alpha \boldsymbol{T}^* + \boldsymbol{L}_{X_\alpha} \boldsymbol{\gamma}_\alpha + \boldsymbol{\delta}_\alpha \tag{16}$$

We can define $\mathbf{p}_\alpha$ as the best linear combination of $\boldsymbol{T}^*$ that the better fits $\boldsymbol{X}_\alpha$.

$$\boldsymbol{L}_{X_\alpha} \mathbf{p}_\alpha = \boldsymbol{X}_\alpha \boldsymbol{T}^{*\prime} \boldsymbol{S}_{T^*}^{-1} \tag{17}$$

For (α) (8) γ_α must be uncorrelated with $\boldsymbol{T}^*$, $\overline{\boldsymbol{Y}}_s$ and therefore can be defined as a linear combination of $\boldsymbol{X}_\alpha$ $\boldsymbol{Q}_{\overline{Y}_s}$ (i.e. by means of the principal component, the canonical correlation, the reduced rank regression, where the linear transformation methods is suggested by the empirical problem studied).
We obtain:

$$\gamma_\alpha = \mathbf{k}_\alpha \boldsymbol{X}_\alpha \boldsymbol{Q}_{\overline{Y}_s} \tag{18}$$

with $\mathbf{k}_\alpha$ vector of parameters and

$$\boldsymbol{L}^*_{X_\alpha} = \boldsymbol{X}_\alpha \boldsymbol{\gamma}_\alpha{}' \boldsymbol{S}_{\gamma_\alpha}^{-1}. \tag{19}$$

From (17), (19) we have:

$$\mathbf{p}_\alpha^* = (\boldsymbol{L}^*_{X_\alpha}{}' \boldsymbol{L}^*_{X_\alpha})^{-1} \boldsymbol{L}^*_{X_\alpha}{}' \boldsymbol{X}_\alpha \boldsymbol{T}^* \boldsymbol{S}_{T^*}^{-1}. \tag{20}$$

For generic λ_η we have from (5), (17), (18) and (19):

$$\boldsymbol{X}_\eta = \boldsymbol{L}_{X_{\alpha_\eta}} \lfloor \mathbf{a}_\eta, \mathbf{p}_\eta \rfloor + \begin{bmatrix} \boldsymbol{\Lambda}_\eta^o \\ \boldsymbol{T}^* \end{bmatrix} + \boldsymbol{L}^*_{X_{\alpha_\eta}} \boldsymbol{\gamma}_\eta + \boldsymbol{\delta}_\eta \tag{21}$$

$$\gamma_\eta = \mathbf{k}_\eta \boldsymbol{X}_\eta \boldsymbol{Q}_{\overline{Y}_s} \tag{22}$$

where:

$$L^*_{x_\eta} = X_\eta \gamma_\eta ' S^{-1}_{\gamma_\eta}, \tag{23}$$

$$L^*_{x_\eta}[a_\eta, p_\eta] = X_\eta F_\eta^{*\prime} S^{*-1}_{F_\eta} \tag{24}$$

with:

$$F_\eta^{*\prime} = [\Lambda_\eta^{\circ\prime}, T^{x\prime}]$$

Therefore, from (23) we have in (24)

$$[a_\eta, p_\eta]^* = \left(L^*_{x_\eta}{}' L_{x_\eta}\right)^{-1} L^{*\prime}_{x_\eta} X_\eta F_\eta^{*\prime} S^{-1}_{F^*_\eta} \tag{25}$$

and therefore from (4), (22), (24) and (25):

$$\lambda_\eta = [a_\eta, p_\eta]^* F^* + \gamma_\eta \tag{26}$$

By means of the (24), (25), we define (21).
At the end from (7) we can define Δ. Give (α), (β) (γ) we have

$$\Delta = XQ_{T^* \cup \Gamma \cup E^*_s} \tag{27}$$

5. Properties of the model

We can show that:

1) All the proprieties of the latent variables are verified.
We have from (12), (14):

$$\text{cov}\,[T^*, E] = \text{cov}\left(T^*, \overline{Y}_S Q_{T^*}\right) = 0\,; \tag{28}$$

from (12), (18):

$$\text{cov}\left[T^*, \Gamma\right] = \text{cov}\left(\frac{1}{s} L_y{}' S^{-1}_{\overline{Y}_S{}'} \overline{Y}_s, KXQ_{\overline{Y}_s}\right) = 0; \tag{29}$$

from (12) and (27):

$$\text{cov}(T^*, \Delta) = \text{cov}\left[T^*, XQ_{T^* \cup \Gamma \cup E^*_s}\right] = 0\,; \tag{30}$$

from (14) and (22):

$$\text{cov}\left[\Gamma, E^*_s\right] = \text{cov}\left(KXQ_{\overline{Y}_s}, \overline{Y}_s Q_T\right) = 0\,; \tag{31}$$

from (22), (27):

$$\text{cov}\,(\Gamma, \Delta) = \text{cov}\left[\boldsymbol{\Gamma}, \boldsymbol{XQ}_{T^* \cup \Gamma \cup E_s^*}\right] = \boldsymbol{0}\,; \tag{32}$$

from (14), (27):

$$\text{cov}(\boldsymbol{\Delta}, \boldsymbol{E}_s^*) = \text{cov}\left(\boldsymbol{XQ}_{T^* \cup \Gamma \cup E_s^*}, \overline{\boldsymbol{Y}}_s \boldsymbol{E}_s^*\right) = \mathbf{0}\,; \tag{33}$$

from (4), (28):

$$\text{cov}\,(\boldsymbol{\Lambda}, \boldsymbol{\Delta}) = \text{cov}\,[(\boldsymbol{I} - \boldsymbol{A})^{-1} \boldsymbol{PT}^* + (\boldsymbol{I} - \boldsymbol{A})^{-1} \boldsymbol{\Gamma},\ \boldsymbol{XQ}_{T^* \cup \Gamma \cup E_s^*}] = \boldsymbol{0}; \tag{34}$$

from (22):

$$\text{cov}\,(\Gamma, \overline{\boldsymbol{Y}}_s) = \text{cov}(\mathbf{K}\boldsymbol{XQ}_{\overline{Y}_s}, \overline{\boldsymbol{Y}}_S) = \boldsymbol{0}\,; \tag{35}$$

2) The matrices $\mathbf{L}_x$, $\mathbf{L}_y$, $\mathbf{L}_x\mathbf{(I\text{-}A)}^{-1}\mathbf{P}$, $\mathbf{L}_x\mathbf{(I\text{-}A)}^{-1}$ are respectively the OLS regression coefficient matrices of Λ on $\boldsymbol{X}$, $\boldsymbol{T}^*$ on $\overline{\boldsymbol{Y}}_s$, $\boldsymbol{T}$ on $\boldsymbol{X}$, Γ on $\boldsymbol{X}$.
The matrix **P** and the vector $\mathbf{a}_\eta$ are the O.L.S. regression coefficients matrices; $\mathbf{P}$, $\mathbf{a}_\eta$ respectively the direct paths of $\boldsymbol{T}^*$ on Λ and of Λ^o_η on λ_η; $\mathbf{(I\text{-}A)}^{-1}\mathbf{P}$, $\mathbf{(I\text{-}A)}^{-1}$ give respectively the total paths of $\boldsymbol{T}^*$ on Λ and of Γ on Λ.

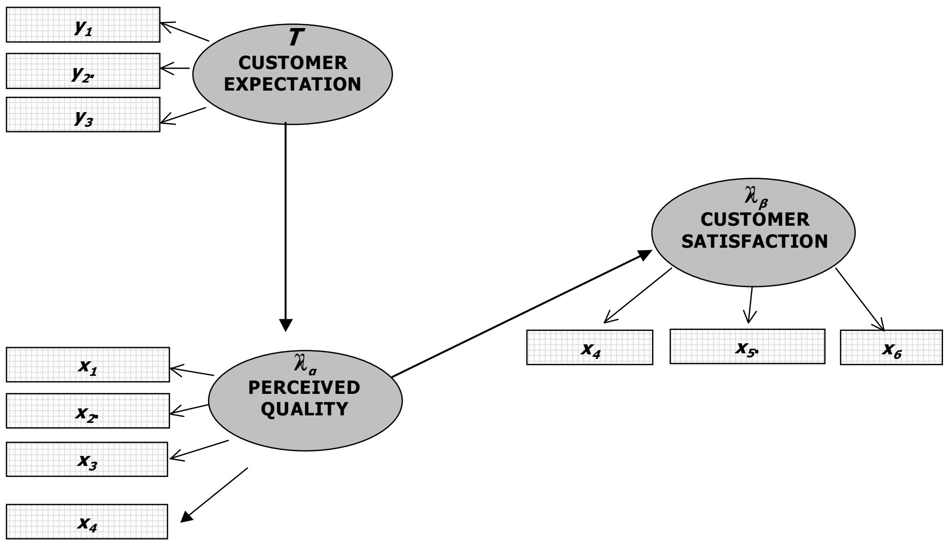
3) The latent variables $\boldsymbol{T}^*$ can be considered causes of the variables $\overline{\boldsymbol{Y}}_s$ by means of the RFCM. The latent variables $\boldsymbol{\Lambda}_\eta$ composed by the linear combination of the latent variables $\boldsymbol{T}^*$ that better fits $\boldsymbol{X}_\eta$ and by γ_η uncorrelated with $\overline{\boldsymbol{Y}}_s$ and $\Lambda_\eta^o = [\lambda_\eta, \ldots, \lambda_{\eta\text{-}1}]$ are causes of $\boldsymbol{X}$. Therefore the measurement models can be considered as two causal models. Also the structural can be considered as causal model because $\boldsymbol{T}^*$ are not linear combination of the Λ.
The solutions allow to obtain the following R.C.D. of the vectorial spaces generated by the observed variables $\boldsymbol{X}$, $\boldsymbol{Y}$

$$\begin{bmatrix} \boldsymbol{X} \\ \overline{\boldsymbol{Y}}_s \end{bmatrix} = \begin{bmatrix} \boldsymbol{X} \\ \overline{\boldsymbol{Y}}_s \end{bmatrix} \boldsymbol{P}_{T^*} + \begin{bmatrix} \boldsymbol{X} \\ \overline{\boldsymbol{Y}}_s \end{bmatrix} \boldsymbol{P}_\Gamma + \begin{bmatrix} \mathbf{0} \\ \boldsymbol{I} \end{bmatrix} \boldsymbol{P}_{E_s^*} + \begin{bmatrix} \boldsymbol{I} \\ \mathbf{0} \end{bmatrix} \boldsymbol{P}_\Delta \tag{36}$$

6. **Example**

Let us consider an example about Customer Satisfaction (figure 2):

Figure 2



y_1 = Expectations of overall quality
y_2 = Expectations of meeting personal requirements (Customization)
y_3 = Expectations for things to go wrong (Reliability)
x_1 = Overall quality
x_2 = Meeting personal requirements
x_3 = Things went wrong
x_4 = Overall satisfaction
x_5 = Confirmation to expectations
x_6 = Distance to ideal product
T = Customer expectation
λ_α = Perceived quality
λ_β = Customer satisfaction.

The latent variable τ is reached by means of RCFM; the errors in equations by means of the Reduced Rank Regression method. Because it is impossible to obtain replicated observations for the observed indicators $\boldsymbol{X}$ we use the method shown in the paragraph 4 for obtaining the latent variable Λ and the errors in equations Γ. The variance-covariance matrices and the regression coefficients matrices are shown in the tables 1, 2.

Table 1

$$\boldsymbol{L_X} = \begin{bmatrix} 1.0266 & 0.2105 \\ 0.9630 & 0.1496 \\ 0.5162 & 0.5506 \\ 0.3730 & 0.8414 \\ 0.3030 & 1.2181 \\ 0.5953 & 0.6851 \end{bmatrix} \quad \boldsymbol{L_Y} = \begin{bmatrix} 6.2035 \\ 0,4918 \\ 0,6000 \end{bmatrix} \quad \boldsymbol{a_\beta} = 0.9458 \quad \mathbf{p} = \begin{bmatrix} 0.9989 \\ -0.6205 \end{bmatrix}$$

Table 2

Correlation Matrix

	T	γ_a	γ_b	δ_{1a}	δ_{2a}	δ_{3a}	δ_{1b}	δ_{2b}	δ_{3b}	ε_1	ε_2	ε_3
T	1	0	0	0	0	0	0	0	0	0	0	0
γ_a	0	1	0	0	0	0	0	0	0	0	0	0
γ_b	0	0	1	0	0	0	0	0	0	0	0	0
δ_{1a}	0	0	0	1,00	-0,03	-0,66	-0,09	-0,03	0,16	0	0	0
δ_{2a}	0	0	0	-0,03	1,00	-0,73	-0,04	0,07	-0,04	0	0	0
δ_{3a}	0	0	0	-0,66	-0,73	1,00	0,09	-0,03	-0,08	0	0	0
δ_{1b}	0	0	0	-0,09	-0,04	0,09	1,00	-0,73	-0,38	0	0	0
δ_{2b}	0	0	0	-0,03	0,07	-0,03	-0,73	1,00	-0,36	0	0	0
δ_{3b}	0	0	0	0,16	-0,04	-0,08	-0,38	-0,36	1,00	0	0	0
ε_1	0	0	0	0	0	0	0	0	0	1,00	-0,43	-0,56
ε_2	0	0	0	0	0	0	0	0	0	-0,43	1,00	-0,51
ε_3	0	0	0	0	0	0	0	0	0	-0,56	-0,51	1,00

References

Bentler P.M. (1982) Linear systems with multiple levels and types of latent variables, in Jöreskog K. G. and Wold H. (Eds.), *Systems under indirect observation: causality, structure, prediction,* North-Holland, Amsterdam, 1, 101-130.

Guttman L. (1955) The determinacy of factor score matrices with implications for five other basic problems of common-factor theory, *The British Journal of Math. Stat. Psych.*, 8(II), pp.65-81.

Haagen K., Oberhofer W. (1992) Common factor model, data analysis technique or what?, *Advances in GLIM and Statistical Modelling,* Eds. L. Fahrmeier, B. Francis, R. Gillchrist, T. Tutz. Springer-Verlag, New York, 151-158

Haagen K., Oberhofer W. (1999) The replicated common factor-analysis model, *Metron*, LVII(3-4), pp.35-45.

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

Haagen K., Vittadini G. (1998) Restricted regression component decomposition, Metron, LVI(1-2), pp.53-75.

Jöreskog K. G. (1973) A general method for estimating a linear structural equation system, in Goldberger A. S. and Duncan O. D. (Eds.), *Structural equation models in the social sciences,* Seminar Press, New York, 85-112.

Jöreskog K. G. (1981) Basic issues in the applications of LISREL, *Data communications computer data analysis,* 1, 1-6

Noonan R., Wold H. (1982) PLS Path modelling with indirectly observed variables, in Joreskog K. G. and Wold H. (Eds.), *System under indirect observation: causality, structure, prediction,* North – Holland, Amsterdam, 2, 75-94.

Schönemann P. H., Steiger J. (1976) Regression Component Analysis, *British Journal of Math. Stat. Psych.*, 29, pp.175-189.

Vittadini G. (1988) On the validity of the indeterminate latent variables in the Lisrel model, *Communications in Statistics - Theory and Methods*, 17(3), pp.861-874.

Vittadini G. (1989) Indeterminacy Problems in the Lisrel Model, *Multivariate Behavioral Research*, 24(4), pp.397-414.

Vittadini G. (1992) Un confronto tra un modello di analisi causale con variabili latenti e metodi alternativi, *Statistica*, LII(3), pp.379-396.

Vittadini, G., Lovaglio P. G. (2001) The estimate of latent variables in a structural model: an alternative approach to PLS in AAVV "*PLS and Related Methods. Proceedings of the PLS' 01 – International Symposium"*, CISIA CERESTA, Montreuil, France pp. 423-434

Wold H. (1982) Soft modelling: the basic design and some extensions, in Jöreskog K., Wold H., eds., *Systems under indirect observation: causality, structure, prediction*, North Holland, vol.II, pp.1-54.