Proceedings of the 90th Annual International Meeting of the Psychometric Society, Minneapolis, United States (2025).

Multidimensional constructs and moderated factor analysis

R. Noah Padgett *

Department of Epidemiology, Harvard T.H. Chan School of Public Health, Harvard University, Boston, MA, USA
*Corresponding author. Email: npadgett@hsph.harvard.edu

Abstract

Moderated nonlinear factor analysis (MNLFA) provides a flexible framework for evaluating measurement invariance and impact across multiple background characteristics simultaneously. However, current MNLFA methodology has been limited to models with at most two latent factors due to the complexity of ensuring positive definiteness of the latent variable covariance matrix when inducing moderation effects on factor correlations. This limitation has restricted the application of MNLFA to the many multidimensional constructs common in psychological and educational assessment. This paper addresses this dimensional constraint by adapting modified Cholesky decomposition approaches from the covariance prediction literature. Specifically, I present a partial-correlation parameterization that separates factor variances from factor correlations, allowing moderation of the Cholesky factors of the correlation matrix while maintaining positive definiteness. The proposed approach extends MNLFA to an arbitrary number of latent factors, enabling researchers to simultaneously investigate measurement non-invariance in item parameters (intercepts, loadings, residual variances) and impact on latent structure (factor means, variances, and correlations) for multidimensional constructs. A Bayesian estimation approach using spike-and-slab priors for regularization is implemented in Stan, with the marginal-likelihood formulation providing improved computational efficiency compared to conditional-likelihood approaches. Practical guidance is provided for prior specification and identification constraints with suggestions for future research on estimation performance. This methodological extension broadens the applicability of MNLFA to complex multidimensional measurement models commonly encountered in scale development and validation research.

Keywords: measurement invariance, differential item functioning, MNLFA, Bayesian, multidimensional, factor analysis

1. Introduction

A foundational component of nearly all scale development processes is investigating measurement invariance across subpopulations. Invariance of scale properties is sometimes referred to as fairness in an educational setting, as defined by the Standards (AERA, APA, NCME, 2024, p. 50) or to provide validity evidence for the consistency in interpretation of scores across groups of like respondents. Many lines of research have developed to empirically assess the validity evidence supporting (or against) the conclusion that scores are comparable across groups (van Erp et al., 2018; van de Vijver et al., 2019).

One way validity evidence can be provided is by evaluating the equivalence of sets of parameters in a confirmatory factor model across two or more groups (e.g., is the entire factor loading matrix equal for males and females? Or, are all item residual variances equal over time?). Researchers have provided some excellent guidance and recommendations for scale developers to use, building on this perspective (Vandenberg & Lance, 2000). However, these methods can be clunky when trying to

identify partial invariance, which is the likely scenario in which only a subset of parameters vary meaningfully across groups.

The degree of the non-invariance within subsets of parameters can be of substantive interest depending on the pattern of how parameters vary. For example, in the study of psychopathology, indicators of internalizing tend to cluster into a few types of behaviors which are treated as factors to study different indicators of psychopathology for different types of people. The pattern of how indicators of psychopathology vary across groups (e.g., sex and age) has a meaningful interpretation for the development of targeted interventions. A limitation of the traditional factor analytic approach is the difficulty in identifying these patterns due to the invariance being studied from a set of parameters perspective.

An alternative perspective, often called differential item functioning, focuses on individual indicators (or items) instead of groups of parameters and has been a useful approach for identifying potentially variant indicators across groups (Clauser & Mazor, 1998). Similarities with the multiple group factor analytic approach are discussed at length elsewhere (Thissen, 2025). A focus on indicators has been the predominant perspective in educational testing, aiming to identify and remove items that may be biased against specific student subgroups. However, in psychological assessment, the degree of variance in parameters across groups can be of direct interest, as discussed in the study of psychopathology above. A general approach that has gained attention in recent years for the study of invariance of model parameters across groups is known as moderated nonlinear factor analysis, which is a flexible framework for incorporating a design-like matrix of group/person characteristics to be predictors of factor analysis model parameters (Bauer, 2023).

Moderated nonlinear factor analysis has been applied in many settings to identify variance in measurement model characteristics across several background characteristics simultaneously (Bauer, 2023). Other applications sometimes aim to provide scores on factors that are adjusted for possible partial invariance (Curran et al., 2014). The development of these methods has expanded into evaluating measurement invariance in up to two constructs (Bauer, 2023). Such moderated nonlinear factor analysis models allow for the simultaneous investigation of differences in latent factor parameters alongside measurement model characteristics. Moderation effects on latent factors are sometimes referred to as the *impact* of background characteristics on the construct of interest. Investigating impact, while controlling for possible non-invariance of a small number of measurement model parameters, is of primary substantive interest when applying moderated nonlinear factor analysis.

The applications of moderated nonlinear factor analysis have been limited thus far to only two latent factors due to the inherent complexity of ensuring the latent variable covariance matrix is valid for all subgroups implied by the design matrix of moderating factors. The current approach for two-factor models is to induce on an unbounded continuum and then use Fisher's z-transformation to restrict the parameter to the [-1,1] bound of correlations. Extending to more than two latent factors cannot utilize the same method using independent unbounded scores transformed to the [-1,1] scale because independent applications of Fisher's z-transformations do not ensure the entire correlation matrix is valid. Any methods for inducing impact on individual correlations must simultaneously ensure each correlation is between [-1,1] and that the matrix of correlations is positive definite. Achieving these simultaneous constraints has been a longstanding problem in the simulation of correlation matrices (Archakov & Hansen, 2021; Joe, 2006; Pourahmadi & Wang, 2015) and in covariance prediction (Bucci et al., 2022; Pinheiro & Bates, 1996; Pourahmadi, 1999; Pourahmadi, 2011). Modified conceptualizations of these methods are used in this article to address the dimensional limitation of moderated nonlinear factor analysis. I generalize moderated nonlinear factor analysis to flexibly allow moderation in the latent factor correlation parameters in models with three or more factors.

The remainder of this article is organized as follows. Next, I describe the MNLFA model in more detail for one- and two-dimensional factor models. Then, I described how MNLFA can be

extended to m-dimensional factor models, which is the primary contribution of this article. I provide examples of how the general specification can induce a variety of correlational structures for specific subgroups. I then report on an application of the described multidimensional moderated nonlinear factor model for the study of psychopathology. I end with a discussion of the model and potential limitations and areas of future study.

2. Moderated Linear Factor Analysis

I focus this article on latent variable models with continuous indicators, but these methods generalize to binary and Likert-type responses with some modifications. To begin, consider the measurement model for items $i = 1, \ldots, I$ with observed factor indicators y. The single factor latent variable η for $n = 1, \ldots, N$ cases can be described as

$$\begin{aligned}
y_{ni} &= \nu_{ni} + \lambda_{ni}\eta_n + \varepsilon_{ni} \\
\nu_{ni} &= \nu_{0i} + \mathbf{x}_n^t \boldsymbol{\delta}_{\nu i} \\
\lambda_{ni} &= \lambda_{0i} + \mathbf{x}_n^t \boldsymbol{\delta}_{\lambda i} \\
\varepsilon_{ni}|\mathbf{x}_n &\sim N\left(0.0, \theta_{ni}\right) \\
\theta_{ni} &= \theta_{0i} \exp(\mathbf{x}_n^t \boldsymbol{\delta}_{\theta i}) \\
\eta_n|\mathbf{x}_n &\sim N\left(\alpha_n, \phi_n\right) \\
\alpha_n &= \alpha_0 + \mathbf{x}_n^t \boldsymbol{\delta}_{\alpha} \\
\phi_n &= \phi_0 \exp\left(\mathbf{x}_n^t \boldsymbol{\delta}_{\phi}\right).
\end{aligned}$$

A full list of the meaning of all parameters is provided in the appendix. The major difference between the above model and a traditional single-factor model is the inclusion of moderation effects on model parameters (e.g, $\nu_{ni} = \nu_{0i} + \mathbf{x}_n^t \boldsymbol{\delta}_{\nu i}$ is the moderation on item intercepts). One way to conceptualize this approach to parameterizing the factor model is to think of hierarchical linear models where the intercepts and slopes (factor loadings) are outcomes of interest (see Raudenbush & Bryk, 2002, p. 80). Shifting to a hierarchical/multi-level perspective for these models makes a direct connection between testing for non-invariance across a wide range of possible variables with well-known methods of linear modeling.

A major benefit of connecting evaluating non-invariance to linear modeling is the flexibility that allows for unique testing of potentially nonlinear relationships between person factors and non-invariance in parameters. For example, Molenaar (2021) illustrated how generalized additive models could be used to evaluate non-invariance. This flexibility opens the door to evaluating even more complex interrelationships between model parameters and person characteristics through interactions among characteristics. Along with the benefits, evaluating non-invariance now shares many of the same concerns that arise with respect to collinearity in the design matrix for the predictors and overall model complexity. I will return to discuss collinearity and model complexity when I discuss priors later.

The one-dimensional factor model described above is a valuable baseline for extending to more complex factor structures. The simplest extension is to a two-factor model with simple structure. In such a two-factor model, the only new wrinkle lies in evaluating impact on the factor correlation. Bauer (2017) suggested the use of the Fisher's z-transformation of a unconstrained parameter to handle the bounding constraint of correlations after the factor variances have been separated out.

$$\begin{aligned}
&\begin{bmatrix} \eta_{ni} \\ \eta_{n2} \end{bmatrix} \sim \mathrm{MVN}\left(\begin{bmatrix} \alpha_{n1} \\ \alpha_{n2} \end{bmatrix}, \begin{bmatrix} \phi_{n11} & \phi_{n12} \\ \phi_{n12} & \phi_{n22} \end{bmatrix} \right) \\
&\phi_{n11} = \phi_{011} \exp\left(\mathbf{x}_n^t \boldsymbol{\delta}_{\phi 1}\right)
\end{aligned}$$

$$
\begin{aligned}
\phi_{n22} &= \phi_{022} \exp\left(\mathbf{x}_n^t \boldsymbol{\delta}_{\phi 2}\right) \\
\rho_{n12} &= 1 - \frac{2}{\exp\left[2\left(\gamma_{012} + \mathbf{x}_n^t \boldsymbol{\delta}_{\gamma 21}\right)\right] + 1} \\
\phi_{n12} &= \rho_{n12} \sqrt{\phi_{n11}} \sqrt{\phi_{n22}}
\end{aligned} \tag{1}
$$

Fisher's z-transformation (Eq. 1) provides a straightforward method for inducing varying factor correlations separately from differences in factor variances. However, to extend the model to three or more dimensions in the latent factor structure, Fisher's z-transformation approach does not work to jointly ensure the positive definiteness of the latent variable covariance matrix. Next, I introduce an approach to ensuring the positive definiteness that could be used in m-dimensional latent variable space.

2.1 Multiple latent factors

Extending beyond two latent factors effectively is a complex balancing act of *where* in the model do you allow moderation to be estimated. A fully saturated multidimensional model has moderation in all parameters–factor loadings, item intercepts, item residual variances, factor means, factor variances, and factor correlations. However, as I discuss in more detail in the *Estimation & Empirical Identification* section, the fully saturated model may not be estimable in all situations due to limitations of the data and model at hand. But first, I describe the fully saturated model and the process for inducing moderation in the factor correlation matrix to illustrate the potential for these models to explore a wide range of possible sources of moderation in different parts of the model. Moderation in the item parameters is relatively straightforward and can be shown as below:

$$
\begin{aligned}
\gamma_{ni} &= \nu_{ni} + \boldsymbol{\lambda}_{ni} \boldsymbol{\eta}_n + \varepsilon_{ni} \\
\nu_{ni} &= \nu_{0i} + \mathbf{x}_n^t \boldsymbol{\delta}_{\nu i} \\
\lambda_{nim} &= \lambda_{0im} + \mathbf{x}_n^t \boldsymbol{\delta}_{\lambda im} \\
\varepsilon_{ni} | \mathbf{x}_n &\sim N\left(0.0, \exp(\theta_{0i} + \mathbf{x}_n^t \boldsymbol{\delta}_{\theta i})\right) \\
\boldsymbol{\eta}_n | \mathbf{x}_n &\sim MVN\left(\boldsymbol{\alpha}_n, \boldsymbol{\Phi}_n\right) \\
\boldsymbol{\Phi}_n &= \left(\boldsymbol{\Phi}_n^D\right)^{0.5} \left(\boldsymbol{\Phi}_n^L\right) \left(\boldsymbol{\Phi}_n^L\right)^t \left(\boldsymbol{\Phi}_n^D\right)^{0.5}.
\end{aligned} \tag{2}
$$

The moderation in the item parameters is achieved through a model specification similar to that of the single or two-factor model. A new wrinkle involves specifying how moderation is defined in the factor-loading matrix to account for potential cross-loadings. Simple structure can be assumed where each item loads onto only one factor, or any number of cross-loadings can be specified. The simplest solution is to fix all non-theorized cross-loadings to zero. A more general solution is to freely estimate a subset of loadings which I believe primarily load on different factors, and then use regularizing priors on cross-loadings. The complexity of extending to greater than two factors arises with the person-specific factor covariance matrices ($\boldsymbol{\Phi}_n$).

Inducing moderation in the person-specific factor covariance matrices is the core limitation of multidimensional moderated nonlinear factor analysis. This problem can be restated as the more general issue of predicting several covariances simultaneously (Anderson, 1973; Pourahmadi, 1999). One such promising approach to inducing moderation in covariance matrices is based on a modified Cholesky decomposition that separates the variances and correlations (Pinheiro & Bates, 1996).

Decomposing the covariance structure involves first separating the factor variances and the factor correlations. The factor variances are separated out into a diagonal matrix of the variances ($\boldsymbol{\Phi}_n^D$). The factor correlations are parameterized as a lower-triangular matrix of Cholesky factors of the correlation matrix ($\boldsymbol{\Phi}_n^L$). Together, the full covariance matrix is reconstructed using equation 2. The

value of this decomposition lies in separating the factor variances from the Cholesky decomposition of the factor correlation matrix. This serves to easily induce impact on factor variances using the exponential function to maintain positive variances and to isolate the problem of inducing impact on factor correlations restricted to the Cholesky factors.

Cholesky factors serve as a useful intermediary step between an unconstrained space and a valid correlation matrix. A valid Cholesky factor of a correlation matrix can be constructed using a partial correlation parameterization (Joe, 2006; Lewandowski et al., 2009). First, a vector of unconstrained parameters, let's call this vector $\boldsymbol{\gamma}$, has support of any real number. This vector, $\boldsymbol{\gamma}$, is then mapped to the constrained space of "partial correlation" [-1,1] using Fisher's z-transformation, then a Cholesky factor is built assuming the vector is comprised of partial correlations. Moderating is induced in the first step on the unconstrained space, then the vector of moderated unconstrained parameters are mapped to partial correlations then the Cholesky factor is built. This mapping from the unconstrained scale to the Cholesky factor is shown below:

$$\boldsymbol{\gamma} = [\gamma_{12}, \gamma_{13}, ..., \gamma_{ij}]$$

$$\gamma_{ij} = \gamma_{0ij} + \mathbf{x}^t \boldsymbol{\delta}_{\gamma mm'}$$

$$\gamma^*_{ij} = 1 - \frac{2}{\exp\left[2(\gamma_{0ij} + \mathbf{x}^t \boldsymbol{\delta}_{\gamma ij})\right] + 1}$$

$$\phi^L_{ij} = \begin{cases} 0 & \text{if } j > i \\ 1 & \text{if } 1 = j = i \\ \gamma^*_{ij} & \text{if, } 1 = j < i \\ \prod_{k=1}^{j-1}\left(1 - \left(\gamma^*_{ik}\right)^2\right)^{0.5} & \text{if, } 1 < j = i \\ \gamma^*_{ij}\prod_{k=1}^{j-1}\left(1 - \left(\gamma^*_{ik}\right)^2\right)^{0.5} & \text{if, } 1 < j < i \end{cases}$$

The interpretation of each "partial" correlation (γ^*_{ij}) depends on the column and row of the Cholesky factor it maps to. For instance, for $i = 1, j = 2$, γ^*_{ij} is the correlation between factor 1 and 2, but element $i = 3, j = 2$ is the partial correlation between factors 2 and 3 after partialing out factor 1. The resulting matrix $\boldsymbol{\Phi}^L$ can then be constructed for each person n in a flexible way to induce different correlation structures depending on the design matrix ($\mathbf{x}^t_n$). The above transformation is illustrated in a simple-to-use R function in the Appendix.

Putting the pieces of the multidimensional MNLFA model results in the model illustrated in Figure 1. The diagram illustrates the nested hierarchical nature of the model and helps break down where measurement non-invariance and impact are modeled.

2.1.1 Factor order matters

A potentially major confounding factor in the application of the proposed method for inducing moderated correlations is that the order in which factors are entered into the model determines which factors are partial out. This means that impact of a moderating variable on the correlation between factors 2 and 3 depends on any moderation in the heading correlations between factors 1 and 2 and factors 1 and 3. Factor 1 being "primary" means the subsequent correlation moderation depends on the adequacy of factor one being identified. More work will need to be done to unpack how severe the impact of factor ordering has on identification of moderation. But, as an initial proposal, I would recommend ordering of the factors in increasing order of the degree of suspectibility to inter-factor correlation moderation. This means that we place the more "stable" factors first. Then, these more stable factors are partialed out when examining moderation in the residual correlation between

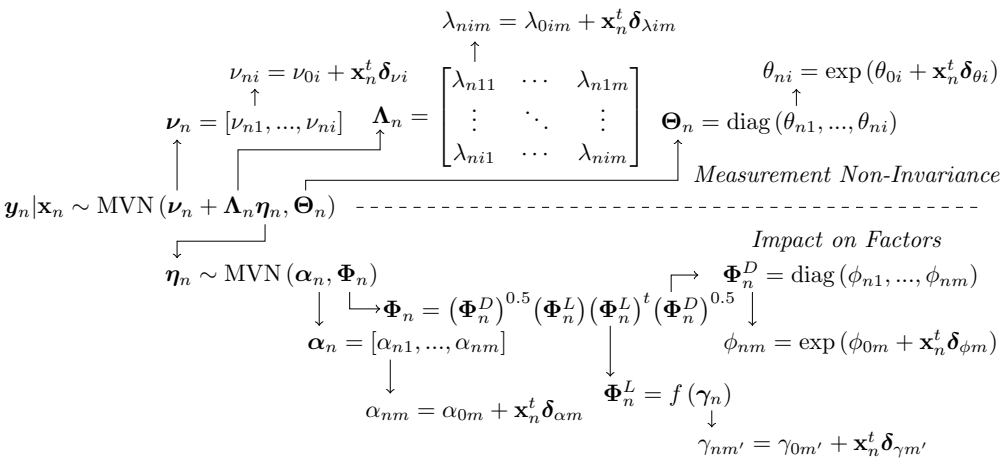

Figure 1. Model specification diagram for multidimensional MNLFA. Measurement non-invariance occurs when the moderation effects ($\boldsymbol{\delta}_\nu$, $\boldsymbol{\delta}_\lambda$, $\boldsymbol{\delta}_\theta$) on item intercepts ($\nu$), factor loadings ($\lambda$), and item residual variances (θ) are non-zero for some elements of each vector/matrix of effects; Impact on factors occurs when the moderation effects ($\boldsymbol{\delta}_\alpha$, $\boldsymbol{\delta}_\phi$, $\boldsymbol{\delta}_\gamma$) on factor means ($\boldsymbol{\alpha}$), factor variances ($\boldsymbol{\Phi}^D$), and factor correlations ($\boldsymbol{\Phi}^L$, Cholesky factorization) are non-zero for some elements of each vector/matrix of effects.

subsequent factors. This allows for more "stable" factors to be partialed out in the computation of later partial correlations, reducing the propagation of moderation effects through the ordering.

2.1.2 Reconstructing the Factor Covariance Matrix

Recovering the covariance matrix from the unconstrained parameters is not exactly intuitive from the equations alone. Suppose the inter-factor partial correlation are -0.50, 0.75, and 0.60. On the unconstrained scale (inverse hyperbolic tangent transform), these partial correlation are approximately -0.549, 0.973, and 0.693. These partial correlations are mapped to the Cholesky factor of the implied correlation matrix by the following:

$$P = \begin{bmatrix} 1 & 0 & 0 \\ -0.50 & 1 & 0 \\ 0.75 & 0.60 & 1 \end{bmatrix}$$

Apply Transformation:

$$L = \begin{bmatrix} 1 & 0 & 0 \\ -0.50 & \left\{1 \times \left(1 - (-0.50)^2\right)^{0.5}\right\} & 0 \\ 0.75 & \left\{0.60 \times \left(1 - (0.75)^2\right)^{0.5}\right\} & \left\{1 \times \left(1 - (0.75)^2\right)^{0.5} \times \left(1 - (0.60)^2\right)^{0.5}\right\} \end{bmatrix}$$

$$L = \begin{bmatrix} 1.00 & 0 & 0 \\ -0.50 & 0.87 & 0 \\ 0.75 & 0.39 & 0.53 \end{bmatrix}$$

therefore, the resulting factor correlation matrix is

$$LL^t = \begin{bmatrix} 1.00 & -0.50 & 0.75 \\ -0.50 & 1.00 & -0.03 \\ 0.75 & -0.03 & 1.00 \end{bmatrix}$$

With factor variances of 1.2, 1.5, and 0.80, the resulting factor covariance matrix is thus

$$
\begin{aligned}
DLL^tD &= \begin{bmatrix} 1.20 & 0.00 & 0.00 \\ 0.00 & 1.50 & 0.00 \\ 0.00 & 0.00 & 0.80 \end{bmatrix}^{0.50} \begin{bmatrix} 1.00 & -0.50 & 0.75 \\ -0.50 & 1.00 & -0.031 \\ 0.75 & -0.031 & 1.00 \end{bmatrix} \begin{bmatrix} 1.20 & 0.00 & 0.00 \\ 0.00 & 1.50 & 0.00 \\ 0.00 & 0.00 & 0.80 \end{bmatrix}^{0.50} \\
&= \begin{bmatrix} 1.20 & -0.67 & 0.73 \\ -0.67 & 1.50 & -0.034 \\ 0.73 & -0.034 & 0.80 \end{bmatrix}
\end{aligned}
$$

2.2 Impact of a continuous variable on correlations

The potential of this approach to inducing varying correlation structures depending on person parameters is shown in Figure 2. The figure plots the induced varying correlations across a continuous design variable x, where on average, the correlations are $\text{cor}(f_1, f_2) = 0.55$, $\text{cor}(f_1, f_3) = 0.65$, and $\text{cor}(f_2, f_3) = 0.83$; but, as shown, the magnitude of correlation can vary quite substantially as a function of a continuous design factor. Notably, although the effect of the predictor is linear in the space of the unconstrained parameters being moderated, the effect on the induced correlation is nonlinear due to the transformation. The nonlinearity is more noticeable for the correlation between factors 2 and 3 ($\text{cor}(f_2, f_3)$) because impact is propagated through the impact already present for the other two correlations; that is, the impact on the correlation between factors 2 and 3 is a function of the impact on the residual correlation $\text{cor}(f_2, f_3)$ and the impact on the correlation between factors 1 and 3 ($\text{cor}(f_1, f_3)$).

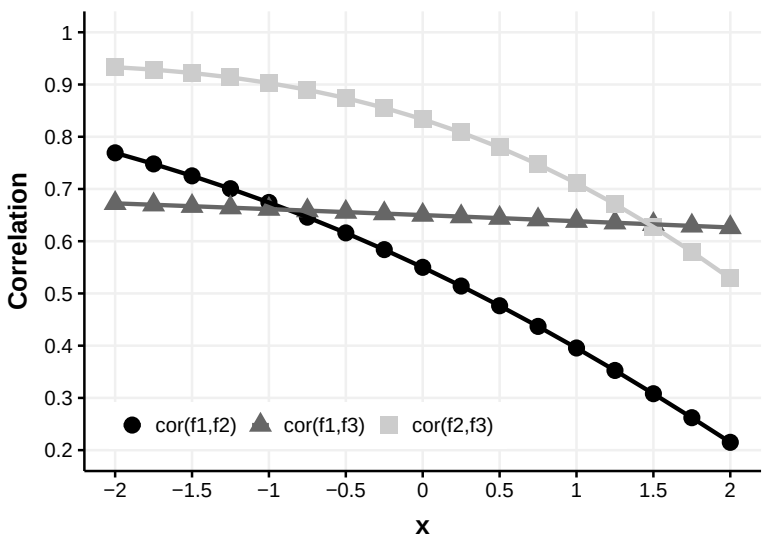

Figure 2. Induced varying correlations among three factors illustration differentiation among factors as a function of a continuous design factor x. Parameters used to simulate data based on the partial correlation parameterization: baseline residual correlations–$\gamma_{12} = 0.55$, $\gamma_{13} = 0.65$, $\gamma_{23} = 0.75$; magnitude of effects on x on residual correlations–$\delta_{\gamma 12} = -0.20$; $\delta_{\gamma 13} = -0.02$; $\delta_{\gamma 23} = -0.20$.

2.3 Estimation

Estimating moderated nonlinear factor analyses is an active area of methodological research (Bauer et al., 2020; Brandt et al., 2025; Chen & Bauer, 2024; Gottfredson et al., 2019; Kolbe et al., 2024). A

common theme across these studies involved the use of regularization to help identify the model given the complexity of the likelihood. An approach with promise involves using Bayesian methods to approximate the posterior distribution and regularizing priors such as spike-and-slab priors for the DIF parameters (Brandt et al., 2025; Chen & Bauer, 2024). The R package (Belzak, 2023) implements Bayesian moderated nonlinear factor analyses for models with up to two latent factors relying on regularizing priors for DIF parameters (Belzak, 2023) regDIF uses the general Bayesian MCMC package Just Another Gibbs Sampler (JAGS) to sample the marginal posterior distributions. JAGS is incredibly powerful and flexible, but can be painfully slow for models with large numbers of parameters. To help overcome this limitation, I implemented the multidimensional moderated nonlinear factor analysis in Stan (Carpenter et al., 2017), which is theoretically more efficient and can more easily explore the complex posterior distributions. I implemented the multidimensional MNLFA model in Stan (described next) and discuss issues encountered and remedies.

2.4 Bayesian approach

A Bayesian latent variable model can be implemented in several different ways depending on how the likelihood is set-up. Initially, I built the multidimensional MNLFA using a *conditional* likelihood approach to bring the implementation in line with how the package 'brms' is likely to implement latent variable models in version 3 (https://github.com/paul-buerkner/brms/issues/304).

The likelihood is implemented as shown in Figure 1. A complexity when implementing the model in Stan is the need to create independent foundation parameters that are then transformed into the parameters necessary for the likelihood function contribution itself. Also known as a *non-centered* parameterization so that the parameters being directly sampled have no direct dependence. This is accomplished largely by using independent normal parameters that are then transformed to get the quantities of interest. For instance, assuming there are N respondents and M latent variables, then the non-centered parameterization for the latent factors can be implemented to theoretically increase efficiency.

The non-centered parameterization is very memory intensive for moderated item factor analysis (MNLFA with binary/ordinal items) due to the need to supply 'nuisance" parameters (u) to utilize the underlying latent response distribution of the probit model. Alternatively, one could use a logistic regression-based approach to bypass this requirement, as done by Chen and Bauer (2024). Though it is unclear how well this approach will extend to multidimensional constructs with complex covariance structures.

Implementing the *conditional* likelihood parameterization had several issues. First, it was PAINFULLY slow, partially due to the huge number of parameters (number of cases $\times$ number of latent factors). This becomes worse in the case of binary items (probit model) where a nusance parameter is used under the hood for each response (not just each person). To help overcome this limitation, I rewrote the model to utilize the *marginal* likelihood, similar to what is utilized in lavaan (Rosseel, 2012) for estimating simple factor models, as I describe in more detail in the next section. Secondly, the conditional likelihood approach led to computer memory issues. When trying to implement the model above, I could not get the model to run with sample sizes above 5000 (individual cases, not individual responses even for "small" models with only 10-20 items) without requesting over 1000 GB of RAM on a cluster computer. Due to these issues, I decided to try an alternative parameterization.

2.4.1 Marginal likelihood approach

The marginal likelihood implemented in Stan is as follows. If the observed factor indicators are continuous variables, the person-wise marginal likelihood can be expressed as

$$\boldsymbol{\gamma}_n \sim \text{MVN}\left(\boldsymbol{\nu}_n + \boldsymbol{\Lambda}_n \boldsymbol{\alpha}_n, \boldsymbol{\Lambda}_n \boldsymbol{\Phi}_n \boldsymbol{\Lambda}_n^t + \boldsymbol{\Theta}_n\right)$$
$$\nu_{n,i} = \nu_{0i} + \mathbf{x}_n^t \boldsymbol{\delta}_{\nu i}$$

$$
\begin{aligned}
\lambda_{n,im} &= \lambda_{0im} + \mathbf{x}_n^t \boldsymbol{\delta}_{\lambda im} \\
\theta_{0,i} &= \theta_{0i} \exp(\mathbf{x}_n^t \boldsymbol{\delta}_{\theta i}) \\
\alpha_{n,m} &= \alpha_{0m} + \mathbf{x}_n^t \boldsymbol{\delta}_{\alpha m} \\
\boldsymbol{\Phi}_n &= \left(\boldsymbol{\Phi}_n^D\right)^{0.5} \left(\boldsymbol{\Phi}_n^L\right) \left(\boldsymbol{\Phi}_n^L\right)^t \left(\boldsymbol{\Phi}_n^D\right)^{0.5} \\
\boldsymbol{\Phi}_n^D &= \mathrm{diag}(\phi_{n1}, ..., \phi_{nm}) \\
\phi_{n,m} &= \phi_{0m} \exp(\mathbf{x}_n^t \boldsymbol{\delta}_{\phi m}) \\
\boldsymbol{\Phi}_n^L &= f(\boldsymbol{\gamma}_n) \\
\gamma_{n,m'} &= \gamma_{nm'} + \mathbf{x}_n^t \boldsymbol{\delta}_{\gamma m'}
\end{aligned}
\tag{3}
$$

The function $f(.)$ in equation 3 is the transformation from unconstrained partial correlations to the lower-triangular Cholesky factor of the factor correlation matrix. Using this marginal likelihood framing ensures the dimension of the parameter space does not increase with sample size, leading to more efficient sampling. A major benefit of this approach is that common SEM/CFA identification rules (Bollen, 1989) can be used to ensure the baseline parameters are easily identified, thereby facilitating the empirical identification of the remaining parameters. Examples of identification approaches for the baseline model are:

Fixed baseline factor scale. A fixed baseline factor variance is also possible. Such as fixing the factor variance to 1.0, and allowing the factor variance to vary as a function of background characteristics. This can work well in conjunction with a semi-informative prior on the factor loadings to establish the scale of the factor variables while still allowing for population differences in scale as a function of background characteristics. These differences would then be interpretable as changes in scales compared to the average across populations.

Marker variable/fixed loading. Constraining the baseline factor loadings of each variable/item's primary factor loading to be a fixed value or sign-constrained. This can be accomplished using a constraint on the parameter itself, or using a semi-informative prior that does not put much, if any, weight below zero; e.g., $\lambda_{0,im} \sim \mathrm{Normal}(1.0, 0.5)$.

2.4.2 Spike-and-slab priors for DIF & Impact

Spike-and-slab priors have seen a growing interest in the area of DIF detection (Belzak & Bauer, 2020; Brandt et al., 2018, 2025; Chen et al., 2021). The spike-and-slab prior can be implemented in a variety of ways depending on the needs of the situation. One way previously implemented is

$$
\begin{aligned}
\delta &\sim \omega \mathrm{Normal}\left(0, \psi\tau\sigma^2\right) \\
\psi &\sim \mathrm{exponential}\,(0.5) \\
\omega &\sim \mathrm{Beta}\,(1, 1) \\
\tau^{-1} &\sim \mathrm{Gamma}\,(9, 4)
\end{aligned}
$$

where δ is a placeholder for any of the DIF/impact effects, ω is the inclusion probability, ψ and τ are regularizing variance parameters across a group of jointly regularized parameters, and σ^2 is the item specific residual variance.

When implementing the above parameterization of the spike-and-slab prior, I found that the individual regularizing variance parameters were not unique leading to sampling difficulties in Stan. An alternative parameterization designed specifically for Stan is described next. Thomson et al. (2019) provided a logit-normal continuous parameterization that is especially efficient for use in Stan since all sampled parameters are continuous with unbounded support that can then be transformed to the

standard spike-and-slab structure.

$$\delta \sim \omega \times \text{Normal}(0, \tau)$$
$$\text{logit}(\omega) = \log\left(\frac{\omega_\mu}{1-\omega_\mu}\right) + \text{Normal}(0,1) \times \omega_{sd}$$

where ω is the inclusion probability with hyperparameters ω_μ for prior average inclusion probability and ω_{sd} for the standard deviation of the logit of the inclusion probability, and τ^2 is the prior standard deviation of the DIF/impact effect. The above prior parameterization has τ fixed (e.g., $\tau = 5.0$), as done in all the simulations conducted by Thompson and colleagues, which greatly improves the speed and sampling efficiency of the spike and slab prior. The standard deviation prior can itself be given a hyper-prior, such as a Gamma distribution, to provide an "adaptive" component (Brandt et al., 2025). However, in experimentation, I've found that including the "adaptive" piece leads to a significant increase in computation time and slightly more diffuse posterior distributions for the DIF/impact effects. A major advantage of the fixed variance approach is only needing to use standard normal priors then shifting/rescaling as needed combined with the inclusion probability to achieve a more or less regularizing prior distribution; that is, changing the prior inclusion probability (mean, ω_μ, and standard deviation, ω_{sd}) and increasing or decreasing τ for more or less regularizing. Together, this leads to a relatively straightforward interpretation of the regularizing prior that is easy to specify and modify as needed for sensitivity analyses.

3. Six Factor Global Flourishing Study Example

3.1 Data

To illustrate the method, I used data from the Global Flourishing Study (Johnson et al., 2024; VanderWeele et al., 2025). The Global Flourishing Study is a five-year, longitudinal study investigating the predictors of human flourishing across approximately 200,000 participants from 20+ geographically and culturally diverse countries and territories. The current analyses are restricted to participants who completed Wave 2 of the survey and responded to all twelve Secure Flourishing Index items. The final sample size is 124,993. The average age of the sample is 47.2 (SD=17.3) and comprises 51.2% females. The Secure Flourishing Index is a 12-item scale measuring 6 domain of flourishing: life satisfaction and happiness (LS/H), physical and mental health (P/MH), meaning and purpose (MP), character and virtue (CV), close social relationships (CSR), and financial and material stability (F/MS). Each domain is measured by two indicators, and all items use a 0-10 point scale (VanderWeele, 2017).

For estimating the model, the six factors were ordered as follows: [LS/H, P/MH, MP, CV, CSR, F/MS]. This ordering was chosen on substantive grounds, placing factors hypothesized to have stronger correlations earlier in the sequence, as the partial correlations for later factors are conditioned on all earlier factors.

3.2 Results

The results are reported in Padgett et al. (2026), and we provide a highlight of the moderation in the factor correlations. The correlations between several factors were identified to vary by age (see Figure 3). As described in more detail in Padgett et al. (2026), the magnitude of the correlation between MP and CV depends on age: younger respondents (below -1.5 SD below the average age) have a stronger correlation, participants of average age (between -1.5 and 1.5 SD around the average age) exhibit approximately the same correlation of 0.75, and older respondents (above 1.5 SD) show a markedly lower correlation. The nonlinear relationship could indicate that older participants increasingly differentiate Meaning/Purpose from Character/Virtue as concepts, more so than younger participants. Specifically, the posterior implied correlation between Meaning/Purpose from Character/Virtue for participants of an average age was 0.762 (95% CI: 0.757, 0.766). At -2.0 SD below the mean age,

this correlation was estimated at 0.787 (95% CI: 0.774, 0.799), compared to 0.704 (95% CI: 0.680, 0.728) at +2.0 SD, reflecting a change of approximately 0.08 correlation units across this age range.

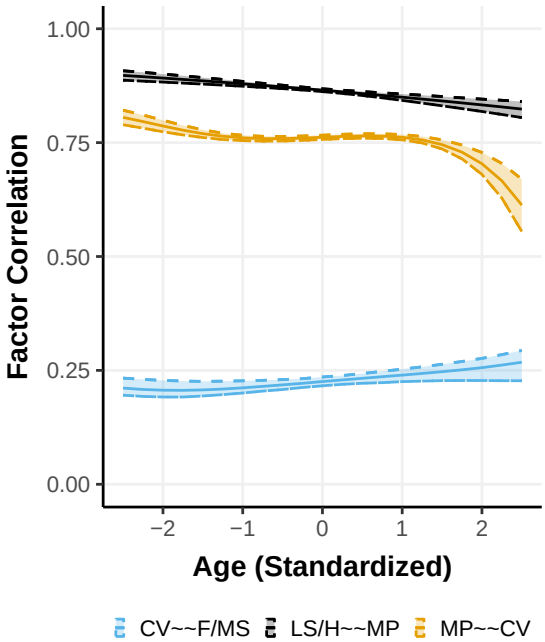

Figure 3. The correlations among Character & Virtue (CV), Meaning & Purpose (MP), Financial & Material Stability (F/MS), and Life Satisfaction & Happiness were found to vary by age after controlling for differential item functioning. Dashed lines represent the 95% credible interval band of the back-transformed posteriors of the unconstrained parameter space.

4. Conclusion

The conceptualization of measurement non-invariance as a function of several background characteristics is a flexible analytic tool (Bauer, 2023). Despite advances in moderated factor-analytic approaches, the complex nature of extending these tools to multidimensional constructs has been a limitation to the wider use. This paper outlines how this limitation can be overcome by incorporating a unique transformation previously used to generate random correlation matrices. Using the methods outlined in this paper, researchers will be able to jointly test measurement invariance and impact on the latent structure using a flexible design matrix of effects.

The estimation approach using a Bayesian model with spike-and-slab priors is an open question. Although my experience implementing the methods described suggests that the proposed approach might be useful for relatively small samples (N < 10,000), larger sample sizes or models may need alternative estimation methods due to computational constraints. Using a penalized maximum likelihood approach might be one such way to achieve empirical identification of the full model (Asparouhov & Muthén, 2024; Lüdtke et al., 2021). Regardless of the estimation approach, more work is needed to make the specification and workflow for using these models practical for researchers. This paper provides a foundation for which such new tools can be built by illustrating how these models can be estimated.

Funding Statement No funding was provided for the current paper.

Competing Interests None.

References

Anderson, T. W. (1973). Asymptotically efficient estimation of covariance matrices with linear structure. *Annals of statistics*, *1*(1), 135–141. https://doi.org/10.1214/aos/1193342389

Archakov, I., & Hansen, P. R. (2021). A new parametrization of correlation matrices. *Econometrica: journal of the Econometric Society*, *89*(4), 1699–1715. https://doi.org/10.3982/ecta16910

Asparouhov, T., & Muthén, B. (2024). Penalized structural equation models. *Structural equation modeling: a multidisciplinary journal*, *31*(3), 429–454. https://doi.org/10.1080/10705511.2023.2263913

Bauer, D. J. (2017). A more general model for testing measurement invariance and differential item functioning. *Psychological methods*, *22*(3), 507–526. https://doi.org/10.1037/met0000077

Bauer, D. J. (2023). Enhancing measurement validity in diverse populations: Modern approaches to evaluating differential item functioning. *The British journal of mathematical and statistical psychology*, *76*(3), 435–461. https://doi.org/10.1111/bmsp.12316

Bauer, D. J., Belzak, W. C. M., & Cole, V. (2020). Simplifying the assessment of measurement invariance over multiple background variables: Using regularized moderated nonlinear factor analysis to detect differential item functioning. *Structural equation modeling: a multidisciplinary journal*, *27*(1), 43–55. https://doi.org/10.1080/10705511.2019.1642754

Belzak, W. C. M. (2023). The regDIF R package: Evaluating complex sources of measurement bias using regularized differential item functioning. *Structural equation modeling: a multidisciplinary journal*, *30*(6), 974–984. https://doi.org/10.1080/10705511.2023.2170235

Belzak, W. C. M., & Bauer, D. J. (2020). Improving the assessment of measurement invariance: Using regularization to select anchor items and identify differential item functioning. *Psychological methods*, *25*(6), 673–690. https://doi.org/10.1037/met0000253

Bollen, K. A. (1989). *Structural equations with latent variables*. John Wiley & Sons.

Brandt, H., Cambria, J., & Kelava, A. (2018). An adaptive bayesian lasso approach with spike-and-slab priors to identify multiple linear and nonlinear effects in structural equation models. *Structural equation modeling: a multidisciplinary journal*, *25*(6), 946–960. https://doi.org/10.1080/10705511.2018.1474114

Brandt, H., Chen, S. M., & Bauer, D. J. (2025). Bayesian penalty methods for evaluating measurement invariance in moderated nonlinear factor analysis. *Psychological methods*, *30*(3), 482–512. https://doi.org/10.1037/met0000552

Bucci, A., Ippoliti, L., & Valentini, P. (2022). Comparing unconstrained parametrization methods for return covariance matrix prediction. *Statistics and computing*, *32*(5). https://doi.org/10.1007/s11222-022-10157-4

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A probabilistic programming language. *Journal of statistical software*, *76*(1). https://doi.org/10.18637/jss.v076.i01

Chen, S. M., & Bauer, D. J. (2024). Modeling construct change over time amidst potential changes in construct measurement: A longitudinal moderated factor analysis approach. *Psychological methods*. https://doi.org/10.1037/met0000685

Chen, S. M., Bauer, D. J., Belzak, W. C. M., & Brandt, H. (2021). Advantages of spike and slab priors for detecting differential item functioning relative to other bayesian regularizing priors and frequentist lasso. *Structural Equation Modeling: A Multidisciplinary Journal*, *29*(1), 122–139. https://doi.org/10.1080/10705511.2021.1948335

Clauser, B. E., & Mazor, K. M. (1998). Using statistical procedures to identify differentially functioning test items. an NCME instructional module. *Educational Measurement: Issues and Practice*, *17*(1), 31–44.

Curran, P. J., McGinley, J. S., Bauer, D. J., Hussong, A. M., Burns, A., Chassin, L., Sher, K., & Zucker, R. (2014). A moderated nonlinear factor model for the development of commensurate measures in integrative data analysis. *Multivariate behavioral research*, *49*(3), 214–231. https://doi.org/10.1080/00273171.2014.889594

Gottfredson, N. C., Cole, V. T., Giordano, M. L., Bauer, D. J., Hussong, A. M., & Ennett, S. T. (2019). Simplifying the implementation of modern scale scoring methods with an automated R package: Automated moderated nonlinear factor analysis (aMNLFA). *Addictive behaviors*, *94*, 65–73. https://doi.org/10.1016/j.addbeh.2018.10.031

Joe, H. (2006). Generating random correlation matrices based on partial correlations. *Journal of multivariate analysis*, *97*(10), 2177–2189. https://doi.org/10.1016/j.jmva.2005.05.010

Johnson, B. R., Ritter, Z., Fogleman, A., Markham, L., Stankov, T., Srinivasan, R., Honohan, J., Ripley, A., Phillips, T., Wang, H., & VanderWeele, T. J. (2024). The global flourishing study. https://doi.org/10.17605/OSF.IO/3JTZ8

Kolbe, L., Molenaar, D., Jak, S., & Jorgensen, T. D. (2024). Assessing measurement invariance with moderated nonlinear factor analysis using the R package OpenMx. *Psychological methods*, *29*(2), 388–406. https://doi.org/10.1037/met0000501

Lewandowski, D., Kurowicka, D., & Joe, H. (2009). Generating random correlation matrices based on vines and extended onion method. *Journal of Multivariate Analysis*, *100*(9), 1989–2001. https://doi.org/10.1016/j.jmva.2009.04.008

Lüdtke, O., Ulitzsch, E., & Robitzsch, A. (2021). A comparison of penalized maximum likelihood estimation and markov chain monte carlo techniques for estimating confirmatory factor analysis models with small sample sizes. *Frontiers in Psychology*, *12*, 615162. https://doi.org/10.3389/fpsyg.2021.615162

Molenaar, D. (2021). A flexible moderated factor analysis approach to test for measurement invariance across a continuous variable. *Psychological methods*, *26*(6), 660–679. https://doi.org/10.1037/met0000360

Padgett, R. N., Johnson, B. R., & VanderWeele, T. J. (2026). Evaluating measurement (non-)invariance of the secure flourishing index: A bayesian moderated factor analysis approach [PsyArXiV]. https://doi.org/10.31234/osf.io/y83zn_v2

Pinheiro, J. C., & Bates, D. M. (1996). Unconstrained parametrizations for variance-covariance matrices. *Statistics and computing*, *6*(3), 289–296. https://doi.org/10.1007/bf00140873

Pourahmadi, M. (1999). Joint mean-covariance models with applications to longitudinal data: Unconstrained parameterisation. *Biometrika*, *86*(3), 677–690. https://doi.org/10.1093/biomet/86.3.677

Pourahmadi, M. (2011). Covariance estimation: The GLM and regularization perspectives. *Statistical science: a review journal of the Institute of Mathematical Statistics*, *26*(3), 369–387. https://doi.org/10.1214/11-sts358

Pourahmadi, M., & Wang, X. (2015). Distribution of random correlation matrices: Hyperspherical parameterization of the cholesky factor. *Statistics & probability letters*, *106*, 5–12. https://doi.org/10.1016/j.spl.2015.06.015

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd). SAGE Publications.

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software*, *48*(2), 1–36. http://www.jstatsoft.org/v48/i02/

Thissen, D. (2025). A review of some of the history of factorial invariance and differential item functioning. *Multivariate behavioral research*, *60*(2), 211–235. https://doi.org/10.1080/00273171.2024.2396148

Thomson, W., Jabbari, S., Taylor, A. E., Arlt, W., & Smith, D. J. (2019). Simultaneous parameter estimation and variable selection via the logit-normal continuous analogue of the spike-and-slab prior. *Journal of the Royal Society, Interface*, *16*(150), 20180572. https://doi.org/10.1098/rsif.2018.0572

van Erp, S., Mulder, J., & Oberski, D. L. (2018). Prior sensitivity analysis in default bayesian structural equation modeling. *Psychological Methods*, *23*(2), 363–388. https://doi.org/10.1037/met0000162

van de Vijver, F. J. R., Avvisati, F., Davidov, E., Eid, M., Fox, J.-P., Le Donné, N., Lek, K., Meuleman, B., Paccagnella, M., & van de Schoot, R. (2019). *Invariance analyses in large-scale studies*. Organisation for Economic Co-operation; Development OECD.

Vandenberg, R. J., & Lance, C. E. (2000). A review and synthesis of the measurement invariance literature: Suggestions, practices, and recommendations for organizational research. *Organizational research methods*, *3*(1), 4–70. https://doi.org/10.1177/109442810031002

VanderWeele, T. J. (2017). On the promotion of human flourishing. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(31), 8148–8156. https://doi.org/10.1073/pnas.1702996114

VanderWeele, T. J., Johnson, B. R., Bialowolski, P. T., Bonhag, R., Bradshaw, M., Breedlove, T., Case, B., Chen, Y., Chen, Z. J., Counted, V., Cowden, R. G., de la Rosa, P. A., Felton, C., Fogleman, A., Gibson, C., Grigoropoulou, N., Gundersen, C., Jang, S. J., Johnson, K. A., … Yancey, G. (2025). The global flourishing study: Study profile and initial results on flourishing. *Nature Mental Health*, *3*(6), 636–653. https://doi.org/10.1038/s44220-025-00423-5

Appendix 1. Symbol meaning

- $\mathbf{y}_n$, vector of item responses for person n
- $\mathbf{x}_n$, vector of background characteristics for person n
- ν_{ni}, item intercept for person n on item i
- ν_{0i}, baseline item intercept for person n on item i
- $\boldsymbol{\delta}_{\nu i}$, vector of DIF effects associated with each person variable in $\mathbf{x}$ for item intercepts of item i
- λ, item factor loading
- η, value on latent factor (factor score)
- θ variance of residuals
- $\boldsymbol{\alpha}$, latent factor mean (baseline is fixed to 0.0)
- $\boldsymbol{\Phi}$, factor covariance matrix, separated out into factor variance $\boldsymbol{\Phi}^D$ and Cholesky factor $\boldsymbol{\Phi}^L$
- $\boldsymbol{\gamma}$, vector of unconstrained partial correlations

Appendix 2. Build Positive Definite Correlation Matrix - R Code

```
#' R Code -- Builds a Cholesky factor (lower triangular matrix of
    dimension K) from a vector of unconstrained parameters (length K*(K
    -1)/2)
#' @param x, vector of unconstrainted partial correlations length(x)
    must equal K*(K-1)/2
#' @param K, dimension of matrix (KxK)
```

```
construct_chol_mat_lkj <- function(x,K){
  z <- matrix(0, nrow=K,ncol=K) ; w <- matrix(0, nrow=K,ncol=K) ;
  r <- matrix(0, nrow=K,ncol=K) ;  x_tanh <- numeric(length(x));
  x_tanh <- tanh(x); #hyperbolic tan tranform of y (unconstrained
      partial correlations)
  z[lower.tri(z,FALSE)] <- x_tanh; diag(z) = 1.0;
  for(i in 1:K){
    for(j in 1:K){
      if(i < j) w[i,j] = 0.0;
      if(1 == j && j == i) w[i,j] = 1.0;
      if(1 == j && j < i) w[i,j] = z[i,j];
      if(1 < j && j == i){
        w[i,j] = cumul_prod( (1.0 - (z[i,1:(j-1)])^2.0)^0.5 );
      }
      if(1 < j && j < i){
        w[i,j] = z[i,j] * cumul_prod( (1.0 - (z[i,1:(j-1)])^2.0)^0.5 )
  }}}
  w;
}
#' Compute cumulative product of a vector x
cumul_prod <- function(x){
  K = length(x); z = 1;
  for(k in 1:K) z = z*(x[k]);
  return(z);
}
# Ex 1. 2x2 matrix => a single correlation
corr_mat <- matrix(c(1,0.70,0.70,1.0),ncol=2)
x <- atanh(0.70) # inverse Fisher transform correlation to get input
    for cholesky construction
L <- construct_chol_mat_lkj(x, 2)
L %*% t(L)
all.equal(L %*% t(L) , corr_mat)
# Ex 2. 3x3 matrix => a single correlation
K <- 3; x <- atanh(c(-0.50, 0.75, 0.60))
L <- construct_chol_mat_lkj(x, K); L
mycorr <- L %*% t(L)
det(mycorr) > 0 ## is the correlation matrix invertible?
# Ex 3. 6x6 matrix
set.seed(1)
K <- 6; x <- rnorm(K*(K-1)/2) # generate a random set of partial cor.
L <- construct_chol_mat_lkj(x, K); L
mycorr <- L %*% t(L)
det(mycorr) > 0 ## is the correlation matrix invertible?
## Ex 4. Three dimensions with diagonal D matrix
K <- 3; x <- atanh(c(-0.50, 0.75, 0.60))
L <- construct_chol_mat_lkj(x, K); L
D <- sqrt(diag(c(1.2, 1.5, 0.8)))
mycov <- D %*% L %*% t(L)%*% D; mycov
det(mycov) > 0 ## is the covariance matrix invertible?
```