Journal of Multivariate Analysis 144 (2016) 110–128

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

# Bayesian factor analysis with uncertain functional constraints about factor loadings

Hea-Jung Kim [a], Taeryon Choi [b,*], Seongil Jo [b]

[a] Department of Statistics, Dongguk University, Seoul, Republic of Korea

[b] Department of Statistics, Korea University, Republic of Korea

ARTICLE INFO

*Article history:*
Received 9 October 2014
Available online 2 December 2015

*AMS subject classifications:*
62F15
62H25
62H12

*Keywords:*
Factor loading
Hierarchical Bayesian model
Markov chain Monte Carlo
Rectangular screened scale mixtures
Uncertain constraint

ABSTRACT

Factor analysis with uncertain functional constraints about factor loading matrix is considered from a Bayesian viewpoint, in which the uncertain prior information is incorporated in the analysis. We propose a hierarchical screened scale mixture of normal factor (*HSMF*) model for flexible inference of the constrained factor loadings, factor scores, and specific variances as well as the covariance matrix of the factors. The proposed model makes provisions for robust factor analysis with uncertainty about the functional constraints. A number of inferential aspects of the proposed model are investigated in order to render the proposed analysis optimal. These include the closure properties of a class of rectangle-screened scale mixture of multivariate normal (*RSMN*) distributions which is useful for statistical inference of the *HSMF* model, eliciting the prior and posterior evolutions of the uncertainly constrained factor loadings, and providing the efficient Bayesian estimation procedure by using the MCMC methods. Empirical analysis for Bayesian factor models with synthetic data and real data applications is given to illustrate the usefulness of the proposed model.


## 1. Introduction

Factor analysis is one of the most commonly used statistical techniques for formulating theories in the multivariate behavioral and social sciences [4]. It is a casual modeling technique that attempts to explain dependence among multivariate observations through covariance relationships in terms of a smaller number of latent random factors [26].

In matrix form, the observation model for the basic normal factor analysis is given by

$$\mathbf{x}_j - \boldsymbol{\mu}_{\mathbf{x}} = \Lambda \mathbf{f}_j + \boldsymbol{\varepsilon}_j, \quad j = 1, \dots, n, \tag{1.1}$$

where $\mathbf{x}_j$ is a $p \times 1$ random observation vector with mean $\boldsymbol{\mu}_{\mathbf{x}}$ and covariance matrix $\Sigma_{\mathbf{x}}$, $\Lambda = \{\lambda_{ik}\}$ is a $p \times m$ matrix of factor loadings with rank $m \le p$, $\mathbf{f}_j \overset{\text{i.i.d.}}{\sim} \mathcal{N}_m(\mathbf{0}, \Phi)$ is an $m \times 1$ vector of latent random factors, and $\boldsymbol{\varepsilon}_j \overset{\text{i.i.d.}}{\sim} \mathcal{N}_p(\mathbf{0}, \Psi)$ is a vector of noise terms, independent of $\mathbf{f}_j$, where $\Psi$ is a diagonal matrix. The basic model in (1.1) has two specific features to be considered for statistical inference, one about the restriction of the factor loadings matrix $\Lambda$ and the other relating to the distributional assumption of the noise terms $\boldsymbol{\varepsilon}_j$. Restriction of the factor loadings matrix $\Lambda$ needs to be made to define a unique model free from identifiability and to allow useful interpretation of the model, while the normal noise terms are exploited to deal with the normal linear factor model.

* Corresponding author.
*E-mail address:* trchoi@gmail.com (T. Choi).

http://dx.doi.org/10.1016/j.jmva.2015.11.006
0047-259X/

Indeed, identifiability of the factor model in (1.1) may be assessed by comparing the number of parameters in $\Lambda$, $\Psi$, and $\Phi$ with the $p(p+1)/2$ elements that are contained in the empirical covariance matrix. In line with a confirmatory approach, fixed values (zero or one) may be preset for some of the loadings in $\Lambda$ [30]. Specifically, since the model in (1.1) is invariant under transformation of the form $\Lambda^* = \Lambda\mathbf{P}^\top$ and $\mathbf{f}^* = \mathbf{Pf}$, where $\mathbf{P}$ is any orthogonal $k \times k$ matrix, the approaches by Geweke and Zhou [8] and Aguilar and West [1] and Lopes and West [20], for example, were to assume that $\Lambda$ is a block lower triangular matrix, assumed to be of full rank with diagonal elements strictly positive, which turns out to be useful for both identification and interpretation.

Modification of the distributional assumptions of the normal linear factor model (1.1) was made for practical applications. For example, [10], and [23] considered factor analysis using dynamic factor models wherein $\mathbf{f}_j$'s are dependent; Wedel et al. [32] and Wedel and Kamakura [33] also noted the assumption that $\mathbf{f}_j$ are normal may need to be modified for certain applications; Fokoué [6] and McLachlan et al. [21] considered a finite mixture of factor models to obtain a flexible factor analysis model. Copious references to the literature on the application and analysis of the Bayesian factor model in (1.1) as well as its more elaborated models can be found in [25,26,5].

The present paper, however, considers yet another model in which prior knowledge about the factor loadings is available in the form of uncertain multivariate inequality constraints. Specifically, suppose that the prior knowledge about $\Lambda$ in (1.1) is uncertain and likely to have the following functional constraint:

$$\{\boldsymbol{\lambda}; L(\boldsymbol{\lambda}) \in \mathbf{C}_q(\boldsymbol{\alpha}, \boldsymbol{\beta})\}, \tag{1.2}$$

where $\boldsymbol{\lambda} = \mathrm{Vec}(\Lambda^\top)$, $L(\boldsymbol{\lambda})$ is a linear function of $\boldsymbol{\lambda}$, $\mathbf{C}_q(\boldsymbol{\alpha}, \boldsymbol{\beta})$ denotes a $q$-variate rectangle set, i.e., $L(\boldsymbol{\lambda}) \in \mathbf{C}_q(\boldsymbol{\alpha}, \boldsymbol{\beta}) \equiv \{(v_1, \ldots, v_q); \alpha_i < v_i < \beta_i,\ i = 1, \ldots, q,\ q \le mp\}$, and $(v_1, \ldots, v_q)^\top = L(\boldsymbol{\lambda})$. Here, if $D$ is a $p \times q$ matrix then by $\mathrm{Vec}(D)$ we mean the $pq \times 1$ vector formed by stacking the columns of $T$ under each other; that is if $D = (\boldsymbol{d}_1, \ldots, \boldsymbol{d}_q)$, where $\boldsymbol{d}_i$ is $p \times 1$ for $i = 1, \ldots, q$, then $\mathrm{Vec}(D) = (\boldsymbol{d}_1^\top, \ldots, \boldsymbol{d}_q^\top)^\top$. There could be at least three reasons for employing the functional constraint (1.2) for the factor model: (i) to provide a realistic model that explains the structure of empirical data, (ii) to achieve identifiability of the model, and (iii) to obtain more meaningful interpretations. Indeed, in either confirmatory or exploratory factor analysis, the location and scale of the latent factor often need to be constrained through the factor variances or loadings (e.g. [22,29,18]).

Thus, alternative methods for dealing with the constraints in Bayesian confirmatory factor analysis need to be developed, and indeed, such practical considerations and issues with identifiability constraints are the motivation for development of a more flexible method of constrained factor analysis, which is tackled in this paper. In particular, a Bayesian approach to flexibly incorporating the uncertain prior knowledge about the constraints on $\Lambda$ with possibly non-normal sample information is proposed, which is the main contribution of this paper to the literature on Bayesian factor analysis. Accordingly, we contrast the proposed model with other existing Bayesian factor models for dealing with uncertain constraints in the factor loading and incorporating the heavy-tailed distribution in terms of hierarchical scale mixture of multivariate normal distributions into the factor analysis. The remaining part of this paper is organized as follows. In Section 2, we give a basic description of a class of the *RSMN* (rectangle-screened scale mixture of multivariate normal) distributions and provide its useful properties which will be the base of the modeling and estimation of our constrained factor model. In Section 3, we specifically propose the *HSMF* (hierarchical screened scale mixture of normal factor) model based on the properties of the class of *RSMN* distributions and explore its theoretical properties for estimation by analytically deriving the posterior distribution and marginal distribution of the *HSMF* model. The *HSMF* model adopts a flexible two-stage prior to elicit the uncertainty of the functional constraint under non-normal and/or heavy-tailed prior information on $\boldsymbol{\lambda}$ in (1.2) (see, e.g. [24,14]). Further, it goes through robust factor modeling (see, e.g., [36]) subject to uncertainty in the parametric restriction to avoid anomalies generated from the non-normal sample information. Bayesian factor models with uncertain functional constraints are illustrated in the context of the *HSMF* model in Section 4 through empirical analysis. Finally, concluding remarks with a discussion are made in Section 5.

## 2. Preliminaries

Before presenting the *HSMF* model with the uncertain multivariate rectangle constraints of the factor loading coefficients $\boldsymbol{\lambda}$, a brief review of some of the important properties of the class of *RSMN* distributions is provided. The properties are useful for the specifications and estimation of the *HSMF* model.

Assume that the joint distribution of respective $q \times 1$ and $p \times 1$ vector variables $\mathbf{U}$ and $\mathbf{V}$ is $F \in \mathcal{F}$, where

$$\mathcal{F} = \left\{F : \mathcal{N}_s\left(\boldsymbol{\psi}, \kappa(\eta)\Gamma\right),\ \eta \sim G(\eta) \text{ with } \kappa(\eta) > 0, \text{ and } \eta > 0\right\}, \tag{2.1}$$

$s = (q+p)$, $\eta$ is a mixing variable with the cdf $G(\eta)$, $\kappa(\eta)$ is a suitably chosen weight function, and $\boldsymbol{\psi}$ and $\Gamma$ are partitioned as

$$\boldsymbol{\psi} = \begin{pmatrix} \boldsymbol{\psi}_{\mathbf{U}} \\ \boldsymbol{\psi}_{\mathbf{V}} \end{pmatrix} \quad \text{and} \quad \Gamma = \begin{pmatrix} \Gamma_{\mathbf{U}} & \Delta^\top \\ \Delta & \Gamma_{\mathbf{V}} \end{pmatrix}$$

corresponding to the orders of $\mathbf{U}$ and $\mathbf{V}$. Notice that $\mathcal{F}$, defined by (2.1), denotes a class of scale mixture of multivariate normal (*SMN*) distributions (see, e.g., [9] for details), equivalently denoted as $\mathcal{SMN}_s(\boldsymbol{\psi}, \Gamma, \kappa, G)$ in the remaining part of the paper.