

# A Modified Algorithm for the Computation of the Covariance Matrix Implied by a Structural Recursive Model with Latent Variables Using the Finite Iterative Method

M'barek Iaousse<sup>1,\*</sup>, Amal Hmimou<sup>1</sup>, Zouhair El Hadri<sup>2</sup>, Yousfi El Kettani<sup>1</sup>

<sup>1</sup>*Department of Mathematics, Ibn Tofail University, Morocco*

<sup>2</sup>*Center of Mathematical Researches and Applications, Faculty of sciences Mohammed V University In Rabat, Morocco*

**Abstract** Structural Equation Modeling (SEM) is a statistical technique that assesses a hypothesized causal model by showing whether or not, it fits the available data. One of the major steps in SEM is the computation of the covariance matrix implied by the specified model. This matrix is crucial in estimating the parameters, testing the validity of the model and, make useful interpretations. In the present paper, two methods used for this purpose are presented: the Jöreskog's formula and the finite iterative method. These methods are characterized by the manner of the computation and based on some a priori assumptions. To make the computation more simplistic and the assumptions less restrictive, a new algorithm for the computation of the implied covariance matrix is introduced. It consists of a modification of the finite iterative method. An illustrative example of the proposed method is presented. Furthermore, theoretical and numerical comparisons between the exposed methods with the proposed algorithm are discussed and illustrated.

**Keywords** Structural Equation Modeling, Latent Variable, Path Analysis, Covariance Matrix, Finite Iterative Method.

**AMS 2010 subject classifications** 62J10, 62H20

**DOI:** 10.19139/soic-2310-5070-937

## 1. Introduction

Structural Equation Modeling (SEM) is a set of statistical techniques to test the conformity of a hypothesized causal model to some available data [1, 2, 4]. Structural equation models are characterized by the presence of both observed and latent variables. As a consequence, their procedures are more complicated compared to the classical statistical methods such as linear regression, ANOVA,... In fact, most of the classical models lie in the general case of SEM [2]. Currently, SEM is widely used in many areas. To name a few: environmental sustainability [26], ecological studies [8, 17], social sciences [9, 24] and psychology [3]. SEM combines Path Analysis [18, 19, 20, 21, 9] with Confirmatory Factor Analysis [25, 16] by taking into account that some variables are not directly observed. The general idea of SEM is that the covariance matrix of the observed variables can be expressed as a function of the model parameters [10, 12, 11, 14, 13, 15]. In other words: If  $\Sigma$  is the population covariance matrix of the model variables, and  $\widehat{\Sigma}(\theta)$  is the covariance matrix implied by the model. Then the fundamental hypothesis of SEM can be expressed as:

$$\Sigma = \widehat{\Sigma}(\theta) \quad (1)$$

\*Correspondence to: M'barek Iaousse (Email: iaousse@gmail.com). Department of Mathematics, Ibn Tofail University. Campus Universitaire B.P 242, 14000 Kenitra.

such that  $\theta$  is the vector containing all model parameters.

An SEM model is decomposed into two types of models: a) the structural model, which describes the relations between latent variables. Latent variables are factors that can not be directly measured and; b) the measurement model, which describes the relations between each latent variable and its manifest variables. Manifest (or observed) variables are directly measured in some scale and available on data. Variables are grouped into two categories: i) endogenous variables, which are variables that we want to examine (confirm or disconfirm) their causes and; ii) exogenous variables, which are not determined by the model. In other words, their causes are outside the model for some reason (we don't know their causes or we don't want to specify their causes in the model). Mathematically, we can summarize the above definition of an SEM model as follows:

A Structural Equation Model for a set of variables

$$\{\xi^t = \{\xi_1, \dots, \xi_q\}, \eta^t = \{\eta_1, \dots, \eta_p\}, \mathbf{X}^t = \{X_1, \dots, X_{n_x}\}, \mathbf{Y}^t = \{Y_1, \dots, Y_{n_y}\}\}$$

is defined as the set of equations of the form:

$$\eta = B\eta + \Gamma\xi + \zeta \quad (2)$$

$$X = \Lambda_x\xi + \delta \quad (3)$$

$$Y = \Lambda_y\eta + \epsilon \quad (4)$$

Such that:  $\zeta$  is the vector of disturbances that refer to all non-included causes in the structural model and  $\delta$  and  $\epsilon$  are the vectors of the measurement errors of each measurement model. All those vectors are error terms.

The following matrix notations are adopted:  $B$  is the matrix ( $p \times p$ ) of structural coefficients between endogenous variables.  $\Gamma$  is the matrix ( $p \times q$ ) of structural coefficients between exogenous and endogenous variables.  $\Lambda_x$  is the matrix ( $n_x \times q$ ) of loadings between exogenous latent variables and their associated manifest variables.  $\Lambda_y$  is the matrix ( $n_y \times p$ ) of loadings between endogenous latent variables and their associated manifest variables. The covariance matrix of the vector of exogenous variables is denoted  $\Phi$  ( $q \times q$ ). The covariance matrix of the vector of disturbances (i.e.  $\zeta$ ) is denoted  $\Psi$  ( $p \times p$ ). The covariance matrix of the vector of measurement errors associated with exogenous variables (i.e.  $\delta$ ) is denoted  $\Theta_\delta$  ( $n_x \times n_x$ ). Finally, The covariance matrix of the vector of measurement errors associated with endogenous variables (i.e.  $\epsilon$ ) is denoted  $\Theta_\epsilon$  ( $n_y \times n_y$ ).

Another important aspect of SEM is the path diagram. It is defined as the graphical representation of the model. Many researchers [2] think that it can be more representative than the structural equations. A path diagram is a directed graph (example Figure 1). The vertices correspond to variables (latent variables, observed variables, and error terms). It is an illustration of the direction of a hypothesized causation between each pair of variables. The edges are i) the one-headed arrow which means the sense of causality between variables and; ii) the two-headed arrow which means that the relationship between its ending vertices (variables) is not analyzed by the model (unanalyzed association). The relations described by the path diagram is complete so that, for example, the non-existence of an arrow between two variables means the nonexistence of a direct association between them. An SEM model is said to be recursive if for every pair of variables if the one causes the second then the second can never be a cause of the first. Otherwise, the model is said to nonrecursive. mathematically, a model is said to be recursive if there exists a permutation of the lines of the matrix  $B$  so that it will be strictly lower triangular (in all what follows, we will consider that the matrix  $B$  permuted so that it is strictly lower triangular for a recursive model). Graphically, a model is said to be recursive if its path diagram acyclic (i.e. contain no feedback loops nor reciprocal causations). SEM models consist of 6 major steps: model specification, model identification, model estimation, model validation, model modification and, interpretation [27].

The estimation is a crucial step in SEM. It consists of expressing the Covariance matrix implied by the model (i.e.  $\hat{\Sigma}$ ) as a function of model (the vector of model parameters is denoted  $\theta$ ) parameters and then find the estimates of those parameters. This is usually done by minimizing a discrepancy function between the population covariance

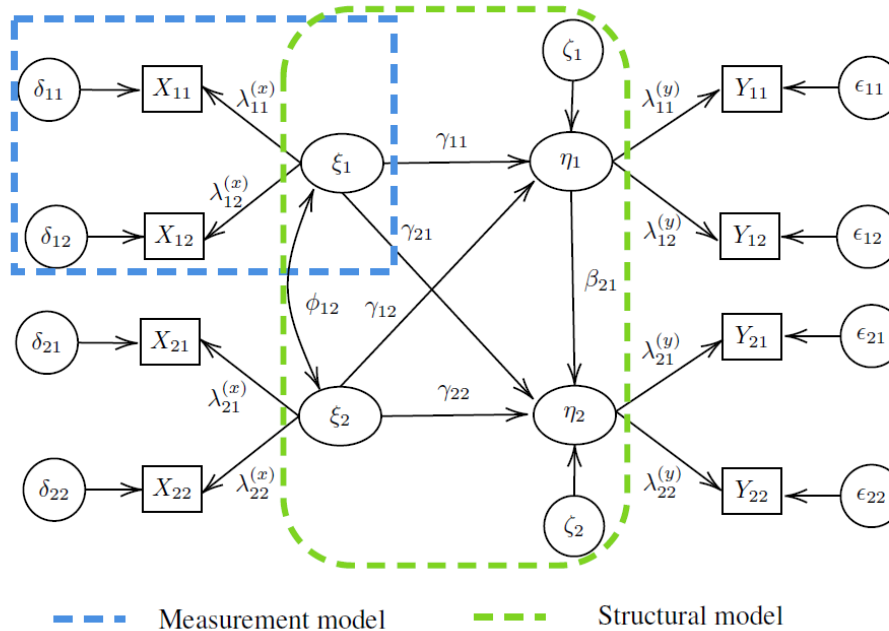


Figure 1. An Example of SEM model with four latent variables and eight observed variables

matrix of model variables  $\Sigma$  and the covariance matrix reproduced by the model (the covariance matrix implied by the model  $\widehat{\Sigma}$ )<sup>†</sup>.

The main topic of the present paper is: For an SEM model, given its parameters, how to compute  $\widehat{\Sigma}$ ? The first answer to this question was proposed by [10]. Then [5] and [6] proposed an iterative method called Finite Iterative Method (FIM) for Path Analysis models (SEM with observed variables). FIM has been then extended to SEM models [7]. This paper aims to introduce a modified FIM algorithm which has some advantages, that will be discussed later, compared to the previous methods.

For this purpose, the manuscript is partitioned as follows: section 2 discusses briefly the implied covariance matrix and the previous methods used for its computation. section 3 exposes the modified algorithm of the FIM method with its demonstration. sections 4 and 5 illustrate the seen methods as well as the comparison between them through an example. Finally, section 6 concludes with discussions and perspectives.

## 2. The implied covariance matrix: the Jöreskog’s formula and the finite iterative method

The present section introduces the concept of the implied Covariance matrix and the methods used for its computation, namely Jöresko’s formula [10] and the Finite Iterative Method [7].

### 2.1. Definition

The estimation step consists of finding the parameters that make the model fit the data. In other words, it aims to find values of the parameters that make the structure of the model as close as possible to the data structure. The model structure is characterized by the assumption that the distribution of the observed variables is multivariate

<sup>†</sup> Another approach in SEM PLS Path Modeling. This approach introduced by [28] is the variance-based method. While the treated approach in this manuscript is the covariance-based method introduced by [10]. The authors suggest [33] to the reader for PLS Path Modeling approach

and sufficiently well described by its mean and covariance [11].

This is translated by the fundamental hypothesis of SEM (1). The implied covariance matrix is the theoretical covariance matrix between variables generated by an SEM model. It is written as a function of the free model parameters. For instance, the covariance matrix implied by the model defined by (2), (3) and (4) is written as follows:

$$\widehat{\Sigma} = \begin{pmatrix} \mathbb{E}[\mathbf{X}\mathbf{X}^t] & \mathbb{E}[\mathbf{X}\mathbf{Y}^t] \\ \mathbb{E}[\mathbf{Y}\mathbf{X}^t] & \mathbb{E}[\mathbf{Y}\mathbf{Y}^t] \end{pmatrix} \quad (5)$$

On the one hand, it can be determined algebraically by writing the vector of endogenous variables as a function of the vector of exogenous ones [10]. On the other hand, it can also be constructed graphically using the tracing rules of Path Analysis [22, 29], by calculating every covariance (or correlation in Path Analysis models) between each pair of model variables. [5, 6], and [7] proposed the Finite Iterative Method, which uses these rules to build an implied correlation matrix with an iterative algorithm. Further discussion in terms of the comparison between Jöreskog's formula and the Finite iterative method is given in [23].

## 2.2. Jöreskog's formula

As above mentioned, The Jöreskog's formula is a compact representation of the Implied covariance matrix given in (5). In order to compute the implied covariance matrix, Jöreskog starts from the following assumptions:

- All variables are supposed to be centred to their means.
- All disturbances and measurement errors are uncorrelated with each other. That means :

$$\Psi = \mathbb{E}[\zeta\zeta^t], \Theta_\delta = \mathbb{E}[\delta\delta^t] \text{ and } \Theta_\epsilon = \mathbb{E}[\epsilon\epsilon^t]$$

are diagonal matrices and

$$\mathbb{E}[\zeta\epsilon^t] = \mathbf{0}, \mathbb{E}[\zeta\delta^t] = \mathbf{0} \text{ and } \mathbb{E}[\delta\epsilon^t] = \mathbf{0}^\ddagger$$

- Exogenous variables are uncorrelated with disturbances and with measurement errors. That is :

$$\mathbb{E}[\xi\zeta^t] = \mathbf{0}, \mathbb{E}[\xi\delta^t] = \mathbf{0} \text{ and } \mathbb{E}[\xi\epsilon^t] = \mathbf{0}$$

- The matrix  $(I - B)$  is nonsingular.

The equation (2) can be written in the compact form:

$$\boldsymbol{\eta} = (I - B)^{-1}\Gamma\xi + (I - B)^{-1}\zeta \quad (6)$$

Using equation (3), (4) and (6) and taking the mathematical expectation of  $\mathbf{X}\mathbf{X}^t$ ,  $\mathbf{X}\mathbf{Y}^t$  and  $\mathbf{Y}\mathbf{Y}^t$ , the implied covariance matrix in (5) will be :

$$\widehat{\Sigma}(\Theta) = \begin{pmatrix} \Lambda_x\Phi\Lambda_x^t + \Theta_\delta & \Lambda_x\Phi\Gamma^t(I - B^t)^{-1}\Lambda_y^t \\ \Lambda_y(I - B)^{-1}\Gamma\Phi\Lambda_x^t & \Lambda_y(I - B)^{-1}(\Gamma\Phi\Gamma^t + \Psi)(I - B^t)^{-1}\Lambda_y^t + \Theta_\epsilon \end{pmatrix} \quad (7)$$

## 2.3. The finite iterative method

The finite Iterative Method is an iterative method which does not take into account the error terms in the computation of the implied correlation matrix. To be more clear, the parameters which are related to the error terms are constrained (set as a function of the rest of model variables). The starting point of the method is to separate the model into two types of submodels: a) the structural model (the inner model) and b) the measurement models. (see Figure 2 as an example).

<sup>‡</sup>From now on in this paper, all null vectors and null matrices are represented by a bold zero (**0**). The reader shall detect its order according to the equation in which it appears.

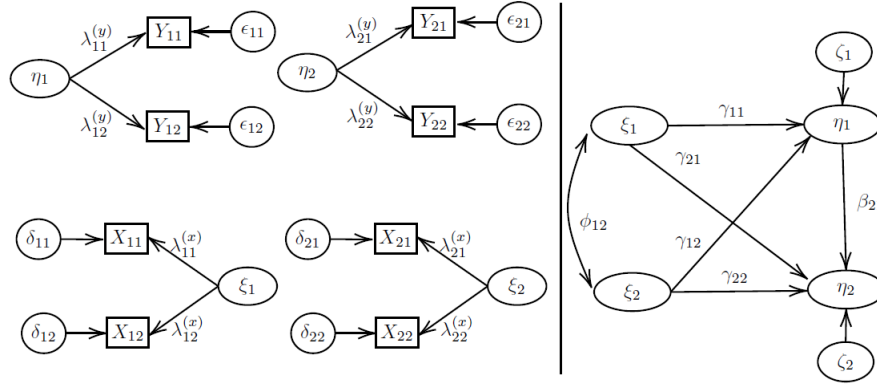


Figure 2. The separation of the structural model and the measurement models

In addition to that FIM has its own notations and assumptions. Indeed, All latent variables are represented by  $\xi$  (i.e.  $\xi = (\xi_1, \dots, \xi_q, \eta_1, \dots, \eta_p)$ ). The vector of all manifest variables is denoted  $X$  and the vector all measurement errors is denoted  $\delta$ . Finally for a matrix  $M$ , the following notation is adopted  $M_{i_1:i_2, j_1:j_2} = (M_{ij})_{i_1 \leq i \leq i_2, j_1 \leq j \leq j_2}$ . The equation (2) is becoming

$$\xi_{q+1:q+p} = A\xi + \zeta \tag{8}$$

such that:

$$A = (\Gamma B)$$

$A : (p \times (q + p))$  is the matrix of the structural coefficients.

Besides, the equations (3) and (4) become:

$$X = \Lambda\xi + \delta \tag{9}$$

Where

$$\Lambda = \begin{pmatrix} \Lambda_x & \mathbf{0} \\ \mathbf{0} & \Lambda_y \end{pmatrix}$$

is the matrix  $((n_x + n_y) \times (n_x + n_y))$  of all the loadings. The FIM algorithm proposed by [7] is based on the following assumptions:

- The model is supposed to be recursive.
- All manifest variables are centred and all latent variables are standardized.
- Disturbances are uncorrelated with the latent variables explaining its associated endogenous latent variables. In other words:

$$\forall j \in \{1, \dots, p\} \forall i \in \{1, \dots, p + j - 1\} : \mathbb{E}[\zeta_j \xi_i] = 0$$

- The measurement errors are uncorrelated with each other (i.e.  $\Theta = \mathbb{E}[\delta\delta^t] = I_{q+p}$ ).
- All latent variables are uncorrelated with the measurement errors (i.e.  $\mathbb{E}[\xi\delta^t] = \mathbf{0}$ ).

The first stage of the computation of the covariance matrix implied by a structural recursive model is the implied correlation matrix  $\widehat{R}^{(LV)}$  of the structural model (since the latent variables are standardized) using Algorithm 1 proposed by [5].

Then the implied covariance matrix of the whole model is given as follows:

$$\widehat{\Sigma}^{FIM} = \Lambda \widehat{R}^{(LV)} \Lambda^t + \Theta \tag{10}$$

such that  $\widehat{R}^{LV} = \mathbb{E}[\xi\xi^t]$  is the matrix computed in the first stage by Algorithm 1. Besides, elements of  $\Theta$  are constrained by the following equation:

$$\Theta = \text{diag}(s_{jh} - \lambda_{jh}^2)_{1 \leq j \leq q+p, 1 \leq h \leq n_j} \tag{11}$$

**Algorithm 1** Finite Iterative Method to compute  $\widehat{R}^{(LV)}$  [7]

```

initialize:  $\widehat{R}_{1:q,1:q}^{(LV)} = \Phi$  and  $j = 1$ 
repeat
  1.  $\widehat{R}_{q+j,1:q+j-1}^{(LV)} = A_{j,1:q+j-1} \widehat{R}_{1:q+j-1,1:q+j-1}^{(LV)}$ 
  2.  $\widehat{R}_{1:q+j-1,q+j}^{(LV)} = (\widehat{R}_{q+j,1:q+j-1}^{(LV)})^t$ 
  3.  $\widehat{R}_{q+j,q+j}^{(LV)} = 1$ 
  4.  $j \leftarrow j + 1$ 
until  $j = p$ 
    
```

**3. A modified Algorithm of FIM**

In this section we will present another view to the extension of FIM to SEM models which have more simplistic algorithm and less restrictive assumptions compared the FIM. The main Idea is to adapt Algorithm 1 with new input so it can be applied in one stage instead of two stages as in section (2). The algorithm (Algorithm 2 bellow) will compute the implied correlation matrix of all model variables and extract the covariance matrix of the observed variables.

**3.1. Notations and assumptions**

First, all variables, latent and observed, that appear dependent in the equations (2), (3) and (4) (i.e.  $\eta$ ,  $X$  and  $Y$ ) will be gathered in a vector denoted  $\eta^*$ . The same for disturbances and measurement errors (i.e.  $\zeta$ ,  $\delta$ ,  $\epsilon$ ), they will be gathered in one vector of errors terms denoted  $\zeta^*$ . Two new matrices  $B^* : ((p + n_x + n_y) \times (p + n_x + n_y))$  and  $\Gamma^* : ((p + n_x + n_y) \times q)$  will be defined as follows:

$$B^* = \begin{pmatrix} B & 0 & 0 \\ 0 & 0 & 0 \\ \Lambda_y & 0 & 0 \end{pmatrix} \tag{12}$$

$$\Gamma^* = \begin{pmatrix} \Gamma \\ \Lambda_x \\ 0 \end{pmatrix} \tag{13}$$

Indeed, the equations (2), (3) and (4) will be gathered in the following one:

$$\eta^* = B^* \eta^* + \Gamma^* \xi + \zeta^* \tag{14}$$

Which can also written as :

$$\eta^* = A^* \begin{pmatrix} \xi \\ \eta^* \end{pmatrix} + \zeta^* \tag{15}$$

Where  $A^*$ , a matrix of order  $((p + n_x + n_y) \times (q + p + n_x + n_y))$  defined as :

$$A^* = \begin{pmatrix} \Gamma^* B^* \\ \Gamma & B & 0 & 0 \\ \Lambda_x & 0 & 0 & 0 \\ 0 & \Lambda_y & 0 & 0 \end{pmatrix} \tag{16}$$

The necessary assumptions for our method are:

- The model is supposed to be recursive.
- All manifest variables are centred to their mean.

- In each equation, the error term is uncorrelated with the independent variables. In other words, using (15) :

$$\forall j \in \{1, \dots, p + n_x + n_y\} : \forall h \in \{1, \dots, q + p + n_x + n_y\} : \mathbf{A}_{j,h}^* \neq 0 \Rightarrow \mathbb{E}[\zeta_j^* \xi_h^*] = 0$$

such that  $\xi_h^* \in \{\xi_1, \dots, \xi_q, \eta_1, \dots, \eta_p, X_1, \dots, X_{n_x}, Y_1, \dots, Y_{n_y}\}$  for each  $h$  in  $\{1, \dots, q + p + n_x + n_y\}$ .

### 3.2. A modified FIM algorithm

Our proposed algorithm (Algorithm 2) is based on the implied covariance matrix of the whole model variables (latent and manifest except error terms). It is expressed as follows:

$$\widehat{\Sigma}^* = \mathbb{E} \left[ \begin{pmatrix} \xi \\ \eta^* \end{pmatrix} \begin{pmatrix} \xi^t & \eta^{*t} \end{pmatrix} \right] = \begin{pmatrix} \mathbb{E}[\xi\xi^t] & \mathbb{E}[\xi\eta^t] & \mathbb{E}[\xi\mathbf{X}^t] & \mathbb{E}[\xi\mathbf{Y}^t] \\ \mathbb{E}[\eta\xi^t] & \mathbb{E}[\eta\eta^t] & \mathbb{E}[\eta\mathbf{X}^t] & \mathbb{E}[\eta\mathbf{Y}^t] \\ \mathbb{E}[\mathbf{X}\xi^t] & \mathbb{E}[\mathbf{X}\eta^t] & \mathbb{E}[\mathbf{X}\mathbf{X}^t] & \mathbb{E}[\mathbf{X}\mathbf{Y}^t] \\ \mathbb{E}[\mathbf{Y}\xi^t] & \mathbb{E}[\mathbf{Y}\eta^t] & \mathbb{E}[\mathbf{Y}\mathbf{X}^t] & \mathbb{E}[\mathbf{Y}\mathbf{Y}^t] \end{pmatrix} \quad (17)$$

Once the matrix (17) is obtained, we extract from it the covariance matrix of the observed variables ( $\widehat{\Sigma}$ ). The model became defined by (14).

---

#### Algorithm 2 Modified Finite Iterative Method to compute $\widehat{\Sigma}$

---

0. **initialize:**  $\widehat{\Sigma}_{1:q,1:q}^* = \Phi$ ,  $\mathbf{G} = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix}$  and  $j = 1$

**repeat**

1.  $\widehat{\Sigma}_{q+j,1:q+j-1}^* = \mathbf{A}_{j,1:q+j-1}^* \widehat{\Sigma}_{1:q+j-1,1:q+j-1}^*$
2.  $\widehat{\Sigma}_{1:q+j-1,q+j}^* = (\widehat{\Sigma}_{q+j,1:q+j-1}^*)^t$
3.  $\widehat{\Sigma}_{q+j,q+j}^* = \sigma_{\eta_j}^2$
4.  $j \leftarrow j + 1$

**until**  $j = p + n_x + n_y$

5. **compute:**  $\widehat{\Sigma} = \mathbf{G}\widehat{\Sigma}^*\mathbf{G}^t$

---

#### Theorem 1

For a given recursive structural model, if  $\widehat{\Sigma}$  and  $\widehat{\Sigma}^{MFIM}$  are respectively the implied covariance matrix and the matrix computed using Algorithm 2. Then we have the following equality:

$$\widehat{\Sigma}^{MFIM} = \widehat{\Sigma} \quad (18)$$

#### Proof

Consider a recursive structural model defined by equations (2), (3) and (4). Using notations in 3.1, the model became deined by (14). Now let  $\widehat{\Sigma}^*$  be its implied covariance matrix and  $\widehat{\Sigma}^{MFIM*}$  be the matrix computed using algorithm 2. The idea is to prove, first of all, that:

$$\widehat{\Sigma}^{MFIM*} = \widehat{\Sigma}^* \quad (19)$$

In other words,

$$\forall j \in \{0, \dots, p + n_x + n_y\} : \widehat{\Sigma}_{1:q+j,1:q+j}^{MFIM*} = \widehat{\Sigma}_{1:q+j,1:q+j}^*$$

For this pupose, we will proceed by induction on  $j \in \{0, \dots, p + n_x + n_y\}$ .

★ For  $j = 0$  :

- according to the initialization step of Algorithm 2 :

$$\widehat{\Sigma}_{1:q,1:q}^{MFIM*} = \Phi$$

- by the definition of the implied covariance matrix:  $\widehat{\Sigma}_{1:q,1:q}^* = \mathbb{E}[\xi\xi^t] = \Phi$

Therefore :

$$\widehat{\Sigma}_{1:q,1:q}^{MFIM^*} = \widehat{\Sigma}_{1:q,1:q}^* \tag{20}$$

★ Now suppose that for  $j \in \{0, \dots, p + n_x + n_y - 1\}$ :

$$\widehat{\Sigma}_{1:q+j,1:q+j}^{MFIM^*} = \widehat{\Sigma}_{1:q+j,1:q+j}^* \tag{21}$$

★ lets prove that :

$$\widehat{\Sigma}_{1:q+j+1,1:q+j+1}^{MFIM^*} = \widehat{\Sigma}_{1:q+j+1,1:q+j+1}^* \tag{22}$$

Since:

1. The matrix  $\widehat{\Sigma}^{MFIM^*}$  is symetric by its construction from Algorithm 2,
2. the matrix  $\widehat{\Sigma}^*$  is symetric by the definition of a covariance matrix,
3. the diagonal of the matrice are equal:

$$\widehat{\Sigma}_{h,h}^{MFIM^*} = \sigma_h^2 = \mathbb{E}[(\eta_h^*)^2] = \widehat{\Sigma}_{h,h}^*$$

for each  $h \in \{q + 1, \dots, q + p + n_x + n_y\}$ .

Then, proving (22) is reduced to proving :

$$\widehat{\Sigma}_{q+j+1,1:q+j}^{MFIM^*} = \widehat{\Sigma}_{q+j+1,1:q+j}^* \tag{23}$$

Denote  $\eta^* = (\eta, X, Y) = (\eta_1, \dots, \eta_p, X_1, \dots, X_{n_x}, Y_1, \dots, Y_{n_y}) = (\eta_1^*, \dots, \eta_{p+n_x+n_y}^*)$

Therefore :

- According to Algorithm 2:

$$\widehat{\Sigma}_{q+j+1,1:q+j}^{MFIM^*} = A_{q+j+1,1:q+j}^* \widehat{\Sigma}_{1:q+j,1:q+j}^{MFIM^*}$$

according the induction hypothesis (21), this means :

$$\widehat{\Sigma}_{q+j+1,1:q+j}^{MFIM^*} = A_{q+j+1,1:q+j}^* \widehat{\Sigma}_{1:q+j,1:q+j}^* \tag{24}$$

- From the definition of the covariance matrix of model variables (17) :

$$\widehat{\Sigma}_{q+j+1,1:q+j}^* = \mathbb{E}[\eta_{j+1}^* (\xi^t, \eta_1^*, \dots, \eta_j^*)] \tag{25}$$

- From (15)

$$\eta_{j+1}^* = A_{j+1,1:q+p+n_x+n_y}^* \begin{pmatrix} \xi \\ \eta^* \end{pmatrix} + \zeta_{j+1}^* \tag{26}$$

There are two cases for  $j + 1$ :

- **case 1** :  $j + 1 \in \{1, \dots, p\}$
- **case 2** :  $j + 1 \in \{p + 1, \dots, p + n_x + n_y\}$

**Case 1:** Since the model is recursive, the matrix  $B$  is stricly lower triangular. Therefore, according to (12), the matrix  $B^*$  is stricly lower triangular. That means :

$$\forall h \in \{1, \dots, p\} : B_{h,h:p} = 0$$

On the other hand, from (16) :  $A_{h,q+p+1:q+p+n_x+n_y}^* = 0$

by taking  $h = j + 1$  we conclude that :  $A_{j+1,q+p+n_x+n_y}^* = 0$

**Case 2:** from (16) :  $\forall h \in \{p + 1, \dots, p + n_x + n_y\} : A_{h,h:q+p+n_x+n_y}^* = 0$ .

by taking  $h = j + 1$  we conclude that :  $A_{j+1,q+p+n_x+n_y}^* = 0$ .



In both cases :

$$\mathbf{A}_{j+1, q+p+n_x+n_y}^* = \mathbf{0}$$

using this equation in (26) we get :

$$\begin{aligned} \eta_{j+1}^* &= \mathbf{A}_{j+1, 1:q+p+n_x+n_y}^* \begin{pmatrix} \xi \\ \eta^* \end{pmatrix} + \zeta_{j+1}^* \\ &= \mathbf{A}_{j+1, 1:j}^* \begin{pmatrix} \xi \\ \eta_1^* \\ \vdots \\ \eta_j^* \end{pmatrix} + \zeta_{j+1}^* \end{aligned} \tag{27}$$

Now we put (27) in (25), we use the assumption that  $\zeta_{j+1}^*$  is uncorrelated with the independent variables in the  $j^{th}$  equation and we get :

$$\begin{aligned} \widehat{\Sigma}_{q+j+1, 1:q+j}^* &= \mathbb{E}[(\mathbf{A}_{j+1, 1:q+j}^* (\xi^t, \eta_1^*, \dots, \eta_j^*)^t + \zeta_{j+1}^*) (\xi^t, \eta_1, \dots, \eta_j)] \\ &= \mathbf{A}_{j+1, 1:q+j}^* \mathbb{E}[(\xi^t, \eta_1^*, \dots, \eta_j^*)^t (\xi^t, \eta_1, \dots, \eta_j)] \\ &= \mathbf{A}_{j+1, 1:q+j}^* \widehat{\Sigma}_{1:q+j, 1:q+j}^* \end{aligned} \tag{28}$$

From (28) and (24), we conclude that (23) is correct, which means : (22) is correct.

By the principle of induction, from (20), (21) and, (22) : (19) is correct.

Finally, we apply the fifth step of Algorithm 2 by multiplying both sides of (19) with  $\mathbf{G}$  at the left and  $\mathbf{G}^t$  at the right:

$$\begin{aligned} \widehat{\Sigma}^{MFIM} &= \mathbf{G} \widehat{\Sigma}^{MFIM} \mathbf{G}^t \\ &= \mathbf{G} \widehat{\Sigma}^* \mathbf{G}^t \\ &= \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbb{E}[\xi \xi^t] & \mathbb{E}[\xi \eta^t] & \mathbb{E}[\xi \mathbf{X}^t] & \mathbb{E}[\xi \mathbf{Y}^t] \\ \mathbb{E}[\eta \xi^t] & \mathbb{E}[\eta \eta^t] & \mathbb{E}[\eta \mathbf{X}^t] & \mathbb{E}[\eta \mathbf{Y}^t] \\ \mathbb{E}[\mathbf{X} \xi^t] & \mathbb{E}[\mathbf{X} \eta^t] & \mathbb{E}[\mathbf{X} \mathbf{X}^t] & \mathbb{E}[\mathbf{X} \mathbf{Y}^t] \\ \mathbb{E}[\mathbf{Y} \xi^t] & \mathbb{E}[\mathbf{Y} \eta^t] & \mathbb{E}[\mathbf{Y} \mathbf{X}^t] & \mathbb{E}[\mathbf{Y} \mathbf{Y}^t] \end{pmatrix} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \\ \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \\ &= \begin{pmatrix} \mathbb{E}[\mathbf{X} \mathbf{X}^t] & \mathbb{E}[\mathbf{X} \mathbf{Y}^t] \\ \mathbb{E}[\mathbf{Y} \mathbf{X}^t] & \mathbb{E}[\mathbf{Y} \mathbf{Y}^t] \end{pmatrix} \\ &= \widehat{\Sigma} \end{aligned} \tag{29}$$

Therefore: from (29), (18) is correct. □

In order to simplify Algorithm 2, we propose the following corollary to make it expressed as a recursive function:

*Corollary 1*

Let  $\widehat{\Sigma}$  be the covariance matrix implied by a structural recursive model, and let

$$\begin{aligned} f: \{0, \dots, p + n_x + n_y\} &\rightarrow \mathcal{M}(\mathbb{R}) \\ j &\mapsto f(j) \end{aligned}$$

be a recursive function such that:

$$f(j) = \begin{cases} \Phi & \text{if } j = 0 \\ \left( \frac{f(j-1)}{\mathbf{A}_{j, 1:q+j-1}^* f(j-1)_{1:q+j-1, 1:q+j-1}} \mid \frac{(\mathbf{A}_{j, 1:q+j-1}^* f(j-1)_{1:q+j-1, 1:q+j-1})^t}{\sigma_j^2} \right) & \text{if } j \geq 1 \end{cases} \tag{30}$$

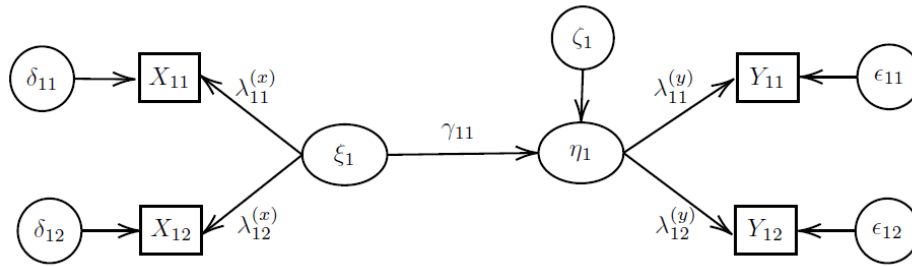


Figure 3. An illustrative example of an SEM model with two latent variables and four manifest variables

Then:

$$\widehat{\Sigma} = \mathbf{G}f(p + n_x + n_y)\mathbf{G}^t \tag{31}$$

*Proof*

On the one hand, from Theorem 1:

$$\widehat{\Sigma}^{MFIM} = \widehat{\Sigma}$$

On the other hand, by the definition of the function  $f$  :

$$\forall j \in \{1, \dots, p + n_x + n_y\} : f(j) = \widehat{\Sigma}_{1:q+j, 1:q+j}^{MFIM\star}$$

Then

$$f(p + n_x + n_y) = \widehat{\Sigma}_{1:q+p+n_x+n_y, 1:q+p+n_x+n_y}^{MFIM\star} = \widehat{\Sigma}^{MFIM\star}$$

Finally:

$$\mathbf{G}f(p + n_x + n_y)\mathbf{G}^t = \widehat{\Sigma}^{MFIM} = \widehat{\Sigma}$$

since each recursive algorithm can be represented by a recursive function. Therefore (31) is correct. □

#### 4. Example for illustration

In this section will illustrate this proposed method using Figure 3 for a better understanding of our proposed method (Algorithm 2).

The model contain, for representability and simplicity reasons, tow latent variables (exogenous variable  $\xi_1$  and endogenous variable  $\eta_1$ ) and for manifest variables (tow variables related to the exogenous variable  $X_{11}$  and  $X_{12}$  and two variables related to the endogenous variable  $Y_{11}$  and  $Y_{12}$ ) notations of section 3.1 are replaced, for the same reasons, by the following ones:

$$a = \lambda_{11}^{(x)}, b = \lambda_{12}^{(x)}, c = \gamma_{11}, d = \lambda_{11}^{(y)} \text{ and, } e = \lambda_{12}^{(y)}$$

Hence the matrix  $\mathbf{A}^*$  is given as follows:

$$\mathbf{A}^* = \begin{pmatrix} c & 0 & 0 & 0 \\ a & 0 & 0 & 0 \\ b & 0 & 0 & 0 \\ 0 & d & 0 & 0 \\ 0 & e & 0 & 0 \end{pmatrix}$$

Denote the variance of  $\xi_1$  by  $\phi_1$  and the variance of the variable  $\eta^*$  by  $\sigma_{\eta^*}^2$  for each  $\eta^*$  in  $\{\eta_1, X_{11}, X_{12}, Y_{11}, Y_{12}\}$ . Applying steps 0 to 4 of Algorithm 2 lead to the following matrix:

$$\widehat{\Sigma}^* = \begin{pmatrix} \phi_1 & & & & & & \\ c\phi_1 & \sigma_{\eta_1}^2 & & & & & \\ a\phi_1 & ac\phi_1 & \sigma_{X_{11}}^2 & & & & \\ b\phi_1 & bc\phi_1 & ab\phi_1 & \sigma_{X_{12}}^2 & & & \\ dc\phi_1 & d\sigma_{\eta_1}^2 & acd\phi_1 & bcd\phi_1 & \sigma_{Y_{11}}^2 & & \\ ec\phi_1 & e\sigma_{\eta_1}^2 & ace\phi_1 & bce\phi_1 & de\sigma_{\eta_1}^2 & \sigma_{Y_{12}}^2 & \end{pmatrix} \tag{32}$$

Applying the fifth step of Algorithm 2 lead to the implied covariance matrix as follows :

$$\widehat{\Sigma} = \begin{pmatrix} \sigma_{X_{11}}^2 & & & & & \\ ab\phi_1 & \sigma_{X_{12}}^2 & & & & \\ acd\phi_1 & bcd\phi_1 & \sigma_{Y_{11}}^2 & & & \\ ace\phi_1 & bce\phi_1 & de\sigma_{\eta_1}^2 & \sigma_{Y_{12}}^2 & & \end{pmatrix} \tag{33}$$

### 5. Theoretical and numerical comparison

The present section exposes some comparisons between the methods discussed previously. More precisely, we will show some theoretical comparison in terms of the manner of the computation as well as the assumptions adopted by each method. Then we will show, through a simulation example the practical differences between them in different situations.

#### 5.1. Theoretical comparison

The differences between Jöreskog’s formula, the Finite Iterative Method, and our proposed modification of FIM can be seen in two main levels:

- Computational level:

At this level, Jöreskog’s formula compute the covariance matrix implied by a structural model by block. It gives a compact expression of our desired covariance matrix. In addition, it requires the inversion of the matrix  $(I - B)$  [7]. On the other side, both algorithms (Algorithm 1 and Algorithm 2) construct the covariance matrix iteratively. Jöreskog’s formula requires  $\Gamma, B, \Lambda_x, \Lambda_y, \Phi, \Psi, \Theta_\delta$  and,  $\Theta_\epsilon$ . For its part, FIM requires  $\Gamma, B, \Lambda_x, \Lambda_y, \Phi, \Theta_\delta$  and,  $\Theta_\epsilon$ . However, using the constraint in (11) will make  $\Theta_\delta$  and  $\Theta_\epsilon$  not needed for the computation [7]. Our proposed method have the same model parameters as the FIM plus the variances of the manifest and latent variables. In fact, in FIM these parameters are not took into account because they are constrained to be equal to the empirical variances for manifest variables and 1 for latent variables. For us, no need to impose this constraint nor the constraint (11) explicitly.

- Assumptions level:

The main advantage of Jöreskog’s formula is that the model does not need to be recursive. A less restrictive assumption adopted by Jöreskog is that the matrix  $(I - B)$  is not singular. On the other hand, FIM and our proposed method require a recursive model. However, research on the computation of the covariance matrix implied by a structural nonrecursive model is underway. Besides, Jöreskog assumes that all error terms (i.e. measurement errors and disturbances) are uncorrelated with each other and uncorrelated with all exogenous variables. FIM assumes that measurement errors are uncorrelated with each other and each disturbance is uncorrelated with all independent variables that are in the same equation containing this error term. This is less restrictive compared to Jöreskog’s formula. Least restrictive assumptions are ours. In fact, we assume each error term is uncorrelated with all independent variables that are in the same equation containing this error term without adding the assumption that measurement errors are uncorrelated with each other. Table (1) summarize all these comparisons.

Method	Computation	Assumptions
Jöreskog	<ul style="list-style-type: none"> <li>• Compact</li> <li>• By block</li> </ul>	<ul style="list-style-type: none"> <li>• <math>(I - B)</math> is invertible (less restrictive)</li> </ul>
		<ul style="list-style-type: none"> <li>• All error terms are uncorrelated with each other</li> <li>• All error terms are uncorrelated with all exogenous variables</li> <li>• <math>\Rightarrow</math>(Restrictive)</li> </ul>
FIM	<ul style="list-style-type: none"> <li>• Step1: iterative</li> </ul>	<ul style="list-style-type: none"> <li>• Recursive model (restrictive)</li> <li>• All measurement errors are uncorrelated with each other</li> <li>• All disturbances are uncorrelated with all independent variables appearing with it in the same equation</li> <li>• <math>\Rightarrow</math>(Less restrictive)</li> </ul>
	<ul style="list-style-type: none"> <li>• Step2: compact</li> </ul>	<ul style="list-style-type: none"> <li>• All latent variables are standardized</li> </ul>
MFIM	<ul style="list-style-type: none"> <li>• Iterative</li> </ul>	<ul style="list-style-type: none"> <li>• Recursive model (restrictive)</li> </ul>
		<ul style="list-style-type: none"> <li>• Regressional assumptions (in each equation, the error term is uncorrelated with the independent variables)</li> <li>• <math>\Rightarrow</math>(Least restrictive)</li> </ul>

Table 1. Theoretical comparison between Jöreskog’s formula (acronym Jöreskog), Finite Iterative Method (acronym FIM) and, our proposed method (acronym MFIM).

It should be mentioned that if the causal model is true at the population level, all the covariances computed by different methods should be equal. This is because the covariance matrix is unique for each model. These differences seen between the calculated covariance matrices are due to: a) the assumptions in which based each method and b) the constraints (which are correct in the true model at the population level) that are used in each computation. The second constraint is in terms of the variances of disturbances, we explain this constraint as follows:

From (8),  $\forall j \in \{1, \dots, p\}$  :

$$\begin{aligned} \eta_j &= \mathbf{A}_{1,1:q+p} \begin{pmatrix} \xi \\ \eta \end{pmatrix} + \zeta_j \\ &= \mathbf{A}_{1,1:q+j-1} (\xi^t \quad \eta_1 \quad \dots \quad \eta_{j-1}) + \zeta_j \end{aligned}$$

Therefore :

$$\sigma_{\eta_j}^2 = \mathbb{E}[\eta_j^2] = \mathbf{A}_{1,1:q+j-1} \widehat{\Sigma}_{1:q+j-1,1:q+j-1} \mathbf{A}_{1,1:q+j-1}^t + \psi_j$$

Which means :

$$\psi_j = \sigma_{\eta_j}^2 - \mathbb{E}[\eta_j^2] = \mathbf{A}_{1,1:q+j-1} \widehat{\Sigma}_{1:q+j-1,1:q+j-1} \mathbf{A}_{1,1:q+j-1}^t \tag{34}$$

In addition to that, for identification reasons, one may set free the variances of the latent variables and scale it by making one of its loadings constrained to one [15, 2]. This is possible when applying Jöreskog’s formula or Algorithm 2. However, since the variances of latent variables are set to one in the finite iterative method (Algorithm 1), we can not constrain the loadings to one.

### 5.2. Numerical comparison

In this section, we will compare all methods using a simulated example. We have generated, in three stages, 100 vectors of model parameters ( the vector of model parameters  $\theta$  contain all free parameters of the matrices  $B, \Gamma, \Lambda_x, \Lambda_y, \Psi, \Theta_\delta$  and,  $\Theta_\epsilon$  ). We calculate the implied covariance matrix via i) Jöreskog’s formula, ii) the finite iterative method and, iii) our proposed method. In the first stage, we generated the parameters such that the measurement errors are constrained by the equation (11). In the second stage, the disturbances are constrained by

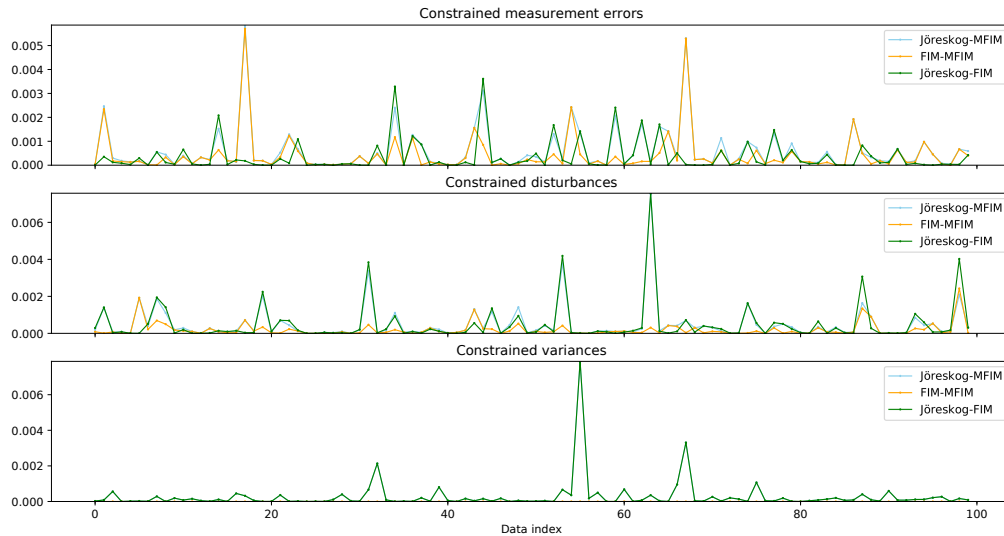


Figure 4. Theoretical comparison between Jöreskog’s formula (acronym Jöreskog), Finite Iterative Method (acronym FIM) and, our proposed method (acronym MFIM).

the formula (34). While the third simulation will make the variances constrained to one ( the second assumption in FIM method)). The adapted model is Figure 3. After that, we compute the distance between these matrices using the discrepancy function defined for given two squared matrices  $M_1$  and  $M_2$ :

$$F(M_1, M_2) = \frac{1}{2}Tr((M_1 - M_2)^t(M_1 - M_2))$$

Figure 4 shows the results of the simulation. Practically, the methods lead to closed matrices (the maximum difference obtain in 300 simulations is .0078). When the error terms (measurement errors and disturbances) are constrained the distance between the method became closer. This is because we impose the constraints to be valid (as in the true model at the population level as aforementioned). When the variances are constrained to one (the second assumption of FIM is valid), Our proposed method became practically identical to FIM (the maximum distance between them is bellow  $2 \times 10^{-32}$ ). In this third simulation, the curve of the difference between matrices computed by the finite iterative method and our proposed method (acronym FIM-MFIM) does not show because it’s near zero. Besides, the curve of the difference between the matrice computed by Jöreskog’s formula and our proposed method (acronym Jöreskog-MFIM) does not show because it’s approximately identical to the difference between Jöreskog’s formula and FIM (acronym Jöreskog-FIM).

### 6. Summary, discussion and perspectives

We presented SEM models, their fields of application, their notation, the fundamental hypothesis of SEM and, the typology of variables and models. We discussed the steps of SEM and focused on the estimation part. Especially the computation of the covariance matrix implied by an SEM model. Two major methods are exposed: the Jöreskog’s formula [10] and the Finite Iterative Method [7].

We showed that Jöreskog’ formula computes the implied covariance matrix by block, it requires the matrix  $(I - B)$  to be invertible instead of the recursiveness of the model. It assumes no correlation among all error terms and between exogenous variables and error terms. No need for the latent variables to be standardized.

We presented the Finite Iterative Method for SEM models. It is an extension of the work of [5, 6]. The method computes the implied covariance matrix through two steps: the first step is the computation of the implied correlation matrix of the structural model using Algorithm 1. Then, the second step is the computation of the covariance matrix of the model using (10). It assumes: all latent variables are standardized, noncorrelation among measurement errors and, each disturbance is uncorrelated with all independent variables containing this disturbance.

We proposed a modified algorithm (Algorithm 1) for the computation of the covariance matrix using FIM. The advantages of this algorithm are: i) it computes the desired matrix in one step without separating the model in the structural part and the measurement part and ii) it is less restrictive in terms of the assumptions compared to both methods. The only required assumptions are the regressional assumptions (in each equation, the error term is uncorrelated with the independent variables in that equation).

A Python program is written to implement all three methods as functions. In the case of our proposed method, we implemented it using the function  $f$  in Corollary 1. The program is available upon request.

In terms of perspectives, current research on the development of FIM for nonrecursive models is on. This work will focus on finding some characteristics of a nonrecursive model that make it possible to compute its implied covariance matrix. In addition to that, we think that FIM will serve to compute the implied covariance matrix for formative models (this idea is already described by [30]). This idea is under study and will introduce new researchers to adapt the covariance-based approach in their studies. Especially in the fields that use PLS Path Modeling (variance-based approach of SEM) due to its advantage in these kinds of models. The authors suggest, for instance, [32] and [31] where the differences between reflective and formative approaches are explained. The more recent advancement on this topic will be considered by the authors for the future development of formative blocks in FIM.

#### REFERENCES

1. H. Baumgartner and C. Homburg, *Applications of structural equation modeling in marketing and consumer research: A review*, International Journal of Research in Marketing, vol. 13, no. 2, pp. 139 – 161, 1996.
2. K. A. Bollen, *Structural equations with LV*, New York: Wiley, 1989.
3. Breckler, S. J., *Applications of correlation structure modeling in psychology: cause for concern?*, Psychol. Bull, vol. 107, pp. 260–273, 1990.
4. Bagozzi, R. P. and Yi, Y., *Specification, evaluation, and interpretation of structural equation models.*, Journal of the Academy of Marketing Science, vol. 40, pp. 8–34, 2012.
5. Elhadri, Z and Hanafi, M., *The Finite Iterative Method for calculating the correlation matrix implied by a recursive path model*, Electronic Journal of Applied Statistical Analysis, vol. 08, no. 01, pp. 84–99, 2015.
6. Elhadri, Z and Hanafi, M., *Extending the Finite Iterative Method for Computing the Covariance Matrix Implied by a Recursive Path Model*, 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, vol. 02, pp. 29–43, 2016.
7. Z. Elhadri and M. Hanafi, and P. Dolce and Y. Elkettani and M. Iaousse, *Computation of the Covariance Matrix Implied by a Structural Recursive Model with Latent Variables through the Finite Iterative Method*, In *The 5th Edition of the International Conference on Optimization and Applications (ICOA 2019)*, pp. 272–280, 2019.
8. Grace, J.B. and Anderson, T.M. and Olef, H. and Scheiner, S.M., *On the specification of structural equation models for ecological systems*, Ecological Monographs, vol. 80, no. 01, pp. 67–87, 2010.
9. Duncan, O. D., *Path analysis: Sociological examples*, American Journal of Sociology, vol. 72, pp. 1–16, 1966.
10. Jöreskog, K. G., *A general method for the analysis of covariance structures*, Biometrika, vol. 57, pp. 239–251, 1970.
11. Jöreskog, K. G., *Structural Analysis of Covariance and Correlation Matrices*, Psychometrika, vol. 43, no. 4, pp. 443–477, 1978.
12. Jöreskog, K. G., *Structural Equation Models in the Social Sciences: Specification, Estimation and Testing*, In *Applications of Statistics*, R. Krishnaiah (Ed). Amsterdam: North-Holland, pp. 265–287, 1977.
13. Jöreskog, K. G. and Sorbom, D., *LISREL 8: Structural equation modeling with the SIMPLIS command language.*, Mahwah, NJ: Lawrence Erlbaum Associates, 1993.
14. Jöreskog, K.G. and Wold, H., *Systems under indirect observation: Causality, structure, prediction, Part I and Part II*, Amsterdam: North Holland, 1982.
15. Jöreskog, K.G. and Olsson, U.H. and Wallentin, F.Y., *Multivariate Analysis with LISREL*, Springer Publishing Company, 2016.
16. Lawley, D. N., *The estimation of factor loadings by the method of maximum likelihood*, In *Proceedings of the Royal Society of Edinburgh*, vol. 60, pp. 64–82, 1940.
17. Pugesek, B.H., *Structural Equation Modeling Applications in Ecological and Evolutionary Biology*, New York : Cambridge University Press, 2003.
18. Wright, S., *Correlation and causation*, Journal of Agricultural Research, vol. 20, no. 17, pp. 557–585, 1921.
19. Wright, S., *The Theory of Path Coefficients- A reply to Niles' Criticism*, Genetics, vol. 08, no. 03, pp. 239–255, 1923.

20. Wright, S., *The Method of Path Coefficients*, Annals of Mathematical Statistics, vol. 05, no. 03, pp. 161–215, 1934.
21. Wright, S., *Path Coefficients and Path Regressions. Alternative or Complementarity Concepts?*, Biometrics, vol. 16, no. 02, pp. 189–202, 1960.
22. Kenny, D. A., *Correlation and Causality*, Wiley, New York, 1979.
23. M. Iamousse, A. Hmimou, Z. El hadri and Y. El kettani, *On the Computation of the Correlation Matrix Implied by a Recursive Path Model*, In: 2020 IEEE 6th International Conference on Optimization and Applications (ICOA), Beni Mellal, Morocco, 2020, pp. 1-5.
24. Hauser, R. M. et al., *Education, occupation, and earnings*, New York: Academic Press, 1975.
25. Spearman, C., *General intelligence objectively determined and measured*, American Journal of Psychology, vol. 15, pp. 201–293, 1904.
26. Mardani, Abbas and Streimikiene, Dalia and Zavadskas, Edmundas and Cavallaro, Fausto and Nilashi, Mehrbakhsh and Jusoh, Ahmad and Zare, Habib, *Application of Structural Equation Modeling (SEM) to Solve Environmental Sustainability Problems: A Comprehensive Review and Meta-Analysis*, Sustainability, vol. 9, no. 10, pp. 1814, 2017.
27. Weston, Rebecca and Gore, Paul A., *A Brief Guide to Structural Equation Modeling*, The Counseling Psychologist, vol. 34, no. 5, pp. 719–751, 2006.
28. Wold, Herman, *Model Construction and Evaluation When Theoretical Knowledge Is Scarce*, In Evaluation of Econometric Models, Elsevier, pp. 47–74, 1980.
29. Mulaik, Stanley A., *Linear causal modeling with structural equations*, CRC Press, 2009.
30. Dolce, Pasquale and Lauro, Natale Carlo, *Comparing maximum likelihood and PLS estimates for structural equation modeling with formative blocks*, Quality & Quantity, vol. 49, no. 3, pp. 891–902, 2015.
31. Dolce, P., Esposito Vinzi, V., and Lauro C., *Non-Symmetrical Composite-based Path Modeling*, Advances in Data Analysis and Classification, vol. 12, no. 3, pp. 759–784, 2018.
32. Henseler, J., *Bridging Design and Behavioral Research With Variance-Based Structural Equation Modeling*, Journal of Advertising, vol. 46, no. 1 pp. 178–192, 2017.
33. Tenenhaus M., Esposito VV., Chatelin YM., Lauro C., *PLS path modeling*, Computational Statistics & Data Analysis, vol. 48, no. 1, pp. 159–205, 2005.