Electronic Journal of Applied Statistical Analysis
EJASA, Electron. J. App. Stat. Anal.
http://siba-ese.unisalento.it/index.php/ejasa/index
e-ISSN: 2070-5948
DOI: 10.1285/i20705948v13n2p413

Properties of the correlation matrix implied by a recursive path model obtained using the Finite Iterative Method
By El Hadri et al.

Published: 14 October 2020

Electronic Journal of Applied Statistical Analysis
Vol. 13, Issue 02, October 2020, 413-435
DOI: 10.1285/i20705948v13n2p413

Properties of the correlation matrix implied by a recursive path model obtained using the Finite Iterative Method

Zouhair El Hadri*[a], M'barek Iaousse[b], Mohamed Hanafi[c], Pasquale Dolce[d], and Yousfi El Kettani[b]

[a] *Center of Mathematical Researches and Applications., Faculty of sciences Mohammed V University In Rabat. Morocco*
[b] *Department of Mathematics, Ibn Tofail University of Kenitra*
[c] *Oniris StatSC Nantes, France*
[d] *Department of Public Health, University of Naples Federico II, Naples, Italy*

Published: 14 October 2020

The present paper announces and demonstrates several properties of the correlation matrix implied by a recursive path model. The main result is that this matrix is affine with respect to the model parameters. These properties concern especially the estimation step in which one should seeks for numerical value for each parameter. Several minimization algorithms are available in softwares for this task, which use iterative procedure that generally requires the computation of successive gradient vectors and the Hessian matrices. The properties established in this paper provide more essence and flexibility within these types of computation.

keywords: Path Analysis, Finite Iterative Method, implied correlation matrix, Unweighted Least Squares function, First & second derivative.

Introduction.

Path analysis, discovered by Wright (1921, 1923, 1934) is a panoply of statistical techniques used to examine cause and effect relations between a set of observed variables

*Corresponding author: z.elhadri@fsr.ac.ma

ISSN: 2070-5948
http://siba-ese.unisalento.it/index.php/ejasa/index

(El Hadri and Hanafi, 2015). On the one hand, Path analysis can be seen as a straightforward extension of multiple regression. On the other hand, it is the special case of structural equation modelling (SEM) since all variables are observed (Bollen, 1989).

Nowadays, Path analysis is applied in many areas such as ecology (Pugesek, 2003; Eisenhauer et al., 2015), social sciences (Duncan, 1966; Hauser, 1975), and psychology (Breckler, 1990).

Path analysis consists of three major elements: (a) the path diagram , (b) decomposing correlations as a function of model parameters, (c) and the separation of the effect of one variable on another into direct, indirect, and total effects.

Figure 1 represents an example of path model with two exogenous variables and three endogenous variables. Furthermore, a path model is analysed upon five steps: specification, identification, estimation, testing, and modification (Schumacker and Lomax, 2004).

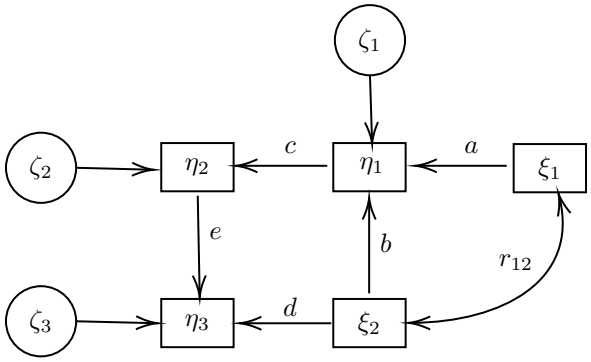

Figure 1: Path analysis model with two exogenous variables and three endogenous variables.

The estimation step is the core of the modelling process. It consists on finding numerical value for each of the model parameters by minimizing a fit function that measures the difference between two matrices: On one side, (i) $\boldsymbol{R}$, the empirical correlation matrix obtained from data, and, on the other side, (ii) $\widehat{\boldsymbol{R}}$, the implied correlation matrix computed as a function of the model parameters. Therefore, the main questions are:
(1) How to compute $\widehat{\boldsymbol{R}}$ when all parameters are known?
(2) How to define a fit function as function of $\widehat{\boldsymbol{R}}$ and $\boldsymbol{R}$?
(3) And finally, how to find the optimal value for each parameter?

To answer the question (1), El Hadri and Hanafi (2015) have presented a method called the Finite Iterative Method (FIM) to calculate $\widehat{\boldsymbol{R}}$ when all parameters are given. Concerning the question (2), several functions are available in the SEM literature, see Bollen (1989). The maximum likelihood F_{ML}, the Generalised Least Squares F_{GLS} and the Unweighted Least Squares function F_{ULS} are widely used. This latter is considered in the present paper and is defined as:

$$F_{ULS}(\boldsymbol{\theta}) = \frac{1}{2} Tr(\widehat{\boldsymbol{R}}(\boldsymbol{\theta}) - \boldsymbol{R})^2 \tag{1}$$

Since this function depends on several variables $\boldsymbol{\theta} = (\theta_1, \theta_2, \dots \theta_T)$ ($\boldsymbol{\theta} = (a, b, c, d, e)$ for the model in Figure 1), then the minimization of F_{ULS} is performed through an optimization algorithm which is the answer to question (3). In this paper, the Newton Raphson algorithm defined as follows is considered.

$$\boldsymbol{\theta^{(s+1)}} = \boldsymbol{\theta^{(s)}} - \left(\boldsymbol{H^{(s)}}\right)^{-1} \boldsymbol{g^{(s)}} \tag{2}$$

where $\boldsymbol{\theta}^{(s)}$ is a vector containing all free parameters in the model at the s^{th} iteration, $\boldsymbol{g}^{(s)}$ is the gradient vector at the s^{th} iteration, and $\boldsymbol{H}^{(s)}$ is the Hessian matrix at the s^{th} iteration. At the initialization step, $\boldsymbol{\theta}^{(1)}$ is chosen arbitrarily , then new vectors $\boldsymbol{\theta}^{(2)}, \boldsymbol{\theta}^{(3)}, \dots$ are successively generated in a way that $F_{ULS}(\boldsymbol{\theta}^{(1)}) \geq F_{ULS}(\boldsymbol{\theta}^{(2)}) \geq F_{ULS}(\boldsymbol{\theta}^{(3)}) \dots$. This process is repeated until convergence is achieved.

Here, elements of $\boldsymbol{g}$ are the first derivatives of F_{ULS} with respect to each parameter :

$$g_i = \frac{\partial F_{ULS}}{\partial \theta_i} = Tr\left[(\widehat{\boldsymbol{R}} - \boldsymbol{R})\frac{\partial \widehat{\boldsymbol{R}}}{\partial \theta_i}\right] \tag{3}$$

And elements of $\boldsymbol{H}$ are the second derivatives of F_{ULS} with respect to each pair of parameters :

$$H_{ij} = \frac{\partial^2 F_{ULS}}{\partial \theta_i \partial \theta_j} = Tr\left[(\widehat{\boldsymbol{R}} - \boldsymbol{R})\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial \theta_i \partial \theta_j} + \frac{\partial \widehat{\boldsymbol{R}}}{\partial \theta_i}\frac{\partial \widehat{\boldsymbol{R}}}{\partial \theta_j}\right] \tag{4}$$

As a consequence, and since $\boldsymbol{R}$ is constant (does not depends on $\boldsymbol{\theta}$), the computation of these derivatives are reduced to the computation of $\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial \theta_i}$ and $\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial \theta_i \partial \theta_j}$ which are respectively the first and second derivation of $\widehat{\boldsymbol{R}}$. Rationally, this requires the explicit expression of $\widehat{\boldsymbol{R}}$ with respect to each parameter θ_i, and with respect to each pair of parameters (θ_i, θ_j). The classical method of estimating a path model uses the finite difference method to calculate the first derivative of $\widehat{\boldsymbol{R}}$ (Jöreskog et al., 1981):

$$\frac{\partial}{\partial \theta_i}\widehat{\boldsymbol{R}}(\boldsymbol{\theta}) = \frac{\widehat{\boldsymbol{R}}(\boldsymbol{\theta} + \boldsymbol{\epsilon}) - \widehat{\boldsymbol{R}}(\boldsymbol{\theta})}{\epsilon_i} \tag{5}$$

Such that $\boldsymbol{\epsilon} = (0, 0, \dots, \epsilon_i, \dots, 0)$ and ϵ_i is a small real number in the i^{th} position. This approximation is due to the fact that $\widehat{\boldsymbol{R}}$ can't be explicitely expresssed in term of θ_i. Besides, it uses an approximation (Jöreskog et al., 1981) in (4) such that:

$$H_{ij} = \frac{\partial^2 F_{ULS}}{\partial \theta_i \partial \theta_j} \simeq Tr\left[\frac{\partial \widehat{\boldsymbol{R}}}{\partial \theta_i}\frac{\partial \widehat{\boldsymbol{R}}}{\partial \theta_j}\right] \tag{6}$$

Contrariwise, since FIM ovecomes these limitations, the aim of the present work is answer the following question: Does this matrix have specific properties that allow us to draw conclusions about the problem of parameter estimation?

The answers to these questions are the main contribution of the present work.

The paper is structured as follows: Section 1 recalls the FIM algorithm for the computation of $\widehat{\boldsymbol{R}}$. Sections 2 introduces basic properties of $\widehat{\boldsymbol{R}}$. In the section 3 we present alternatives useful procedures to compute the derivatives of $\widehat{\boldsymbol{R}}$. Section 4 and 5 illustrate these properties through example, using simulated data and considering real data. In section 6, some conclusions and perspectives are drawn. Finally, all proofs are given in 6 Appendix.

1 Computation of the implied correlation matrix

As mentioned in the introduction, the estimation step consists of finding the parameters that fit the model. In other words, it aims to find values of the parameters that make the data structure as close as possible to the structure of the model. The data structure is characterized by the assumption that the distribution of the observed variables is multivariate and sufficiently well described by its mean and covariance (Jöreskog, 1978). Since the present paper deals with standardized variables, the data structure is defined by its correlation matrix. Consequently, the purpose is to determine the correlation matrix implied by the model.

1.1 Implied correlation matrix

The implied correlation matrix noted by $\widehat{\boldsymbol{R}}$ is computed from the free model parameters. Formally, let:

$$\begin{cases} \eta_1 = \gamma_{11}\xi_1 + \ldots + \gamma_{1p}\xi_p + \zeta_1 \\ \vdots \\ \eta_q = \beta_{q1}\eta_1 + \ldots + \beta_{q,q-1}\eta_{q-1} + \gamma_{q1}\xi_1 + \ldots + \gamma_{qp}\xi_p + \zeta_q \end{cases} \tag{7}$$

be the system of the structural equations for a recursive path model where p and q are respectively the number of exogenous and endogenous variables. And let:

$$\eta = \boldsymbol{B}\eta + \boldsymbol{\Gamma}\xi + \zeta \tag{8}$$

be the compact representation of the system (7). η, ξ and ζ represent respectively the vectors of the endogenous variables, exogenous variables, and disturbances. Assumptions on these vectors are made (see El Hadri and Hanafi (2015)). Then $\widehat{\boldsymbol{R}}$ is defined as :

$$\widehat{\boldsymbol{R}} = \begin{pmatrix} \mathbb{E}(\xi\xi^t) & \mathbb{E}(\xi\eta^t) \\ \mathbb{E}(\eta\xi^t) & \mathbb{E}(\eta\eta^t) \end{pmatrix} \tag{9}$$

This matrix can be determined in a compact way by writing the vector of endogenous variables as a function of the vector of exogenous ones (Jöreskog, 1970, 1978; Jöreskog and Wold, 1982; Jöreskog and Sorbom, 1993).

It can also be constructed using the Wright's rules (Wright, 1921, 1923, 1934, 1960), by calculating every correlation between each pair of model variables. El Hadri and Hanafi

(2015, 2016) proposed FIM, which uses these rules to build $\widehat{\boldsymbol{R}}$ iteratively. Morover, FIM was further discussed by El Hadri et al. (2019) and Iaousse et al. (2020a,b). In what follows, the FIM Algorithm is presented.

1.2 Finite Iterative Method

El Hadri and Hanafi (2015) have demonstrated that under the condition of a recursive model and standardized variables, $\widehat{\boldsymbol{R}}$ can be constructed using the algorithm 1 called FIM Algorithm. To make it clear, if we note by $\boldsymbol{A} = (\boldsymbol{\Gamma}, \boldsymbol{B})$ the matrix of all model parameters then (8) becomes:

$$\eta = \boldsymbol{A} \begin{pmatrix} \xi \\ \eta \end{pmatrix} + \zeta \tag{10}$$

In addition, we note by $\boldsymbol{\Phi} = E(\xi\xi^{'})$ the correlation matrix of the exogenous variables. Finally, for a given matrix X: $X_{r:t,l:s}$ denote the submatrix of X containing rows from r to t and columns from l to s (i.e $X_{r:t,l:s} = (X_{ij})_{(r<i<t,l<j<s)}$). Algorithm 1 shows the steps to construct the implied correlation matrix $\widehat{\boldsymbol{R}}$ (El Hadri and Hanafi (2015)) and Figure 2 is an illustration.

Initialization: $\widehat{\boldsymbol{R}}_{\boldsymbol{1:p,1:p}} = \boldsymbol{\Phi}$;
Repeat for $j = 1, ...q$;
1. $\widehat{\boldsymbol{R}}_{\boldsymbol{p+j,1:p+j-1}} = \boldsymbol{A}_{\boldsymbol{j,1:p+j-1}} \widehat{\boldsymbol{R}}_{\boldsymbol{1:p+j-1,1:p+j-1}}$;
2. $\widehat{\boldsymbol{R}}_{\boldsymbol{1:p+j-1,p+j}} = (\widehat{\boldsymbol{R}}_{\boldsymbol{p+j,1:p+j-1}})^t$;
3. $\widehat{\boldsymbol{R}}_{\boldsymbol{p+j,p+j}} = 1$

Algorithm 1: FIM algorithm.

For instance, the correlation matrix implied by the model in Figure 1 using Algorithm 1 is:

$$\widehat{\boldsymbol{R}} = \begin{pmatrix} 1 & r_{12} & a + br_{12} & ac + bcr_{12} & dr_{12} + ace + bcer_{12} \\ r_{12} & 1 & ar_{12} + b & acr_{12} + bc & d + acer_{12} + bce \\ a + br_{12} & ar_{12} + b & 1 & c & adr_{12} + bd + ce \\ ac + bcr_{12} & acr_{12} + bc & c & 1 & acdr_{12} + bcd + e \\ dr_{12} + ace + bcer_{12} & d + acer_{12} + bce & adr_{12} + bd + ce & acdr_{12} + bcd + e & 1 \end{pmatrix}$$

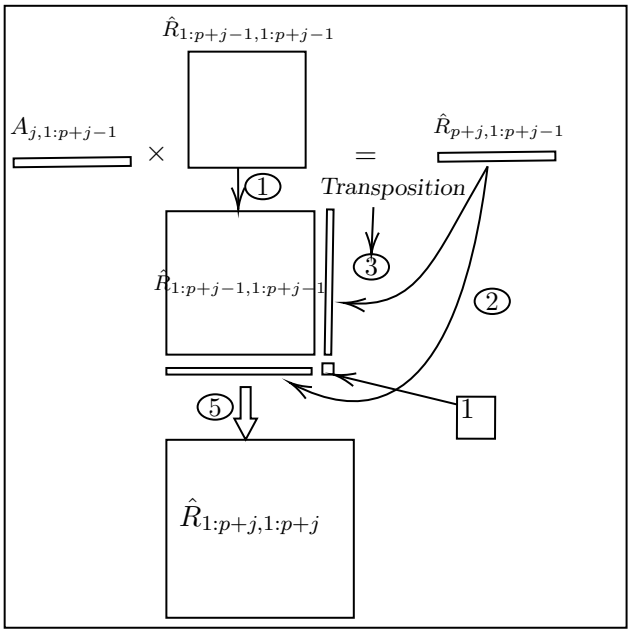

Figure 2: Illustration of Algorithm 1

2 Basic properties of the implied correlation matrix

As aforementioned, the implied correlation matrix $\widehat{\boldsymbol{R}}$ obtained using FIM disposes of some useful properties. Recall that $\widehat{\boldsymbol{R}}$ is function of all parameters. In the present paper, the model parameters are the elements of matrix $\boldsymbol{A}$.

In what follows, we suppose that we are dealing with a recursive path model with equations given in (7) or, equivalently in compact form (8 or 10). In addition, x and y are two parameters of the model such that:

$$x = \boldsymbol{A}_{jk} \text{ such that } (j,k) \in \{1:q\} \times \{1:p+j-1\} \tag{11}$$

and

$$y = \boldsymbol{A}_{st} \text{ such that } (s,t) \in \{1:q\} \times \{1:p+s-1\} \tag{12}$$

The reader may notice that the dimension of the matrix $\boldsymbol{A}$ is $q \times (p+q)$. However, the column indix k of x (respectivelly t for y) is at most $p+j-1$ (respectvelly $p+s-1$) such that the j (respectivelly s) is the row index. This is because the matrix $\boldsymbol{B}$ is strictilly lower triangular (since the model is recursive). Thus any parameter of the form $\boldsymbol{A}_{jk}$ such that $k \geq p+j$ vanishes. In addition, since x and y play a symmetric role, we consider that $s \geq j$.

Lemma 1. *The bloc $\widehat{\boldsymbol{R}}_{1:p+j-1,1:p+j-1}$ does not depend on x.*

Proof. See 6 Appendix. □

Lemma 2. *The bloc $\widehat{\boldsymbol{R}}_{1:p+j,1:p+j}$ is affine on x.*

Proof. See 6 Appendix. □

Theorem 1. *The correlation matrix implied by the model* $\widehat{\boldsymbol{R}}$ *is affine with respect to* x.

Proof. See 6 Appendix. □

A direct consequence of theorem 1 is that $\widehat{\boldsymbol{R}}$ can be expressed as :

$$\widehat{\boldsymbol{R}} = \left(\frac{b\widehat{\boldsymbol{R}}(a) - a\widehat{\boldsymbol{R}}(b)}{b-a} \right) + x \left(\frac{\widehat{\boldsymbol{R}}(b) - \widehat{\boldsymbol{R}}(a)}{b-a} \right) \tag{13}$$

such that a and b are arbitrary distinct values for x and $\widehat{\boldsymbol{R}}(\alpha) = \widehat{\boldsymbol{R}}(x = \alpha)$ is the correlation matrix implied by the model where x is replaced by α. However, setting $a = 0$ and $b = 1$, (13) is reduced to:

$$\widehat{\boldsymbol{R}} = \widehat{\boldsymbol{R}}(0) + x \left(\widehat{\boldsymbol{R}}(1) - \widehat{\boldsymbol{R}}(0) \right) \tag{14}$$

Moreover, we suppose that another parameter is unknown, say y. Then, since $\widehat{\boldsymbol{R}}(0)$ and $\widehat{\boldsymbol{R}}(1)$ are implied correlation matrices, we can use (14) to express them in term of y.
Indeed, $\widehat{\boldsymbol{R}}$ can expressed in term of x and y simultaneously as:

$$\begin{aligned} \widehat{\boldsymbol{R}} &= \widehat{\boldsymbol{R}}(0,0) + x \left[\widehat{\boldsymbol{R}}(1,0) - \widehat{\boldsymbol{R}}(0,0) \right] + y \left[\widehat{\boldsymbol{R}}(0,1) - \widehat{\boldsymbol{R}}(0,0) \right] \\ &\quad + xy \left[\widehat{\boldsymbol{R}}(1,1) - \widehat{\boldsymbol{R}}(1,0) - \widehat{\boldsymbol{R}}(0,1) + \widehat{\boldsymbol{R}}(0,0) \right] \end{aligned} \tag{15}$$

such that $\widehat{\boldsymbol{R}}(\alpha, \beta) = \widehat{\boldsymbol{R}}(x = \alpha, y = \beta)$ is the correlation matrix implied by the model where x and y are respectivelly fixed to α and β.

Indeed, using (14) it holds:

$$\frac{\partial \widehat{\boldsymbol{R}}}{\partial x} = \widehat{\boldsymbol{R}}(1) - \widehat{\boldsymbol{R}}(0) \tag{16}$$

And

$$\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x^2} = 0 \tag{17}$$

And using (15) it holds:

$$\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y} = \widehat{\boldsymbol{R}}(1,1) - \widehat{\boldsymbol{R}}(1,0) - \widehat{\boldsymbol{R}}(0,1) + \widehat{\boldsymbol{R}}(0,0) \tag{18}$$

3 Alternative methods for computing the fisrt and the second derivative of the implied correlation matrix

In this section, we propose alternative strategies to compute $\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial x}$ and $\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}$. We begin by the algorithm 2 below allowing the computation of the first derivative. Let $\boldsymbol{G}$ be a $(p+q)\times(p+q)$ matrix.

$\star$ If $j \neq q$ then :

1. $\boldsymbol{G}_{1:p+j-1,1:p+j-1} = 0$
2. $\boldsymbol{G}_{p+j,1:p+j-1} = \widehat{\boldsymbol{R}}_{k,1:p+j-1}$
3. $\boldsymbol{G}_{1:p+j-1,p+j} = (\boldsymbol{G}_{p+j,1:p+j-1})^t$
4. $\boldsymbol{G}_{p+j,p+j} = 0$
5. for $j^{'} \in \{j+1:q\}$
   a) $\boldsymbol{G}_{p+j^{'},1:p+j^{'}-1} = \boldsymbol{A}_{j^{'},1:p+j^{'}-1}\boldsymbol{G}_{1:p+j^{'}-1,1:p+j^{'}-1}$
   b) $\boldsymbol{G}_{1:p+j^{'}-1,p+j^{'}} = \boldsymbol{G}^t$
   c) $\boldsymbol{G}_{p+j^{'},p+j^{'}} = 0$

$\star$ If $j = q$ then the building stops at step 4.

Algorithm 2: First derivative algorithm.

Theorem 2. *The first derivative can be computed by algorithm 2 above:* $\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial x} = \boldsymbol{G}$.

Proof. See 6 Appendix. □

Remark 1. *We can note that algorithm 2 is an adaptation of Algorithm 1 in the sense that:*

1. *The iterations begin at* $p+j+1$ *instead of* $p+1$,
2. *the iterations are initialized by* $\boldsymbol{G}_{1:p+j,1:p+j}$ *instead of* $\boldsymbol{\Phi}$,
3. *diagonal elements are fixed to* 0 *instead of* 1.

Next, using algorithm 3 below, we can compute iteratively the second derivative with respect to two distinct parameters x and y. Let $\boldsymbol{H}$ be a $(p+q)\times(p+q)$ matrix.

⋆ If $s \neq q$:

1. $\boldsymbol{H}_{1:p+s-1,1:p+s-1} = 0$
2. $\boldsymbol{H}_{p+s,1:p+s-1} = \dfrac{\partial}{\partial x}\widehat{\boldsymbol{R}}_{t,1:p+s-1}$
3. $\boldsymbol{H}_{1:p+s-1,p+s} = (\boldsymbol{H}_{p+s,1:p+s-1})^t$
4. $\boldsymbol{H}_{p+s,p+s} = 0$
5. for $s^{'} \in \{s+1 : q\}$
   a) $\boldsymbol{H}_{p+s^{'},1:p+s^{'}-1} = \boldsymbol{A}_{s^{'},1:p+s^{'}-1}\boldsymbol{H}_{1:p+s^{'}-1,1:p+s^{'}-1}$
   b) $\boldsymbol{H}_{1:p+s^{'}-1,p+s^{'}} = (\boldsymbol{H}_{p+s^{'},1:p+s^{'}-1})^t$
   c) $\boldsymbol{H}_{p+s^{'},p+s^{'}} = 0$

⋆ If $s = q$ then the building stops at step 4.

Algorithm 3: Second derivative algorithm.

Theorem 3. *The second derivative can be computed by algorithm 3 above:* $\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y} = \boldsymbol{H}$.

Proof. See 6 Appendix. □

Remark 2. *Once again, algorithm 3 is an adaptation of Algorithm 1 in the sense that:*

1. *The iterations begin at* $p+s+1$ *instead of* $p+1$,
2. *the iterations are initialized by* $\boldsymbol{H}_{1:p+s,1:p+s}$ *instead of* $\boldsymbol{\Phi}$,
3. *diagonal elements are fixed to* 0 *instead of* 1.

Corollary 1. *If* x *and* y *are associated with the same endogenous variable then* $\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y} = 0$.

Proof. See 6 Appendix. □

4 An example of a recursive path model for illustration

In this section, we consider the model in Figure 1 (see Introduction section). The system of structural equation of this path model is:

$$\begin{cases} \eta_1 = a\xi_1 + b\xi_2 + \zeta_1 \\ \eta_2 = c\eta_1 + \zeta_2 \\ \eta_3 = d\xi_2 + e\eta_2 + \zeta_3 \end{cases} \tag{19}$$

Or, in the compact form:

$$\boldsymbol{\eta} = \begin{pmatrix} 0 & 0 & 0 \\ c & 0 & 0 \\ 0 & e & 0 \end{pmatrix} \boldsymbol{\eta} + \begin{pmatrix} a & b \\ 0 & 0 \\ 0 & d \end{pmatrix} \boldsymbol{\xi} + \boldsymbol{\zeta} \tag{20}$$

such that, $\boldsymbol{\eta} = (\eta_1, \eta_2, \eta_3)^t$, $\boldsymbol{\xi} = (\xi_1, \xi_2)^t$, and $\boldsymbol{\zeta} = (\zeta_1, \zeta_2, \zeta_3)^t$. Thus, the matrix of parameters is

$$\boldsymbol{A} = \begin{pmatrix} a & b & 0 & 0 & 0 \\ 0 & 0 & c & 0 & 0 \\ 0 & d & 0 & e & 0 \end{pmatrix} \tag{21}$$

In addition, the correlation matrix among exogenous variables is

$$\boldsymbol{\Phi} = \begin{pmatrix} 1 & r_{12} \\ r_{12} & 1 \end{pmatrix} \tag{22}$$

where r_{12} is the empirical correlation between the two exogenous variables ξ_1 and ξ_2.

Next, we consider the following values for parameters: $a = 0.45, b = 0.32, c = -0.10, d = 0.72, e = -0.92$ and $r_{12} = 0.60$. In what follows, we propose to apply properties announced in the previous section to compute the following derivatives: $\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial a}$, $\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial a \partial c}$, and $\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial d \partial e}$.

1. Computation of $\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial a}$:

   a) Using Algorithm 1, it holds $\widehat{\boldsymbol{R}}_{1:2,1:2} = \boldsymbol{\Phi} = \begin{pmatrix} 1 & 0.6 \\ 0.6 & 1 \end{pmatrix}$,

   b) Initialization: $\left(\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{1:2,1:2} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$,

   $\star$ $\left(\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{3,1:2} = \widehat{\boldsymbol{R}}_{1,1:2} = \begin{pmatrix} 1 & 0.6 \end{pmatrix}$

   $\star$ $\left(\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{1:2,3} = \begin{pmatrix} 1 \\ 0.6 \end{pmatrix}$

   $\star$ $\left(\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{3,3} = 0$

   Thus, $\left(\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{1:3,1:3} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0.6 \\ 1 & 0.6 & 0 \end{pmatrix}$.

c) Iteration 1:

$$\star \left(\frac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{4,1:3} = \boldsymbol{A}_{2,1:3}\left(\frac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{1:3,1:3} = \begin{pmatrix} 0 & 0 & -0.1 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0.6 \\ 1 & 0.6 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -0.1 & -0.06 & 0 \end{pmatrix}$$

$$\star \left(\frac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{1:3,4} = \begin{pmatrix} -0.1 \\ -0.06 \\ 0 \end{pmatrix} \text{ and } \left(\frac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{4,4} = 0$$

$$\text{Thus } \left(\frac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{1:4,1:4} = \begin{pmatrix} 0 & 0 & 1 & -0.1 \\ 0 & 0 & 0.6 & -0.06 \\ 1 & 0.6 & 0 & 0 \\ -0.1 & -0.06 & 0 & 0 \end{pmatrix}.$$

d) Iteration 2:

$$\star \left(\frac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{5,1:4} = \boldsymbol{A}_{3,1:4}\left(\frac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{1:4,1:4}$$

$$= \begin{pmatrix} 0 & 0.72 & 0 & -0.92 \end{pmatrix}\begin{pmatrix} 0 & 0 & 1 & -0.1 \\ 0 & 0 & 0.6 & -0.06 \\ 1 & 0.6 & 0 & 0 \\ -0.1 & -0.06 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 0.092 & 0.0552 & 0.432 & -0.0432 \end{pmatrix}$$

$$\star \left(\frac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{1:4,5} = \begin{pmatrix} 0.092 \\ 0.0552 \\ 0.432 \\ -0.0432 \end{pmatrix}$$

$$\star \left(\frac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{5,5} = 0$$

$$\text{Thus } \frac{\partial \widehat{\boldsymbol{R}}}{\partial a} = \begin{pmatrix} 0 & 0 & 1 & -0.1 & 0.092 \\ 0 & 0 & 0.6 & -0.06 & 0.0552 \\ 1 & 0.6 & 0 & 0 & 0.432 \\ -0.1 & -0.06 & 0 & 0 & -0.0432 \\ 0.092 & 0.0552 & 0.432 & -0.0432 & 0 \end{pmatrix}.$$

2. Computation of $\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial a \partial c}$:

a) Using Algorithm 2 (or the previous example), it holds $\left(\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{1:3,1:3} = \begin{pmatrix} 0 & 0.0 & 1.0 \\ 0 & 0.0 & 0.6 \\ 1 & 0.6 & 0.0 \end{pmatrix}$,

b) Initialization: $\left(\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial a \partial c}\right)_{1:3,1:3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$,

$\star \quad \left(\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial a \partial c}\right)_{4,1:3} = \left(\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial a}\right)_{3,1:3} = \begin{pmatrix} 1 & 0.6 & 0 \end{pmatrix}$

$\star \quad \left(\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial a \partial c}\right)_{1:3,4} = \begin{pmatrix} 1 \\ 0.6 \\ 0 \end{pmatrix}$

$\star \quad \left(\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial a \partial c}\right)_{4,4} = 0$

Thus, $\left(\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial a \partial c}\right)_{1:4,1:4} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0.6 \\ 0 & 0 & 0 & 0 \\ 1 & 0.6 & 0 & 0 \end{pmatrix}$.

c) Iteration 1:

$$\star \quad \left(\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial a \partial c}\right)_{5,1:4} = \boldsymbol{A}_{3,1:4} \left(\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial a \partial c}\right)_{1:4,1:4}$$

$$= \begin{pmatrix} 0 & 0.72 & 0 & -0.92 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0.6 \\ 0 & 0 & 0 & 0 \\ 1 & 0.6 & 0 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} -0.92 & -0.552 & 0 & 0.432 \end{pmatrix}$$

$$\star \quad \left(\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial a \partial c}\right)_{1:4,5} = \begin{pmatrix} -0.92 \\ -0.552 \\ 0 \\ 0.432 \end{pmatrix}$$

$$\star \quad \left(\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial a \partial c}\right)_{5,5} = 0$$

$$\text{Thus, } \frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial a \partial c} = \begin{pmatrix} 0 & 0 & 0 & 1 & -0.92 \\ 0 & 0 & 0 & 0.6 & -0.552 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0.6 & 0 & 0 & 0.432 \\ -0.92 & -0.552 & 0 & 0.432 & 0 \end{pmatrix}.$$

3. Computation of $\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial d \partial e}$:
Since d and e are associated with the same endogenous variable (see (19) then according to the corollary 1, it holds: $\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial d \partial e} = \mathbf{0}$ where $\mathbf{0}$ is the null matrix of order (5×5).

5 Simulation Study and Empirical Example

In the present section, we propose to compare the results of the present paper to some classical methods. This comparison shall concern two aspects : a) the gradient vectors and the Hessian matrices and b) the parameters estimates of a known model using real data.

5.1 Comparison of the gradient vector and Hessian matrix

We consider the model described in figure 1 and the following simulation: 100 vectors of model parameters are generated. For each vector, the gradient vector given in (3)

and the Hessian matrix given in (4) are computed using two approaches. (i)Classical Approach (CA): The package *NumDeriv* In R software and (ii) Proposal Approach (PA): Theorem 2 and Theorem 3 of the present paper. To compare the two approaches, we compute the following quantities: $\Delta_g = Norm(g_2 - g_1)$ and $\Delta_H = Norm(H_2 - H_1)$ where g_1 (respectively H_1) is the gradient vector (respectively the Hessian matrix) computed using CA and g_2 (respectively H_2) is the gradient vector (respectively the Hessian matrix) computed using PA.

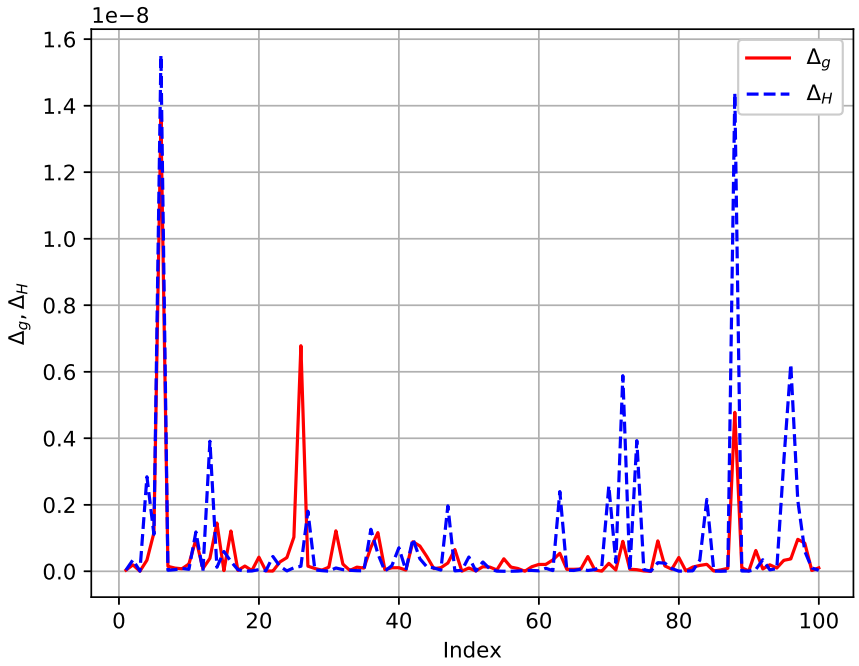

Figure 3: The difference between the gradient vectors (Δ_g) and the Hessian matrices (Δ_H) for 100 simulated vectors of parameters

Values of Δ_g and Δ_H are represented in Figure 3. We can observe that these values do not exceed 1.4×10^{-8} in the case of the gradient vectors and 1.6×10^{-8} for the Hessian matrices. In other words, we can affirm that PA and CA are identical.

5.2 Comparison of estimated parameters

Here, we consider the Union Sentiment Model described in figure 4 (available in $MIIVsem$ package in R). According to Bollen (1989), the union sentiment data come from a study of union sentiment among southern non-union textile workers. It contains 175 observations and 5 variables which are: years in the textile mill, age, deference to managers, support for labour activism, and sentiment toward unions. The causal ordering adopted by Bollen (Bollen, 1989) specifies that age (age) influences deference

($deferenc$) and attitude toward activism ($laboract$), while seniority ($yrsmill$) affects only union sentiment ($unionsen$). The causal ordering among endogenous variables specifies that deference ($deferenc$) causes attitudes toward activism ($laborct$) and unions ($unionsen$); and that activism influences union sentiment.

The parameters of this model are estimated using two methods. (i) the *lavann* package in R software; and (ii) the Newton Raphson algorithm described in (2) using Theorem 2 and Theorem 3 of the present paper. The results are shown in Table 1. This table shows that the estimates are either equal or very close.

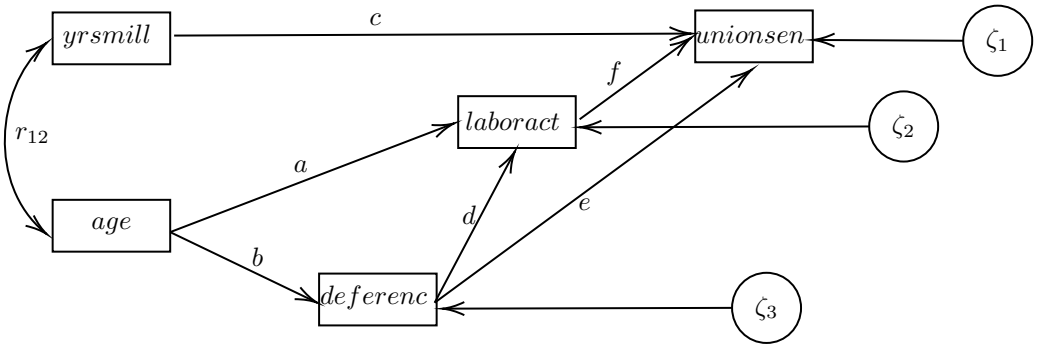

Figure 4: The path diagram for the Union Sentiment Model

Parameter	(i)	(ii)
a	-0.323	-0.323
b	0.278	0.279
c	0.166	0.166
d	-0.332	-0.321
e	-0.141	-0.142
f	0.508	0.507

Table 1: Estimates of Union sentiment paramters's model

6 Conclusion and perspectives

The estimation stage is a crucial step in modelling. Consequently, it is better to make the steps of the computation more accurate and avoid as much as possible approximations. In this paper, we tried to achieve this objective by demonstrating some properties of the implied correlation matrix, which make the computation of the derivatives of the fit function precise. We believe that these properties will lead to new perspective for model estimation. It should be noted that the proposed properties are applicable only to path models with standard assumption. The question that still remain is how could behave the correlation matrix implied by path model with some violated hypothesis (e.g.

a model with correlated disturbances)?

In addition, it should be mentioned that the present paper treats the estimation stage. However, the classical indices of fit can be used to measure how a path model can produce given data. In this regard, further work will focus on how the properties found in the present can be exploited to propose new model fit measures. Moreover, further research will also focus on the exploitation of the results found in this work to the composited-based approach to SEM, as the Partial Least Squares Path Modeling (PLSPM) and its more recent developments (Dolce et al., 2018). In particular, the FIM algorithm could be used as a module (i.e., a sub-iterative algorithm) within the PLSPM algorithms, to obtain new proporties for the PLSPM solutions.

Acknowledgement

The authors are grateful to the editorial team and the anonymous reviewers for their valuable remarks and suggestions.

Appendix

Proof of lemma 1

Proof. We know from (11) that $x = \boldsymbol{A}_{jk}$ *such that* $(j,k) \in \{1:q\} \times \{1:p+j-1\}$.

1. If $j = 1$ then $\widehat{\boldsymbol{R}}_{1:p,1:p} = \boldsymbol{\Phi}$ is constant, i.e the bloc $\widehat{\boldsymbol{R}}_{1:p,1:p}$ does not depend on x.
2. If $j > 1$. We know that x does not appears in all equations of (7) from 1 to $j-1$, thus the rows of $\boldsymbol{A}$ from 1 to $j-1$ do not depends on x. As a consequence,
   - (i) the 1^{st} sub-row $\boldsymbol{A}_{1,1:p}$ of $\boldsymbol{A}$ is constant, and (ii) the bloc $\widehat{\boldsymbol{R}}_{1:p,1:p}$ is constant, thus, the sub-row $\widehat{\boldsymbol{R}}_{p+1,1:p} = \boldsymbol{A}_{1,1:p}\widehat{\boldsymbol{R}}_{1:p,1:p}$ is constant. And by transposition, the sub-column $\widehat{\boldsymbol{R}}_{1:p,p+1}$ is constant. And since $\widehat{\boldsymbol{R}}_{p+1,p+1} = 1$ is constant then the bloc $\widehat{\boldsymbol{R}}_{1:p+1,1:p+1}$ is constant.
   - (i) the 2^{nd} sub-row $\boldsymbol{A}_{2,1:p+1}$ of $\boldsymbol{A}$ is constant, and (ii) from the previous iteration, the bloc $\widehat{\boldsymbol{R}}_{1:p+1,1:p+1}$ is constant, thus, the sub-row $\widehat{\boldsymbol{R}}_{p+2,1:p+1} = \boldsymbol{A}_{2,1:p+1}\widehat{\boldsymbol{R}}_{1:p+1,1:p+1}$ is constant. And by transposition, the sub-column $\widehat{\boldsymbol{R}}_{1:p+1,p+2}$ is constant. And since $\widehat{\boldsymbol{R}}_{p+2,p+2} = 1$ is constant then the bloc $\widehat{\boldsymbol{R}}_{1:p+2,1:p+2}$ is constant.
   - we continue in the same way. (i) the $(j-1)^{th}$ sub-row $\boldsymbol{A}_{j-1,1:p+j-2}$ of $\boldsymbol{A}$ is constant, and (ii) from the previous iteration, the bloc $\widehat{\boldsymbol{R}}_{1:p+j-2,1:p+j-2}$ is constant, thus, the sub-row $\widehat{\boldsymbol{R}}_{p+j-1,1:p+j-2} = \boldsymbol{A}_{j-1,1:p+j-2}\widehat{\boldsymbol{R}}_{1:p+j-2,1:p+j-2}$ is constant. And by transposition, the sub-column $\widehat{\boldsymbol{R}}_{1:p+j-2,p+j-1}$ is constant. And since $\widehat{\boldsymbol{R}}_{p+j-1,p+j-1} = 1$ is constant then the bloc $\widehat{\boldsymbol{R}}_{1:p+j-1,1:p+j-1}$ is constant. See Figure 5.

□

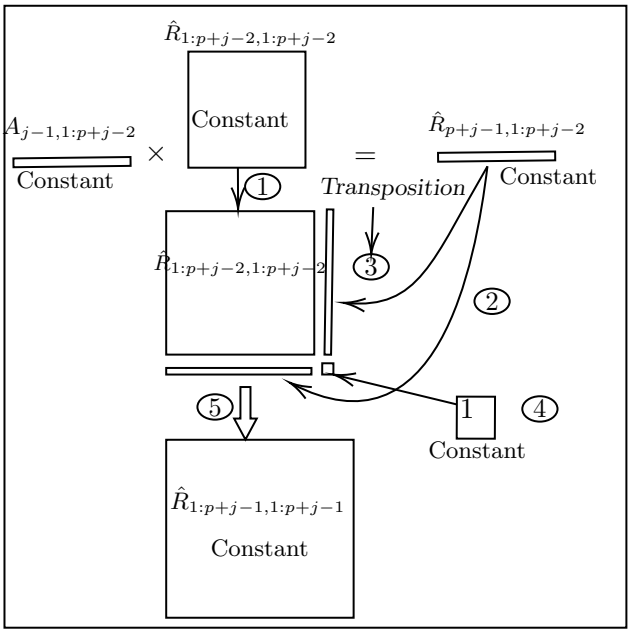

Figure 5: Illustration of lemma 1

Proof of Lemma 2

Proof. We know that (i) the j^{th} sub-row $\boldsymbol{A}_{j,1:p+j-1}$ of $\boldsymbol{A}$ is affine on x, and (ii) from lemma 1, the bloc $\widehat{\boldsymbol{R}}_{1:p+j-1,1:p+j-1}$ is constant, thus, the sub-row $\widehat{\boldsymbol{R}}_{p+j,1:p+j-1} = \boldsymbol{A}_{j,1:p+j-1}\widehat{\boldsymbol{R}}_{1:p+j-1,1:p+j-1}$ is affine on x. And by transposition, the sub-column $\widehat{\boldsymbol{R}}_{1:p+j-1,p+j}$ is affine. And since $\widehat{\boldsymbol{R}}_{p+j,p+j} = 1$ is constant then the bloc $\widehat{\boldsymbol{R}}_{1:p+j,1:p+j}$ is affine. See Figure 6.

□

Proof of Theorem 1

Proof.

- (i) the $(j+1)^{th}$ sub-row $\boldsymbol{A}_{j+1,1:p+j}$ of $\boldsymbol{A}$ is constant, and (ii) from lemma 2, the bloc $\widehat{\boldsymbol{R}}_{1:p+j,1:p+j}$ is affine, thus, the sub-row $\widehat{\boldsymbol{R}}_{p+j+1,1:p+j} = \boldsymbol{A}_{j+1,1:p+j}\widehat{\boldsymbol{R}}_{1:p+j,1:p+j}$ is affine. And by transposition, the sub-column $\widehat{\boldsymbol{R}}_{1:p+j,p+j+1}$ is affine. And since $\widehat{\boldsymbol{R}}_{p+j+1,p+j+1} = 1$ is constant then the bloc $\widehat{\boldsymbol{R}}_{1:p+j+1,1:p+j+1}$ is affine.
- (i) the $(j+2)^{th}$ sub-row $\boldsymbol{A}_{j+2,1:p+j+1}$ of $\boldsymbol{A}$ is constant, and (ii) from the previous iteration, the bloc $\widehat{\boldsymbol{R}}_{1:p+j+1,1:p+j+1}$ is affine, thus, the sub-row :

  $$\widehat{\boldsymbol{R}}_{p+j+2,1:p+j+1} = \boldsymbol{A}_{j+2,1:p+j+1}\widehat{\boldsymbol{R}}_{1:p+j+1,1:p+j+1}$$

  is affine. And by transposition, the sub-column $\widehat{\boldsymbol{R}}_{1:p+j+1,p+j+2}$ is affine. And since $\widehat{\boldsymbol{R}}_{p+j+2,p+j+2} = 1$ is constant then the bloc $\widehat{\boldsymbol{R}}_{1:p+j+2,1:p+j+2}$ is affine.

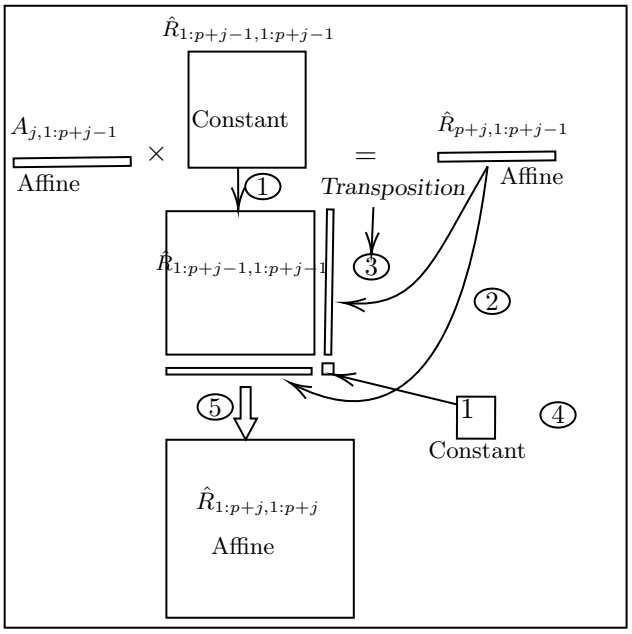

Figure 6: Illustration of lemma 2

- we continue in the same way. (i) the q^{th} sub-row $\boldsymbol{A}_{q,1:p+q-1}$ of $\boldsymbol{A}$ is constant, and (ii) from the previous iteration, the bloc $\widehat{\boldsymbol{R}}_{1:p+q-1,1:p+q-1}$ is affine, thus, the sub-row $\widehat{\boldsymbol{R}}_{p+q,1:p+q-1} = \boldsymbol{A}_{q,1:p+q-1}\widehat{\boldsymbol{R}}_{1:p+q-1,1:p+q-1}$ is affine on x. And by transposition, the sub-column $\widehat{\boldsymbol{R}}_{1:p+q-1,p+q}$ is affine. And since $\widehat{\boldsymbol{R}}_{p+q,p+q} = 1$ is constant then the bloc $\widehat{\boldsymbol{R}} = \widehat{\boldsymbol{R}}_{1:p+q,1:p+q}$ is affine. See Figure 7.

□

Proof of Theorem 2

Proof.

1. Using lemma 1, it comes $\left(\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial x}\right)_{1:p+j-1,1:p+j-1} = 0.$
2. $\left(\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial x}\right)_{p+j,1:p+j-1} = \left(\dfrac{\partial}{\partial x}\boldsymbol{A}_{j,1:p+j-1}\right)\widehat{\boldsymbol{R}}_{1:p+j-1,1:p+j-1} = (0,\dots 0,1,0,\dots,0)\widehat{\boldsymbol{R}}_{1:p+j-1,1:p+j-1}$ where 1 is in the k^{th} position. As a consequence, $\dfrac{\partial}{\partial x}\widehat{\boldsymbol{R}}_{p+j,1:p+j-1} = \widehat{\boldsymbol{R}}_{k,1:p+j-1}$.
3. Flows from the symmetry of $\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial x}$.
4. $\left(\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial x}\right)_{p+j,p+j} = \dfrac{\partial}{\partial x}(1) = 0.$ See Figure 8 (a).

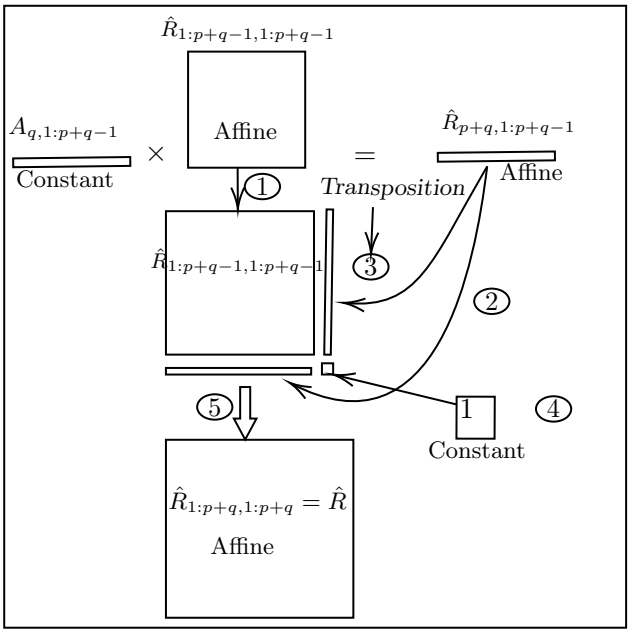

Figure 7: Illustration of theorem 1

5. a) $\left(\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial x}\right)_{p+j',1:p+j'-1} = \boldsymbol{A}_{j',1:p+j'-1}\left(\dfrac{\partial}{\partial x}\widehat{\boldsymbol{R}}_{1:p+j'-1,1:p+j'-1}\right)$.

   b) Flows from the symmetry of $\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial x}$.

   c) $\left(\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial x}\right)_{1:p+j'-1,1:p+j'-1} = \dfrac{\partial}{\partial x}(1) = 0$. See Figure 8 (b).

□

Proof of Theorem 3

Proof. 1. Using lemma 2, the bloc $\widehat{\boldsymbol{R}}_{1:p+s-1,1:p+s-1}$ is constant on y.

Thus $\left(\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}\right)_{1:p+s-1,1:p+s-1} = 0.$

2. Using theorem 2, $\left(\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}\right)_{p+s,1:p+s-1} = \left(\dfrac{\partial}{\partial x}\left(\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial y}\right)\right)_{p+s,1:p+s-1} = \left(\dfrac{\partial \widehat{\boldsymbol{R}}}{\partial x}\right)_{t,1:p+s-1}$

3. Flows from the symmetry of $\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}$

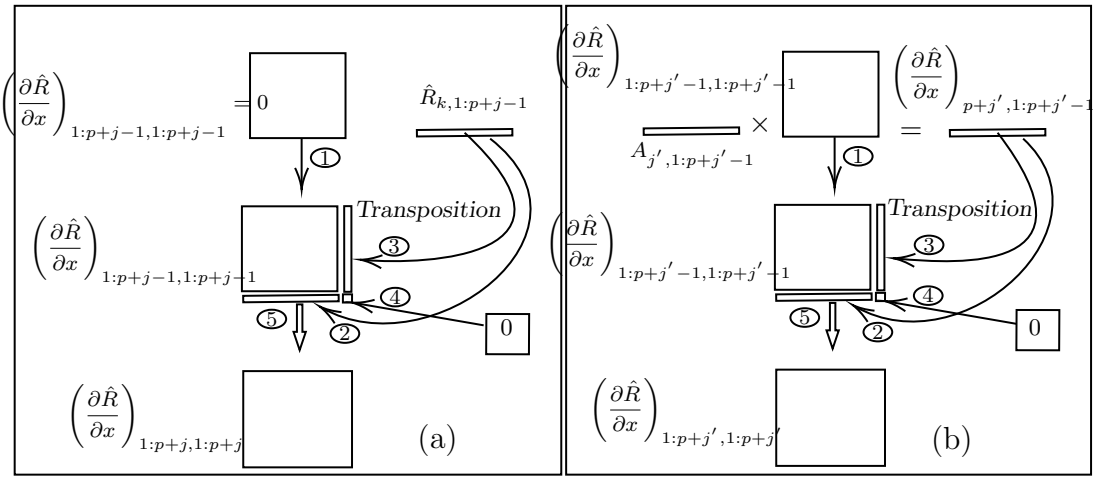

Figure 8: Illustration of Algorithm 2

4. $\left(\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}\right)_{p+s,p+s} = \dfrac{\partial^2}{\partial x \partial y}(1) = 0$. See Figure 9 (a).

5. a)

$$\begin{aligned}
\left(\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}\right)_{p+s',1:p+s'-1} &= \frac{\partial^2}{\partial x \partial y}(\boldsymbol{A}_{s',1:p+s'-1} \widehat{\boldsymbol{R}}_{p+s',1:p+s'-1}) \\
&= \left(\frac{\partial^2}{\partial x \partial y} \boldsymbol{A}_{s',1:p+s'-1}\right) \widehat{\boldsymbol{R}}_{p+s',1:p+s'-1} \\
&\quad + (\frac{\partial}{\partial x} \boldsymbol{A}_{s',1:p+s'-1})(\frac{\partial}{\partial y} \widehat{\boldsymbol{R}}_{p+s',1:p+s'-1}) \\
&\quad + \left(\frac{\partial}{\partial y} \boldsymbol{A}_{s',1:p+s'-1}\right) \left(\frac{\partial}{\partial x} \widehat{\boldsymbol{R}}_{p+s',1:p+s'-1}\right) \\
&\quad + \boldsymbol{A}_{s',1:p+s'-1} \left(\frac{\partial^2}{\partial x \partial y} \widehat{\boldsymbol{R}}_{p+s',1:p+s'-1}\right)
\end{aligned}$$

We know that: $s' > s > j$ thus $\boldsymbol{A}_{s',1:q+s'-1}$ does not depends on x and y.
Hence $\dfrac{\partial}{\partial x} \boldsymbol{A}_{s',1:q+s'-1} = 0$, $\dfrac{\partial}{\partial y} \boldsymbol{A}_{s',1:q+s'-1} = 0$ and $\dfrac{\partial^2}{\partial x \partial y} \boldsymbol{A}_{s',1:q+s'-1} = 0$.
As a consequence, $\left(\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}\right)_{p+s',1:q+s'-1} = \boldsymbol{A}_{s',1:q+s'-1} \left(\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}\right)_{p+s',1:q+s'-1}$

b) Flows from the symmetry of $\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}$

c) $\left(\dfrac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}\right)_{p+s',p+s'} = \dfrac{\partial^2}{\partial x \partial y}(1) = 0$. See Figure 9 (b).

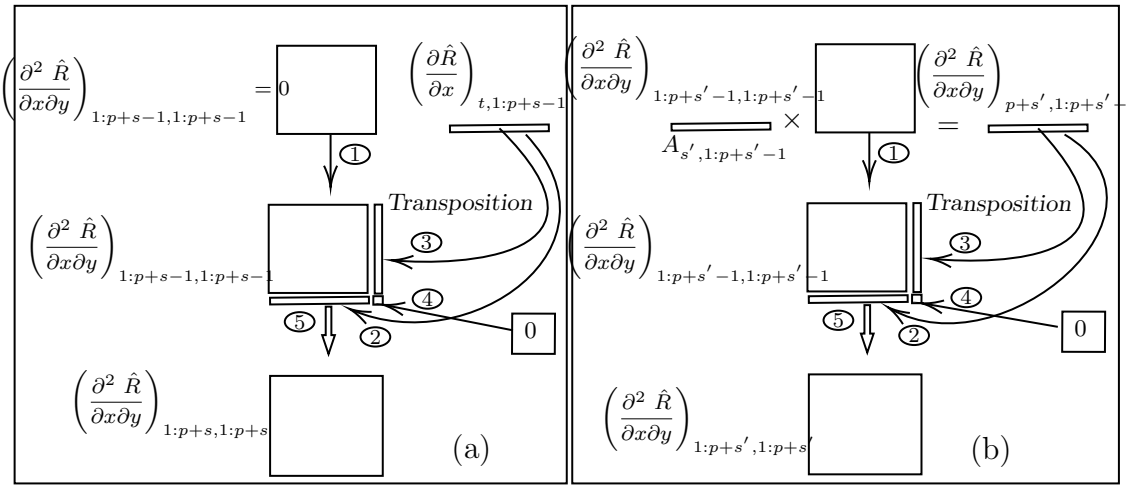

Figure 9: Illustration of Algorithm 3

□

Proof of Corollary 1

Proof. Since x and y are associated with the same endogenous variables then $s = j$ and $t \in \{1 : p + j - 1\}$.

1. If $j < q$.

   a) i. Using theorem 3, $\left(\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}\right)_{1:p+j-1,1:p+j-1} = 0$

   ii. Using theorem 3, $\left(\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}\right)_{p+j,1:p+j-1} = \left(\frac{\partial \widehat{\boldsymbol{R}}}{\partial x}\right)_{t,1:p+j-1}$.

   However since $t \in \{1 : p + j - 1\}$ then using theorem 2, $\left(\frac{\partial \widehat{\boldsymbol{R}}}{\partial x}\right)_{t,1:p+j-1} = 0$.

   iii. From the symmetry of $\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}$ it comes $\left(\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}\right)_{1:p+j-1,p+j} = 0$.

   iv. $\left(\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}\right)_{p+j,p+j} = \frac{\partial^2}{\partial x \partial y}(1) = 0$

   As a consequence $\left(\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}\right)_{1:p+j,1:p+j} = 0..$

b) Using for now the same process as in point 5 of theorem 3, we obtain

$$\left(\frac{\partial^2 \widehat{\boldsymbol{R}}}{\partial x \partial y}\right)_{1:p+j^{'},1:p+j^{'}} = 0, \forall j^{'} \in \{j+1:q\}$$

2. If $j = q$ then the proof stops at point a).

□

References

Bollen, K. A. (1989). *Structural equations with LV*. New York: Wiley.

Breckler, S. J. (1990). Applications of correlation structure modeling in psychology: cause for concern? *Psychol. Bull*, 107:260–273.

Dolce, P., Esposito Vinzi, V., and C., L. (2018). Non-symmetrical composite-based path modeling. *Advances in Data Analysis and Classification*, 12(3):759–784.

Duncan, O. D. (1966). Path analysis: Sociological examples. *American Journal of Sociology*, 72:1–16.

Eisenhauer, N., Bowker, M., Grace, J., and Powell, J. (2015). From patterns to causal understanding: structural equation modeling (sem) in soil ecology. *Pedobiologia*, 58(2):65–72.

El Hadri, Z. and Hanafi, M. (2015). The finite iterative method for calculating the correlation matrix implied by a recursive path model. *Electronic Journal of Applied Statistical Analysis*, 08(01):84–99.

El Hadri, Z. and Hanafi, M. (2016). Extending the finite iterative method for computing the covariance matrix implied by a recursive path model. In *In Saporta, G., Russolillo, G., Trinchera, L., Abdi H, Esposito V.V. (eds.),The Multiple Facets of Partial Least Squares Methods: PLS, Paris, France, 2014*, volume 2, pages 29–43. Springer.

El Hadri, Z., Hanafi, M., Dolce, P., El Kettani, Y., and Iaousse, M. (2019). Computation of the covariance matrix implied by a structural recursive model with latent variables through the finite iterative method. In *In The 5th Edition of the International Conference on Optimization and Applications (ICOA 2019)*, pages 272–280.

Hauser, R. M. e. a. (1975). Education, occupation, and earnings. *New York: Academic Press*.

Iaousse, M., Hmimou, A., El Hadri, Z., and El Kettani, Y. (2020a). A modified algorithm for the computation of the covariance matrix implied by a structural recursive model with latent variables using the finite iterative method. *Statistics, Optimization &; Information Computing*, 8(2):359–373.

Iaousse, M., Hmimou, A., El Hadri, Z., and El Kettani, Y. (2020b). On the computation of the correlation matrix implied by a recursive path model. In *2020 IEEE 6th International Conference on Optimization and Applications (ICOA)*, pages 1–5.

Jöreskog, K. and Wold, H. (1982). *Systems under indirect observation: Causality, structure, prediction, Part I and Part II.* Amsterdam: North Holland.

Jöreskog, K. G. (1970). A general method for the analysis of covariance structures. *Biometrica*, 57:239–251.

Jöreskog, K. G. (1978). Structural analysis of covariance and correlation matrices. *Psychometrika*, 43(4):443–477.

Jöreskog, K. G., Andersen, E. B., Laake, P., Cox, D. R., and Schweder, T. (1981). Analysis of covariance structures [with discussion and reply]. *Scandinavian Journal of Statistics*, 8(2):65–92.

Jöreskog, K. G. and Sorbom, D. (1993). *LISREL 8: Structural equation modeling with the SIMPLIS command language.* Mahwah, NJ: Lawrence Erlbaum Associates.

Pugesek, B. (2003). *Structural Equation Modeling Applications in Ecological and Evolutionary Biology.* New York : Cambridge University Press.

Schumacker, R. E. and Lomax, R. G. (2004). *A beginner's guide to structural equation modeling.* Mahwah, N.J: Lawrence Erlbaum Associates.

Wright, S. (1921). Correlation and causation. *Journal of Agricultural Research*, 20(17):557–585.

Wright, S. (1923). The theory of path coefficients- a reply to niles' criticism. *Genetics*, 8:239–255.

Wright, S. (1934). The method of path coefficients. *Annals of Mathematical Statistics*, 5:161–215.

Wright, S. (1960). Path coefficients and path regressions. alternative or complementarity concepts? *Biometrics*, 16:189–202.