ON THE METHOD FOR FINDING INVARIANTS OF FACTOR ANALYSIS MODELS

S. Stafeev
Institute of Applied Mathematical Research
Petrozavodsk, RUSSIA
e-mail: staf@karelia.ru

Abstract

The method for finding polynomial invariants of factor analysis model with correlated residuals is developed.

1 Model

Consider the following model:

$$\mathbf{X} = \mathbf{AH} + \mathbf{Y}, \tag{1}$$

where $\mathbf{X} = \{X_1, ..., X_n\}^t$ — is a vector with covariance matrix $\Sigma = (\sigma_{ij})$; $\mathbf{H} = \{H_1, ..., H_k\}^t$ is a set of independent normally distributed latent (hidden) variables (factors) with $\mathbf{M}(H_j) = 0$, $\mathbf{Var}(H_j) = 1$, $j = 1, ..., k$; $A = (a_{ij})$ is a factor loading matrix; A vector of residuals $\mathbf{Y} = \{Y_1, ..., Y_n\}^t$ has a nonsingular normal distribution with zero mean vector and a covariance matrix $\Theta_G = (\theta_{ij})$; $\mathbf{Y}$ and $\mathbf{H}$ are independent. Relationships among components of the vector $\mathbf{Y}$ are represented by a covariance graphical model with a structure $G = (V, E)$, where $V = \{1, .., n\}$ and $(i, j) \notin E$ if $\theta_{ij} = 0$.

Usually statistical inference in factor analysis models based on parametric representation and on maximum likelihood estimates [1]. But such approach is not always applicable. For example in case when a sample covariance matrix is singular (when the sample size is smaller then the number of variables n it is always so) the likelihood ratio test cannot be used. One of the approaches for overcoming these difficulties consists in using model invariants, that is, polynomial equality relations that the model imposes on the entries of the covariance matrix of the observed variables. The main problem is that for $k > 2$ finding polynomial invariants is very difficult from the computing point of view [2, 3]. In this paper a simple method for finding model invariants for any k is developed.

2 Polynomial invariants

It is easy to see that covariance matrix of observable variables Σ admits the following decomposition:

$$\Sigma = AA^t + \Theta_G, \tag{2}$$

Let P_G be a set of all positive definite $n \times n$ matrixes admitting decomposition (2) and let $f(\Sigma)$ be polynomial in the entries σ_{ij} of the matrix Σ. The polynomial f is

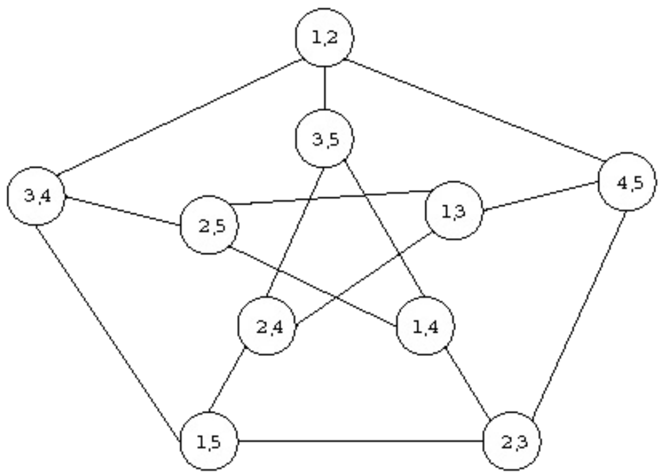

Figure 1: A graph $\mathbf{G}$ for $k = 2$ and $n = 5$.

called an invariant of the model (1) if $f(\Sigma) = 0$ for all $\Sigma \in P_G$. A set of all polynomial invariants forms an ideal $I^G_{k,n}$ in the ring of polynomials $\mathbb{R}[\sigma_{ij}, 1 \leqslant i \leqslant j \leqslant n]$.

The best known invariants are tetrads and pentads which arise in one and two-factor models respectively [1]. For example, if $k = 1$ and $n = 4$ an invariant has the form:

$$\sigma_{12}\sigma_{34} - \sigma_{23}\sigma_{14} = 0.$$

For the case when $k = 2$ and $n = 5$ an invariant has the form:

$$\begin{array}{lll}
\sigma_{12}\sigma_{23}\sigma_{34}\sigma_{45}\sigma_{51}- & \sigma_{12}\sigma_{23}\sigma_{35}\sigma_{14}\sigma_{45}- & \sigma_{12}\sigma_{24}\sigma_{35}\sigma_{34}\sigma_{15}+ \\
\sigma_{12}\sigma_{24}\sigma_{13}\sigma_{45}\sigma_{35}+ & \sigma_{12}\sigma_{25}\sigma_{34}\sigma_{14}\sigma_{35}- & \sigma_{12}\sigma_{25}\sigma_{13}\sigma_{34}\sigma_{45}- \\
\sigma_{13}\sigma_{24}\sigma_{35}\sigma_{14}\sigma_{25}+ & \sigma_{13}\sigma_{25}\sigma_{34}\sigma_{24}\sigma_{15}+ & \sigma_{14}\sigma_{23}\sigma_{13}\sigma_{45}\sigma_{25}- \\
\sigma_{14}\sigma_{25}\sigma_{23}\sigma_{34}\sigma_{15}- & \sigma_{15}\sigma_{23}\sigma_{13}\sigma_{24}\sigma_{45}+ & \sigma_{15}\sigma_{24}\sigma_{23}\sigma_{14}\sigma_{35} = 0
\end{array}$$

Define the basic object of this work. Let $\overline{G} = (V, \overline{E})$ be a complementary graph for G. Define a graph $\mathbf{G} = (\mathbf{V}, \mathbf{E})$, where $\mathbf{V} = \{\mathbf{i} = (i_1, ..., i_k), 1 \leqslant i_1 < i_2 < ... < i_k \leqslant k\}$, $\mathbf{E} = \{(\mathbf{i}, \mathbf{j}) | \mathbf{i} = (i_1, ..., i_k), \mathbf{j} = (j_1, ..., j_k), (i_s, j_l) \in \overline{E}, s, l = 1, ...k\}$. Note that for the first time the graph $\mathbf{G}$ was determined in work [5] in context of a model identifiability. Denote by $|\Sigma_{\mathbf{i},\mathbf{j}}|$ a determinant of the matrix $\Sigma_{\mathbf{i},\mathbf{j}} = (\sigma_{i_m, j_l})^k_{m,l=1}$. Let $C = (V_C, E_C)$, where $V_C = \{\alpha_i, i = 1, ..., m\}$, $E_C = \{(\alpha_i, \alpha_{i+1}), i = 1, ..., m-1, (\alpha_1, \alpha_m)\}$ be a even simple cycle. Define the following polynomial:

$$f_C(\Sigma) = \prod_{(\mathbf{i},\mathbf{j}) \in E'_C} |\Sigma_{\mathbf{i},\mathbf{j}}| - \prod_{(\mathbf{i},\mathbf{j}) \in E''_C} |\Sigma_{\mathbf{i},\mathbf{j}}|,$$

where $E'_C = \{(\alpha_{2j+1}, \alpha_{2j+2}), j = 0, ..., m/2 - 1\}$ and $E''_C = E_C \setminus E'_C$.

Let $T = (V_T, E_T)$ be a graph that consists of a simple odd cycle with a set of edges $E_1 = \{(\alpha_1, \alpha_2), ..., (\alpha_{m_1-1}, \alpha_{m_1}), (\alpha_1, \alpha_{m_1})\}$, a simple chain with a set of edges $E_2 = \{(\alpha_{m_1}, \alpha_{m_1+1}), ..., (\alpha_{m_2-1}, \alpha_{m_2})\}$ and a simple odd cycle with a set of edges $E_3 = \{(\alpha_{m_2}, \alpha_{m_2+1}), ..., (\alpha_{m-1}, \alpha_m), (\alpha_m, \alpha_{m_2})\}$.

Define the following polynomial:

$$f_T(\Sigma) = \prod_{(\mathbf{i},\mathbf{j}) \in E'_T \cap (E_1 \cup E_3)} |\Sigma_{\mathbf{i},\mathbf{j}}| \prod_{(\mathbf{i},\mathbf{j}) \in E'_T \cap E_2} |\Sigma_{\mathbf{i},\mathbf{j}}|^2 - \prod_{(\mathbf{i},\mathbf{j}) \in E''_T \cap (E_1 \cup E_3)} |\Sigma_{\mathbf{i},\mathbf{j}}| \prod_{(\mathbf{i},\mathbf{j}) \in E''_T \cap E_2} |\Sigma_{\mathbf{i},\mathbf{j}}|^2,$$

where $E'_T = \{(\alpha_{2j+1}, \alpha_{2j+2}), j = 0, ..., [m/2] - 1\}$ and $E''_T = E_T \setminus E'_T$.

The following theorem describes a large class of polynomial invariants.

Theorem 1. *Let* $\mathbf{C}(\mathbf{G})$ *be a set of all even simple cycles of the graph* $\mathbf{G}$ *and let* $\mathbf{T}(\mathbf{G})$ *be a set of all subgraphs of the graph* $\mathbf{G}$ *that consist of two simple odd cycles connected by a simple chain. Then*

$$\{f_C(\Sigma), C \in \mathbf{C}(\mathbf{G}), f_T(\Sigma), T \in \mathbf{T}(\mathbf{G})\} \subset I^G_{k,n}.$$

For example, in the case $k = 2$ and $n = 5$ a polynomial $f_C(\Sigma)$ has the form:

$$f^{2,5}_C = \begin{vmatrix} \sigma_{12} & \sigma_{32} \\ \sigma_{14} & \sigma_{34} \end{vmatrix} \begin{vmatrix} \sigma_{12} & \sigma_{52} \\ \sigma_{13} & \sigma_{53} \end{vmatrix} \begin{vmatrix} \sigma_{12} & \sigma_{42} \\ \sigma_{15} & \sigma_{45} \end{vmatrix} - \begin{vmatrix} \sigma_{12} & \sigma_{52} \\ \sigma_{14} & \sigma_{54} \end{vmatrix} \begin{vmatrix} \sigma_{12} & \sigma_{42} \\ \sigma_{13} & \sigma_{43} \end{vmatrix} \begin{vmatrix} \sigma_{12} & \sigma_{32} \\ \sigma_{15} & \sigma_{35} \end{vmatrix}.$$

Unfortunately but the invariant $f^{2,5}_C$ is reducible. That is $f^{2,5}_C$ is the product σ_{12} and pentad. It is true the following general proposition.

Theorem 2. *Let* $\mathbf{m} = \{m_1, ..., m_{k-1}\}$ *and* $\mathbf{d} = \{d_1, ..., d_{k-1}\}$, $\mathbf{m} \cap d = \emptyset$, $i_1 \neq i_2 \neq i_3$, $\{i_1, i_2, i_3\} \notin \mathbf{m} \cup \mathbf{d}$. *If the graph* $\mathbf{G}$ *contains a cycle with a set of edges*

$$\{(< i_1, \mathbf{m} >, < i_2, \mathbf{d} >), (< i_2, \mathbf{d} >, < i_3, \mathbf{m} >), (< i_3, \mathbf{m} >, < i_1, \mathbf{d} >),$$
$$(< i_1, \mathbf{d} >, < i_2, \mathbf{m} >), (< i_2, \mathbf{m} >, < i_3, \mathbf{d} >), (< i_3, \mathbf{d} >, < i_1, \mathbf{m} >)\}.$$

Then

$$|\Sigma_{\langle \mathbf{d}i_1 \rangle, \langle \mathbf{m}i_2 \rangle}||\Sigma_{\langle \mathbf{d}i_3 \rangle, \langle \mathbf{m}i_1 \rangle}||\Sigma_{\langle \mathbf{d}i_2 \rangle, \langle \mathbf{m}i_3 \rangle}| - |\Sigma_{\langle \mathbf{d}i_3 \rangle, \langle \mathbf{m}i_2 \rangle}||\Sigma_{\langle \mathbf{d}i_2 \rangle, \langle \mathbf{m}i_1 \rangle}||\Sigma_{\langle \mathbf{d}i_1 \rangle, \langle \mathbf{m}i_3 \rangle}| =$$
$$= |\Sigma_{\mathbf{d},\mathbf{m}}| \times \textit{irreducible polynomial invariant.}$$

Let S be a sample covariance matrix. For testing the null hypothesis H_f: $f(\Sigma) = 0$ we can use statistics $f(S)$ [4]. Denote by $Var_\Sigma(f(S))$ the variance of $f(S)$. $Var_\Sigma(f(S))$ is a polynomial function of the covariance matrix Σ. Replacing Σ by the sample covariance matrix S in this polynomial yields the estimator $Var_S(f(S))$. In conditions of the null hypothesis standardized statistics $\frac{f(S)}{\sqrt{Var_S(f(S))}}$ has an asymptotic standard normal distribution.

With the help of the graph $\mathbf{G}$ it is possible to formulate sufficient conditions of the model (1) identifiability.

Sufficient conditions of model identifiability. Let the graph $\mathbf{G}$ be connected and contains a simple odd cycle with a set of edges E, where $|\Sigma_{\mathbf{i},\mathbf{j}}| \neq 0$, $(\mathbf{i}, \mathbf{j}) \in E$. Then the model (1) is identifiable.

References

[1] Harman H. (1976) *Modern Factor Analysis.* University of Chicago Press, third edition.

[2] Drton M., Sturmfels B., Sullivant S. (2007) Algebraic factor analysis: tetrads, pentads and beyond. *Probability Theory and Related Fields*. Vol. **138**, pp. 463-493.

[3] Drton M., Sullivant S. (2007) Algebraic statistical models. *Statistica Sinica*. Vol. **17**, pp. 1273-1297.

[4] Drton M., Massam H., Olkin I. (2008) Moments of minors of Wishart matrices. *Annals of Statistics*. Vol. **36**, pp. 2261-2283.

[5] Stafeev S.V. (2007) On the parameter estimation of recursive bow – free models with latent variables. *Proceedings of the Eighth International Minsk Conference Computer Data Analysis and Modeling: Complex Stochastic Data and Systems.* Minsk. - Vol. **1**, pp. 178-181.