STRUCTURAL EQUATION MODELING
2026, VOL. 33, NO. 5, 734–744

OPEN ACCESS

Regularized Matrix Decomposition Structural Equation Modeling

Naoto Yamashita

Kansai University

ABSTRACT

Matrix Decomposition Structural Equation Modeling (MDSEM) provides a data-matrix-based alternative to conventional SEM procedures. MDSEM is shown to be highly vulnerable to multicollinearity among latent variables, often resulting in biased and unstable estimates of parameters. This study proposes Regularized Matrix Decomposition Structural Equation Modeling (RMDSEM), an extension of MDSEM that introduces explicit L_2 regularization. RMDSEM addresses the multicollinearity issue by regularizing the path coefficient matrix while preserving the matrix-decomposition framework. An iterative estimation algorithm and a *K*-fold cross-validation procedure for tuning multiple regularization parameters are derived. Simulation results show that MDSEM and maximum-likelihood SEM yield inflated bias and standard errors under moderate to strong multicollinearity, whereas RMDSEM achieves stable estimation with substantially reduced mean squared errors. Empirical applications further demonstrate that RMDSEM suppresses unreasonable estimates and yields interpretable structural relations.

KEYWORDS

Covariance structure analysis; matrix decomposition; multicollinearity; regularization; structural equation modeling

1. Introduction

Structural Equation Modeling (SEM; Bentler, 1980, 1986; Mulaik, 1986) is a general and flexible framework for modeling multiple observed variables, and it has been applied extensively across disciplines such as psychology and sociology (Bollen, 2002; MacCallum & Austin, 2000). A central assumption in SEM is the existence of latent variables that are not directly observed, and the variation of observed variables can be explained by the latent variables. After specifying a causal model among observed and latent variables, typically as a form of path diagram, SEM fits the model-implied covariance matrix (i.e., the covariance structure) derived from the causal model to the sample covariance matrix.

Let $\mathbf{X}$ $(n \times p)$ denote a centered data matrix for n individuals on p observed variables, and let $\mathbf{S}_\mathbf{X} = n^{-1}\mathbf{X}'\mathbf{X}$ be the sample covariance matrix. Let $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ $(p \times p)$ denote the model-implied covariance matrix, called covariance structure, expressed as a function of the vector of unknown parameters $\boldsymbol{\theta}$ in the hypothetical model. In SEM parameter estimation, an estimate of $\boldsymbol{\theta}$ can be obtained by minimizing, with respect to $\boldsymbol{\theta}$, a discrepancy function $d(\mathbf{S}_\mathbf{X}, \boldsymbol{\Sigma}(\boldsymbol{\theta}))$ that quantifies a certain discrepancy between the two matrices. Common choices of the function d include the negative log-likelihood and least-squares criteria (Kline, 2023). The same logic of parameter estimation is also employed in parameter estimation in a factor analysis model, which constitutes a major class of models subsumed by SEM (Terada, 2025).

1.1. Two Emerging SEM Procedures: GSCA and MDSEM

In contrast to the conventional SEM formulated as a covariance-structure fitting problem, alternative approaches based on different mathematical formulations have been proposed to enable equivalent modeling to SEM. Among them, one notable approach has been proposed by Hwang & Takane (2004), called Generalized Structured Component Analysis (GSCA), in which common factor scores are defined as linear combinations of observed variables. GSCA estimates the loading matrix $\boldsymbol{\Lambda}(p \times r)$, the path coefficient matrix $\mathbf{B}(r \times r)$, and the weight matrix $\mathbf{W}(p \times r)$ by minimizing the least-squares criterion

$$l_G = ||\mathbf{X} - \mathbf{X}\mathbf{W}\boldsymbol{\Lambda}'||^2 + ||\mathbf{X}\mathbf{W} - \mathbf{X}\mathbf{W}\mathbf{B}||^2 \qquad (1)$$

subject to the model specification, where some elements of the parameter matrices are constrained to 0 and only a subset are treated as free parameters to be estimated. A distinguishing feature of GSCA can be found in (1), where the data matrix $\mathbf{X}$ is approximated by the model counterpart, and this is not a covariance-fitting problem. The matrix of r scommon factor scores for n individuals in GSCA, noted as $n \times r$ matrix $\mathbf{F}$, is constrained as a linear combination of p observed variables, expressed as $\mathbf{F} = \mathbf{X}\mathbf{W}$ using a certain weight matrix $\mathbf{W}$. This constraint, referred to as the weighted relation model, leads to a distinguishing property of GSCA: factor score uniqueness. Namely, given a certain $\mathbf{W}$, which is uniquely estimated in GSCA, $\mathbf{F}$ is also uniquely computed as a linear combination of the p variables in $\mathbf{X}$ because of the weighted relation model (Hwang & Takane, 2014).

© 2026 The Author(s).
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

In contrast, no such constraint is imposed in the conventional SEM, and thus SEM inherits the well-known indeterminacy of factor scores as in factor analysis (Steiger, 1979, 1996). Moreover, several extensions have been proposed, including Integrated GSCA for more flexible modeling (Hwang et al., 2021), Bayesian extensions (Choi & Hwang, 2020), and model exploration procedures for GSCA (Yamashita, 2024).

Following GSCA, Yamashita (2024) proposed Matrix Decomposition SEM (MDSEM) as an extension of Matrix Decomposition Factor Analysis (MDFA; Adachi & Trendafilov, 2018). MDSEM does not impose a constraint that defines the factor score matrix as a linear combination of observed variables, as GSCA does. Accordingly, factor scores are not unique in MDSEM, and in this respect MDSEM can be regarded as more theoretically consistent with conventional SEM than GSCA. Although $\mathbf{F}$ and $\mathbf{U}$ appear as variables in the optimization of (2), they are not uniquely determined due to factor-score indeterminacy and thus do not constitute estimated quantities in the usual sense. Instead, the covariance matrix between $\mathbf{X}$ and the factor scores is uniquely determined and serves as the quantity of inferential interest (Yamashita, 2024). In addition, Yamashita (2024) established important properties of MDSEM, including that the sufficient statistic is the sample covariance matrix; these properties do not hold for GSCA.

In MDSEM, for an arbitrary nonnegative constant α, the least-squares criterion

$$l_M = ||\mathbf{X} - (\mathbf{F}\mathbf{\Lambda}' + \mathbf{U}\mathbf{\Psi}^{1/2})||^2 + \alpha||\mathbf{F} - \mathbf{F}\mathbf{B}||^2 \qquad (2)$$

is minimized over $\mathbf{F}$, $\mathbf{\Lambda}$, $\mathbf{B}$, the unique factor score matrix $\mathbf{U}$ $(n \times p)$, and the diagonal matrix $\mathbf{\Psi}^{1/2}$ whose diagonal entries are the square roots of uniquenesses. The factor scores are constrained as

$$\mathbf{1}_n'\mathbf{F} = \mathbf{0}_r', \ \mathbf{1}_n'\mathbf{U} = \mathbf{0}_p', \ \mathrm{diag}(n^{-1}\mathbf{F}'\mathbf{F}) = \mathbf{I}_r, \ n^{-1}\mathbf{U}'\mathbf{U} = \mathbf{I}_p, \ \mathbf{F}'\mathbf{U} = {}_r\mathbf{O}_p. \qquad (3)$$

Here, depending on the specified model, only a certain set of elements in $\mathbf{\Lambda}$ and $\mathbf{B}$ are treated as free parameters. The first term of l_M corresponds to the least-squares criterion minimized in MDFA, whereas the second term represents regression relations among common factor scores. This formulation enables MDSEM to represent causal models among latent variables in the same manner as conventional SEM. In addition, uniquenesses in MDSEM are obtained as $\mathbf{\Psi}$; because the diagonal elements of $\mathbf{\Psi}^{1/2}$ are squared, uniquenesses are guaranteed to be nonnegative. Therefore, MDSEM avoids improper solutions that often arise in SEM.

It should be noted that both GSCA and MDSEM have a distinguishing property against a similar SEM-like procedure, partial least squares path modeling (PLS-PM) (Dijkstra and Henseler, 2015; Hair et al., 2021); GSCA and MDSEM minimize a single optimization criterion defined in (1) or (2), respectively, while PLS-PM does not.

1.2. Multicollinearity in SEM Procedures

One of the major issues in SEM procedures is multicollinearity of observed and latent variables (Hwang, 2009). In general, multicollinearity refers to a situation in regression where two or more independent variables are highly correlated, which is known to yield unstable estimation of regression coefficients. For example, in multiple regression, if the set of independent variables includes pairs whose absolute correlation is close to unity, the estimated partial regression coefficients are estimated with substantial bias and large standard errors, resulting in unstable estimation. Similarly in SEM, when some of the observed variables are highly correlated, estimation of path coefficients would be unstable in models where those observed variables are used as independent variables, namely formative models. Moreover, multicollinearity can also occur among latent variables in SEM. Specifically, in reflective models—regression models in which observed variables are dependent variables and latent variables are independent variables—high correlations among multiple latent variables can induce multicollinearity and destabilize estimation of factor loadings. Furthermore, when such highly correlated latent variables serve as independent variables in a causal model among latent variables, estimation of path coefficients also becomes unstable. Thus, because SEM can be affected by multicollinearity through multiple mechanisms, addressing multicollinearity is of critical practical importance.

Regularization has emerged as a shared strategy for addressing multicollinearity across the two major paradigms of SEM. In covariance-based SEM, Jacobucci et al. (2016) proposed RegSEM, which incorporates L_2 penalties into maximum likelihood estimation, stabilizing the estimation of structural parameters under multicollinearity. In component-based SEM, Hwang (2009) proposed Regularized GSCA (RGSCA) by adding L_2 regularization terms for $\mathbf{\Lambda}$, $\mathbf{W}$, and $\mathbf{B}$ to (1); it minimizes

$$l_{RG} = l_G + \alpha||\mathbf{W}||^2 + \beta||\mathbf{\Lambda}||^2 + \gamma||\mathbf{B}||^2 \qquad (4)$$

where α, β, and γ are nonnegative penalty parameters controlling the strength of regularization; a regularized extension of Integrated GSCA (Cho et al., 2025) has also been developed along the same lines. Despite their differences in formulation, these methods share a common motivation: regularization provides a principled mechanism for stabilizing estimation when latent variables are highly correlated. MDSEM represents a third, matrix-decomposition-based approach to SEM, and extending it with regularization follows the same methodological imperative that has driven analogous developments in both covariance-based and component-based procedures. The present study proposes such an extension, termed Regularized MDSEM (RMDSEM).

1.3. Multicollinearity in MDSEM

Yamashita (2024) indicated that multicollinearity may be a more serious issue in MDSEM than in other SEM procedures due to non-uniqueness of the common factor score. As noted above, GSCA constrains $\mathbf{F} = \mathbf{X}\mathbf{W}$, meaning that

common factor scores are restricted to be linear combinations of observed variables, whereas MDSEM imposes no such constraint. Therefore, relative to GSCA, MDSEM allows one to determine common factor scores more freely. Consequently, it can be possible to obtain $\mathbf{F}$ that makes the second term of (2) attain its minimum value 0, or a value close to 0, with a trivial solution. For example, when $r = 2$, consider a causal model in which the first factor is regressed on the second factor, namely a path coefficient matrix $\mathbf{B}$ given by

$$\mathbf{B} = \begin{bmatrix} 0 & 0 \\ b & 0 \end{bmatrix} \tag{5}$$

where b denotes the path coefficient corresponding to the path from the second factor to the first factor. Let $\mathbf{f}_j$ denote the $j(= 1, \ldots, r)$-th column of $\mathbf{F}$. If we set $\mathbf{f}_1 = b\mathbf{f}_2$, then the second term of (2), namely the residual sum of squares for the regression among common factor scores, becomes 0. In this case, the correlation between the first and second factor scores approaches unity, inducing multicollinearity and unstable estimation of factor loadings and path coefficients involving these factors.

To address this issue, Yamashita (2024) proposed multiplying $||\mathbf{F} - \mathbf{FB}||^2$ by a small nonnegative constant α, as in (2), thereby reducing the relative importance of the regression among common factors in the overall objective function, suppressing correlations among common factors, and consequently preventing multicollinearity. The value of α is specified prior to analysis via cross-validation.

However, as shown below, this approach cannot adequately prevent multicollinearity. The second term of (2) can be written as

$$||\mathbf{F} - \mathbf{FB}||^2 = ||\mathbf{F}(\mathbf{I}_r - \mathbf{B})||^2. \tag{6}$$

This representation reveals that the second term of (2) has an inherently destabilizing effect on the estimation of $\mathbf{B}$. In the alternating optimization of $\mathbf{F}$ and $\mathbf{B}$, detailed in Yamashita (2024), when $\mathbf{B}$ is updated with $\mathbf{F}$ fixed, minimizing this term is equivalent to an ordinary least-squares regression of $\mathbf{f}_1$ on $\mathbf{f}_2$, which pushes the path coefficient b toward the inter-factor correlation ρ, which is generally nonzero in practice. Conversely, when $\mathbf{F}$ is updated with $\mathbf{B}$ fixed, this term drives $\mathbf{f}_1$ toward $b\mathbf{f}_2$, since the former is regressed on the latter, increasing the correlation between factor scores toward unity. It should be noted that $\mathbf{F}$ itself is not estimated due to factor score indeterminacy, but the covariance matrix between latent and observed variables, which is uniquely determined, is estimated. These two steps form a positive feedback loop—b is pulled toward ρ, and ρ is simultaneously driven toward 1—progressively amplifying multicollinearity and therefore destabilizing the estimation of $\mathbf{B}$. One might consider reducing α to suppress this feedback. However, as $\alpha \to 0$, the structural model represented by $\mathbf{B}$ becomes increasingly negligible in the objective function, rendering the estimation of $\mathbf{B}$ itself meaningless. Thus, α governs a fundamental tradeoff: a larger α worsens multicollinearity, whereas a smaller α undermines the estimation of $\mathbf{B}$. This tradeoff implies that tuning α alone cannot simultaneously achieve stable and meaningful estimation of $\mathbf{B}$, and consequently that this strategy has inherent limitations for addressing multicollinearity in MDSEM.

To illustrate MDSEM's vulnerability to multicollinearity, we present an application of MDSEM to the American Customer Satisfaction Index data (Fornell et al., 1996). As described in detail in Section 4.1, this dataset contains responses of 152 respondents to 14 items concerning purchase satisfaction and experience. We applied MDSEM with the number of common factors set to $r = 6$, and Table 1 reports the resulting path coefficients among factors and their bootstrap standard errors. The penalty parameter α in (2) was selected as $\alpha = 0.001$ via 5-fold cross validation. The MDSEM results show relatively large regression coefficients and large standard errors for some paths, indicating instability in estimating some elements of the path coefficient matrix $\mathbf{B}$. In addition, some estimated path coefficients are unreasonable, judging from their signs. For example, customer expectations (CE) negatively affect perceived value (PV) (b_2) and customer satisfaction (CS) (b_3), since the corresponding coefficients are negatively estimated, which is substantively unreasonable. These findings suggest the presence of multicollinearity, and thus constitute a serious practical limitation of MDSEM in interpretation of its result.

1.4. Purpose of the Research

In this study, we propose an L_2-regularized approach to address multicollinearity in MDSEM. The proposed method introduces L_2 regularization exclusively on the path coefficient matrix $\mathbf{B}$, and is termed Regularized MDSEM (RMDSEM). In MDSEM, which is a purely reflective latent variable model, $\mathbf{X}$ serves as an outcome of the latent factors $\mathbf{F}$, and multicollinearity manifests among the latent factors themselves rather than in the loading structure. Consequently, L_2 regularization on $\mathbf{\Lambda}$ is not theoretically warranted: shrinking factor loadings toward zero does not address latent factor collinearity and would merely attenuate the loadings without principled justification. In contrast, GSCA is a circular system in which $\mathbf{X}$ simultaneously defines and is predicted by the components, making regularization of $\mathbf{\Lambda}$ necessary to prevent multicollinearity among predictor variables in that framework. By restricting regularization to $\mathbf{B}$, the factor loading matrix $\mathbf{\Lambda}$ is estimated without shrinkage, which also leaves open the possibility of incorporating L_1 penalties on $\mathbf{\Lambda}$ in future extensions for

Table 1. Results of applying MDSEM to the ACSI data. The six latent variables are as follows: consumer expectations (CE), perceived quality (PQ), perceived value (PV), customer satisfaction (CS), customer complaints (CC), and customer loyalty (CL).

	estimate	s.e.
CE → PQ (b_1)	0.94	0.01
CE → PV (b_2)	−0.31	0.18
CE → CS (b_3)	−0.19	0.07
PQ → PV (b_4)	1.15	0.18
PQ → CS (b_5)	0.97	0.07
PV → CS (b_6)	0.22	0.03
CS → CC (b_7)	−0.60	0.08
CS → CL (b_8)	0.29	0.08
CC → CL (b_9)	−0.70	0.08

exploratory measurement model search and variable selection, as discussed in the final section. We also propose a cross-validation-based procedure for tuning the regularization parameters.

The remainder of this paper is organized as follows. In the next section, we present the proposed method and its mathematical formulation. RMDSEM is formulated as a loss-function minimization problem with L_2 penalty terms. In the same section, we propose an algorithm for solving the minimization problem as well as an algorithm for tuning the penalty parameters. Section 3 reports results from numerical experiments conducted to demonstrate the effectiveness of the proposed method relative to existing approaches, and Section 4 presents an empirical application of the proposed method to real data along with comparisons to existing methods. Finally, the last section provides a general discussion of the findings and their implications.

An R package implementing RMDSEM, along with fully reproducible code for all simulation studies and empirical applications reported in this paper, is publicly available at https://github.com/nyamashita/rmdsem.

2. Proposed Method

In this section, we first present a mathematical formulation of the proposed method as a least-squares minimization problem and derive an iterative algorithm for solving the resulting optimization problem for parameter estimation. We also present an algorithm for simultaneously tuning multiple regularization parameters.

2.1. Formulation

RMDSEM is defined as the problem of minimizing a criterion obtained by adding an L_2 regularization term for $\mathbf{B}$ to the objective function in (2), adopting the same L_2 regularization strategy as in (4). Specifically, RMDSEM is defined as a minimization of

$$l_{RM} = ||\mathbf{X} - (\mathbf{F}\boldsymbol{\Lambda}' + \mathbf{U}\boldsymbol{\Psi}^{1/2})||^2 + \alpha||\mathbf{F} - \mathbf{F}\mathbf{B}||^2 + \lambda_B||\mathbf{B}||^2 \quad (7)$$

subject to the constraints on the factor score matrices (3). The third constraint is required for assuring identifiability of the MDSEM model, while the rest of the constraints are standard assumptions in SEM and factor analysis. As discussed above, multicollinearity in MDSEM manifests as high inter-factor correlations that destabilize the estimation of $\mathbf{B}$. In RMDSEM, we aim to mitigate this issue by imposing L_2 regularization on $\mathbf{B}$ alone.

Here, (7) can be rewritten as

$$\begin{aligned} l_{RM} &= ||\mathbf{X} - (\mathbf{F}^{\#}\mathbf{T}\boldsymbol{\Lambda}' + \mathbf{U}\boldsymbol{\Psi}^{1/2})||^2 + \alpha||\mathbf{F}^{\#}\mathbf{T} - \mathbf{F}^{\#}\mathbf{T}\mathbf{B}||^2 \\ &\quad + \lambda_B||\mathbf{B}||^2 \qquad (8) \\ &= ||\mathbf{X} - \mathbf{Z}^{\#}\tilde{\mathbf{T}}\mathbf{A}||^2 + \alpha||\mathbf{Z}^{\#}\tilde{\mathbf{T}} - \mathbf{Z}^{\#}\tilde{\mathbf{T}}\tilde{\mathbf{B}}||^2 + \lambda_B||\mathbf{B}||^2 \end{aligned}$$

whereas (3) can be rewritten as

$$\mathbf{1}_n'\mathbf{Z}^{\#} = \mathbf{0}_{p+r}', \quad n^{-1}\mathbf{Z}^{\#}\mathbf{Z}^{\#} = \mathbf{I}_{p+r}, \quad \mathrm{diag}(\mathbf{T}'\mathbf{T}) = \mathbf{I}_r \quad (9)$$

where $\mathbf{Z}^{\#} = [\mathbf{F}^{\#}, \mathbf{U}]$, $\mathbf{A} = [\boldsymbol{\Lambda}', \boldsymbol{\Psi}^{1/2}]'$, $\tilde{\mathbf{T}} = \mathrm{bdiag}(\mathbf{T}, \mathbf{I}_p)$, and $\tilde{\mathbf{B}} = \mathrm{bdiag}(\mathbf{B}, \mathbf{I}_p)$. In RMDSEM, given positive constants α and λ_B, we estimate the factor loading matrix $\boldsymbol{\Lambda}$, the uniqueness matrix $\boldsymbol{\Psi}$, the path coefficient matrix $\mathbf{B}$, an orthogonal factor score matrix $\mathbf{Z}^{\#}$, and a nonsingular matrix $\mathbf{T}$. As detailed below, $\mathbf{Z}^{\#}$ is not uniquely determined due to factor-score indeterminacy. In contrast, the covariance matrix between $\mathbf{X}$ and $\mathbf{Z}^{\#}$, denoted by $\mathbf{S}_{\mathbf{XZ}^{\#}}$, is uniquely determined. Therefore, rather than $\mathbf{Z}^{\#}$ itself, we estimate $\mathbf{S}_{\mathbf{XZ}^{\#}}$ instead.

2.2. Parameter Estimation Algorithm

To minimize (7), we first assign randomly generated initial values to each parameter matrix and then iteratively update the parameters using the steps below until the decrease in the objective function converges.

We begin by minimizing (7) with respect to $\mathbf{Z}^{\#}$. The minimizer is obtained via the singular value decomposition

$$n^{-1/2}\mathbf{X}\mathbf{A}\tilde{\mathbf{T}} = \tilde{\mathbf{K}}\tilde{\mathbf{D}}\tilde{\mathbf{L}}' = [\mathbf{K}, \mathbf{K}_{\perp}]\begin{bmatrix}\mathbf{D} & \\ & {}_r\mathbf{O}_r\end{bmatrix}\begin{bmatrix}\mathbf{L}' \\ \mathbf{L}_{\perp}'\end{bmatrix} \quad (10)$$

which yields the optimal $\mathbf{Z}^{\#}$ minimizing (7):

$$\mathbf{Z}^{\#} = n^{1/2}\mathbf{K}\mathbf{L}' + n^{1/2}\mathbf{K}_{\perp}\mathbf{L}_{\perp}'. \quad (11)$$

It is easily verified that $\mathbf{Z}^{\#}$ in (7) satisfies (9) and equivalently (3). Here, $\mathbf{K}$ $(n \times p)$ and $\mathbf{L}$ $((p+r) \times p)$ are matrices of the left and right singular vectors associated with the largest p singular values of $n^{-1/2}\mathbf{X}\mathbf{A}\tilde{\mathbf{T}}$, respectively, and $\mathbf{D}$ is the $p \times p$ diagonal matrix of these singular values in descending order. Moreover, $\mathbf{K}_{\perp}$ $(n \times r)$ and $\mathbf{L}_{\perp}((p+r) \times r)$ are arbitrary column-orthonormal matrices satisfying $\mathbf{K}_{\perp}'\mathbf{K} = {}_r\mathbf{O}_p$ and $\mathbf{L}_{\perp}'\mathbf{L} = {}_{p+r}\mathbf{O}_p$. Because $\mathrm{rank}(n^{-1/2}\mathbf{X}\mathbf{A}\tilde{\mathbf{T}}) \leq p$, the smallest r singular values are equal to zero; consequently, $\mathbf{K}_{\perp}$ and $\mathbf{L}_{\perp}$, which consist of the singular vectors corresponding to these zero singular values, are not unique. Hence, (11) is not unique. This is the same type of factor-score indeterminacy as in MDFA and MDSEM. In contrast, $\mathbf{S}_{\mathbf{XZ}^{\#}}$ is uniquely determined, because

$$\begin{aligned} \mathbf{S}_{\mathbf{XZ}^{\#}} &= n^{-1}\mathbf{X}'\mathbf{Z}^{\#} \qquad (12) \\ &= n^{-1}(n^{1/2}\tilde{\mathbf{K}}\tilde{\mathbf{D}}\tilde{\mathbf{L}}'(\mathbf{A}\tilde{\mathbf{T}})^{+})'(n^{1/2}\mathbf{K}\mathbf{L}' + n^{1/2}\mathbf{K}_{\perp}\mathbf{L}_{\perp}') \qquad (13) \\ &= n^{-1/2}(\mathbf{A}\tilde{\mathbf{T}})^{+'}\tilde{\mathbf{L}}\tilde{\mathbf{D}}\tilde{\mathbf{K}}'\mathbf{K}\mathbf{L}' + n^{-1/2}(\mathbf{A}\tilde{\mathbf{T}})^{+'}\tilde{\mathbf{L}}\tilde{\mathbf{D}}\tilde{\mathbf{K}}'\mathbf{K}_{\perp}\mathbf{L}_{\perp}' \qquad (14) \\ &= n^{-1/2}(\mathbf{A}\tilde{\mathbf{T}})^{+'}\mathbf{L}\mathbf{D}\mathbf{L}' \qquad (15) \end{aligned}$$

where $(\mathbf{A}\tilde{\mathbf{T}})^{+}$ is the Moore-Penrose generalized inverse of $\mathbf{A}\tilde{\mathbf{T}}$ (Yanai et al., 2011).

Next, we minimize (7) with respect to $\boldsymbol{\Lambda}$, $\boldsymbol{\Psi}$, and $\mathbf{B}$. To this end, we rewrite the objective function as follows:

$$\begin{aligned} l_{RM} &= n||\mathbf{M} - \tilde{\mathbf{T}}\mathbf{C}||^2 + \lambda_B||\mathbf{B}||^2 + const. \qquad (16) \\ &\propto ||\mathrm{vec}(\mathbf{M}) - (\mathbf{I}_{2p+r} \otimes \tilde{\mathbf{T}})\mathrm{vec}(\mathbf{C})||^2 + \frac{\lambda_B}{n}||\mathbf{B}||^2 \qquad (17) \\ &\propto ||\mathrm{vec}(\mathbf{M}) - \boldsymbol{\Gamma}\mathbf{c}||^2 + ||\boldsymbol{\Delta}\mathbf{c}||^2 \qquad (18) \end{aligned}$$

where $\mathbf{M} = [\mathbf{S}_{\mathbf{XZ}^{\#}}', \alpha^{1/2}\tilde{\mathbf{T}}]$, and

$$\mathbf{C} = \begin{bmatrix} \boldsymbol{\Lambda}' & \alpha^{1/2}\mathbf{B} & {}_r\mathbf{O}_p \\ \boldsymbol{\Psi}^{1/2} & {}_p\mathbf{O}_r & \alpha^{1/2}\mathbf{I}_p \end{bmatrix} \tag{19}$$

Here, $\mathbf{c}$ is the vector formed by the elements of $\mathbf{C}$ corresponding to the free parameters, which is determined by the pattern of free parameters in $\boldsymbol{\Lambda}$ and $\mathbf{B}$, and $\boldsymbol{\Gamma}$ is the matrix formed by the columns of $\mathbf{I}_{2p+r} \otimes \tilde{\mathbf{T}}$ corresponding to free parameters in vec($\mathbf{C}$). Moreover,

$$\boldsymbol{\Delta} = \mathrm{diag}\left(0, \ldots, 0, 0, \ldots, 0, \sqrt{\frac{\lambda_B}{n}}, \ldots, \sqrt{\frac{\lambda_B}{n}}\right) \tag{20}$$

is a diagonal matrix whose diagonal entries contain the tuning parameters corresponding to each element of $\mathbf{c}$. The first block of zeros corresponds to the free parameters in $\boldsymbol{\Lambda}$, the second to $\boldsymbol{\Psi}^{1/2}$, and the third to $\mathbf{B}$; only $\mathbf{B}$ is subject to L_2 regularization. Using the above notation, the minimizer of (7) with respect to $\mathbf{C}$, equivalently with respect to $\mathbf{c}$, is simply given by

$$\mathbf{c} = (\boldsymbol{\Gamma}'\boldsymbol{\Gamma} + \boldsymbol{\Delta}^2\mathbf{I}_m)^{-1}\boldsymbol{\Gamma}'\mathbf{M} \tag{21}$$

where m denotes the total number of free parameters in $\mathbf{C}$, which equals the dimension of $\mathbf{c}$. Note that the free parameters in $\mathbf{B}$ are divided by $\alpha^{1/2}$, after obtaining the estimate via (21).

Finally, the matrix $\mathbf{T}$ that minimizes (8) is obtained, conditional on the other parameter matrices, by the gradient projection (GP) algorithm (Jennrich, 2002, 2004).

Based on the above discussion, parameter matrices in RMDSEM are estimated by the following iterative algorithm:

Step 1. Assign randomly generated initial values to $\boldsymbol{\Lambda}$, $\boldsymbol{\Psi}^{1/2}$, $\mathbf{B}$, and $\mathbf{T}$.
Step 2. Update $\mathbf{S}_{\mathbf{XZ}}{}^{\#}$ using (15) based on the singular value decomposition of $n^{-1/2}\mathbf{X}\mathbf{A}\tilde{\mathbf{T}}$.
Step 3. Given α and λ_B, update $\boldsymbol{\Lambda}$, $\boldsymbol{\Psi}^{1/2}$, and $\mathbf{B}$ using (21).
Step 4. Given α and λ_B, update $\mathbf{T}$ using the GP algorithm.
Step 5. If the decrease in the function value of (7) is less than or equal to ϵ compared with the one in previous round, terminate the algorithm; otherwise, return to Step 2.

In Steps 2 to 4, each parameter update is guaranteed to not increase the value of the objective function (7). Moreover, by definition (7), we have $l_{RM} \geq 0$, and thus the decrease in the objective function must converge after sufficiently large number of iterations. Because the converged solution is only a local optimum, we run the above algorithm from M different initial values and adopt, as the final solution, the estimate that yields the smallest value of l_{RM} among the resulting M sets of parameter estimates. In the numerical experiments and empirical application reported below, we set the convergence criterion and the number of initializations to $\epsilon = 10^{-6}$ and $M = 100$, respectively. In practice, this multiple-start strategy is particularly important for nonconvex least-squares problems such as (7), because different initializations may lead to substantively different local minima and, consequently, different structural interpretations.

It should be noted that the convergence of the algorithm described above concerns the optimization procedure, and is distinct from the statistical properties of the resulting estimator. The asymptotic properties of the RMDSEM estimator, such as consistency and asymptotic normality, remain to be formally established and constitute an important direction for future theoretical work.

2.3. Algorithm for Hyperparameter Tuning

In RMDSEM, we must choose the values for two hyperparameters, namely α and λ_B, and then run the estimation algorithm described above. In this subsection, we propose an algorithm for identifying suitable values of these hyperparameters.

In this study, we determine the hyperparameters via the K-fold cross-validation procedure described below. Let n_k denote the number of observations in the test set in the k-th fold, where n_k is given by

$$n_k = \begin{cases} \lfloor \frac{n}{K} \rfloor + 1 \text{ if } k \leq n \bmod K \\ \lfloor \frac{n}{K} \rfloor \text{ otherwise} \end{cases} \tag{22}$$

with $n_k \geq n_l$ $(k > l)$. For the k-th fold, let $\mathbf{X}^{(k)}$ $(n_k \times p)$ denote the data matrix obtained by randomly selecting n_k observations from the full data matrix $\mathbf{X}$ $(n \times p)$, and let $\mathbf{X}^{(-k)}$ $((n - n_k) \times p)$ denote the data matrix consisting of the remaining observations. Using $\mathbf{X}^{(-k)}$ as the training data matrix, we apply RMDSEM to $\mathbf{X}^{(-k)}$ under a given set of $\{\alpha, \lambda_B\}$, and denote the resulting parameter estimates by $\hat{\boldsymbol{\theta}}^{(-k)} = \{\hat{\boldsymbol{\Lambda}}^{(-k)}, \hat{\mathbf{B}}^{(-k)}, \hat{\boldsymbol{\Psi}}^{(-k)}, \hat{\mathbf{T}}^{(-k)}\}$. After repeating this procedure for $k = 1, \ldots, K$, we compute the cross-validation error for $\{\alpha, \lambda_B\}$ as

$$CV(\alpha, \lambda_B) = \frac{1}{K}\sum_{k=1}^{K} |\mathbf{S}_{\mathbf{X}}^{(k)} - \Sigma(\hat{\boldsymbol{\theta}}^{(-k)})|^2 \tag{23}$$

with $\mathbf{S}_{\mathbf{X}}^{(k)} = n_k^{-1}\mathbf{X}^{(k)\prime}\mathbf{X}^{(k)}$ the sample covariance matrix of the k-th test set, and $\Sigma(\hat{\boldsymbol{\theta}}^{(-k)})$ the model-implied covariance matrix computed from the training-set parameter estimates, given by

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \mathbf{P}(\mathbf{I}_{p+r} - \mathbf{Q})^{-1'}\begin{bmatrix} \boldsymbol{\Psi} & \\ & \boldsymbol{\Sigma}_f \end{bmatrix}(\mathbf{I}_{p+r} - \mathbf{Q})^{-1}\mathbf{P}' \tag{24}$$

where

$$\mathbf{Q} = \begin{bmatrix} {}_p\mathbf{O}_p & {}_p\mathbf{O}_r \\ \boldsymbol{\Lambda}' & \mathbf{B} \end{bmatrix}, \quad \mathbf{P} = [\mathbf{I}_p, \ {}_r\mathbf{O}_r] \tag{25}$$

and $\boldsymbol{\Sigma}_f = n^{-1}(\mathbf{F} - \mathbf{F}\mathbf{B})'(\mathbf{F} - \mathbf{F}\mathbf{B}) = (\mathbf{I}_r - \mathbf{B})'\mathbf{T}'\mathbf{T}(\mathbf{I}_r - \mathbf{B})$ is the covariance matrix of the residual of structural model. $\boldsymbol{\Sigma}(\boldsymbol{\theta})$ is equivalent to the covariance structure used in conventional SEM procedures based on the reticular action model (McArdle, 1980). Since $\mathbf{B}$ as well as $\mathbf{T}$ appears in $\boldsymbol{\Sigma}(\boldsymbol{\theta})$, this criterion appropriately incorporates the structural parameters of interest into the hyperparameter selection procedure. This criterion is based on the same sufficient

statistic $\mathbf{S_X}$ as the RMDSEM objective function, providing a coherent basis for cross-validation.

Hwang (2009) proposed an analogous K-fold cross-validation approach for selecting hyperparameters in RGSCA.

In the numerical simulations and empirical application reported in the subsequent sections, we tuned the hyperparameters using 5-fold cross validation with $K = 5$.

2.4. Computing Standard Error via Bootstrapping

One of the benefits of using a regularization technique is to stabilize the estimation process in MDSEM. To evaluate how stable the estimation is, we used the bootstrapping method (Efron, 1982) and computed standard errors, as a measure of instability. In the following sections, we used bootstrapping with 1000 replications. The standard deviation of 1000 estimates is used as a standard error, and it is used to construct bootstrap confidence intervals.

3. Numerical Simulation

We conducted a numerical simulation study to evaluate the effectiveness of the proposed method compared to existing procedures, and to clarify its empirical properties. The current section reports the design and results of the simulation.

3.1. Design

In the simulation, we first generated artificial data matrices $\mathbf{X}$, which are modeled by SEM procedures including RMDSEM, in the following manner. Using the path diagram in Figure 1, which includes eight observed variables $x_1, \ldots, x_8$ and four latent variables $f_1, \ldots, f_4$, we constructed a covariance structure $\boldsymbol{\Sigma}$. We then generated n data vectors with p-dimensionality noted as $\mathbf{x}_{(1)}, \ldots, \mathbf{x}_{(n)}$ from the p-dimensional multivariate normal distribution $N_p(\mathbf{0}_p, \boldsymbol{\Sigma})$, and formed the data matrix $\mathbf{X} = [\mathbf{x}_{(1)}, \ldots, \mathbf{x}_{(n)}]'$.

The population covariance matrix $\boldsymbol{\Sigma}$ was computed following the derivation of covariance structures;

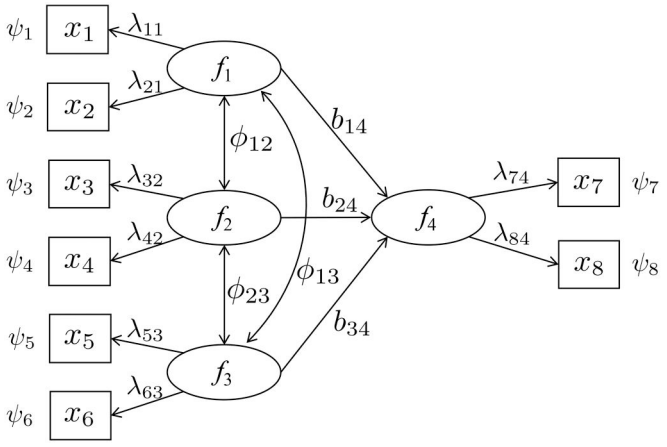

Figure 1. Path diagram used in the numerical simulation with factor loadings $\lambda_{11}, \ldots, \lambda_{48}$, path coefficients $b_{14}, \ldots, b_{34}$, factor residual covariances $\phi_{12}, \ldots, \phi_{23}$, and uniqueness $\psi_1, \ldots, \psi_8$.

$$\boldsymbol{\Sigma} = \boldsymbol{\Lambda}_T' \boldsymbol{\Phi}_T \boldsymbol{\Lambda}_T' + \boldsymbol{\Psi}_T \tag{26}$$

where $\boldsymbol{\Psi}_T$ and $\boldsymbol{\Lambda}_T$ denote the true values of $\boldsymbol{\Psi}$ and $\boldsymbol{\Lambda}$, respectively. The inter-factor correlation matrix $\boldsymbol{\Phi}_T$ is given by

$$\boldsymbol{\Phi}_T = (\mathbf{I}_r - \mathbf{B}_T)^{-1'} \boldsymbol{\Sigma}_f (\mathbf{I}_r - \mathbf{B}_T)^{-1} \tag{27}$$

where $\mathbf{B}_T$ and $\boldsymbol{\Sigma}_f$ denote the true value of $\mathbf{B}$ and the covariance matrix of the factor score residuals $\mathbf{E}_f = \mathbf{F}(\mathbf{I}_r - \mathbf{B})$ implied by the second term of (7), respectively. The true value matrices were specified as

$$\boldsymbol{\Lambda}_T = \begin{bmatrix} \lambda_{11} & 0 & 0 & 0 \\ \lambda_{21} & 0 & 0 & 0 \\ 0 & \lambda_{32} & 0 & 0 \\ 0 & \lambda_{42} & 0 & 0 \\ 0 & 0 & \lambda_{53} & 0 \\ 0 & 0 & \lambda_{63} & 0 \\ 0 & 0 & 0 & \lambda_{74} \\ 0 & 0 & 0 & \lambda_{84} \end{bmatrix}, \tag{28}$$

$$\mathbf{B}_T = \begin{bmatrix} 0 & 0 & 0 & b_{14} \\ 0 & 0 & 0 & b_{24} \\ 0 & 0 & 0 & b_{34} \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{29}$$

$$\boldsymbol{\Sigma}_f = \begin{bmatrix} 1 & \phi_{12} & \phi_{13} & 0 \\ & 1 & \phi_{23} & 0 \\ & & 1 & 0 \\ & & & 1 \end{bmatrix} \tag{30}$$

and $\boldsymbol{\Psi}_T$ is given by $\boldsymbol{\Psi}_T = \mathrm{diag}(\mathbf{I}_8 - \boldsymbol{\Lambda}_T \boldsymbol{\Phi}_T \boldsymbol{\Lambda}_T')$ to keep the diagonal of $\boldsymbol{\Sigma}$ being unity. Here, the diagonal elements of $\boldsymbol{\Sigma}_f$ are set to unity, normalizing the residual variances to one, and the off-diagonal elements ϕ_{ij} represent the residual correlations among factor scores. Because the path coefficient matrix $\mathbf{B}$ in (29) has nonzero elements only in the first three rows of the fourth column, the matrix $\mathbf{I}_4 - \mathbf{B}$ differs from the identity only in its fourth column. Consequently, considering (27) for the first three factors, the residual covariances in $\boldsymbol{\Sigma}_f$ directly equal the inter-factor correlations among factors 1–3.

It should be noted that (26) is equivalent to (24) used in cross-validation, since

$$(\mathbf{I}_{p+r} - \mathbf{Q})^{-1} \mathbf{P}' = \begin{bmatrix} \mathbf{I}_p \\ (\mathbf{I}_r - \mathbf{B})^{-1} \boldsymbol{\Lambda}' \end{bmatrix} \tag{31}$$

and therefore we have

$$\boldsymbol{\Sigma}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{I}_p & \boldsymbol{\Lambda}(\mathbf{I}_r - \mathbf{B})^{-1'} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Psi} & \\ & \boldsymbol{\Sigma}_f \end{bmatrix} \begin{bmatrix} \mathbf{I}_p \\ (\mathbf{I}_r - \mathbf{B})^{-1} \boldsymbol{\Lambda}' \end{bmatrix} \tag{32}$$

$$= \boldsymbol{\Lambda}(\mathbf{I}_r - \mathbf{B})^{-1'} \boldsymbol{\Sigma}_f (\mathbf{I}_r - \mathbf{B})^{-1} \boldsymbol{\Lambda}' + \boldsymbol{\Psi}. \tag{33}$$

The values of the elements of true-value matrices were set according to the experimental conditions described in Table 2. Patterns A, B, and C correspond to weak, moderate, and strong inter-factor correlations among the first three factors, respectively. In particular, in Pattern C, strong multicollinearity among factors is expected to make estimation of path coefficients highly unstable, that is, to increase

Table 2. Three patterns of true parameter values in the numerical simulation.

pattern	$\lambda_{11}, \lambda_{21}$	$\phi_{12}, \phi_{13}, \phi_{23}$
A	0.7	0.4
B	0.8	0.6
C	0.9	0.9

bias in the estimates relative to the true values and/or inflate standard errors.

For parameters not listed in Table 2, we set, for all patterns, $\lambda_{32} = \lambda_{42} = \lambda_{53} = \lambda_{63} = \lambda_{74} = \lambda_{84} = 0.8$, $b_{14} = b_{34} = 0.4$, and $b_{24} = -0.4$.

For each of the three patterns, we considered three sample-size conditions, $n = 50, 100, 200$, and generated 100 data matrices $\mathbf{X}$ for each of the resulting $3 \times 3 = 9$ conditions. For each generated dataset, we applied the proposed method (RMDSEM) and, for comparison, MDSEM, SEM based on maximum likelihood estimation (ML-SEM), and RGSCA, and obtained parameter estimates. For RMDSEM/MDSEM/RGSCA, the tuning parameters were selected using the cross-validation procedure described above and Hwang (2009). For all methods, parameter estimation was performed under the correctly specified model, identical to the data-generating model. As a measure of recovery accuracy relative to the true values, we computed the mean squared error (MSE) of the parameter estimates. In addition, we computed standard errors of the parameter estimates using the bootstrap method. To ensure a fair comparison, all competing methods were provided with the same model structure and were evaluated under identical data-generating conditions; thus, differences in performance can be attributed primarily to the estimation principles and regularization mechanisms rather than to model misspecification. We used 100 replications per condition, following the convention adopted in prior simulation studies of regularized SEM procedures (Hwang, 2009), and confirmed that the conclusions were virtually unchanged when 200 replications were used, showing that 100 replications are sufficient for the simulation.

3.2. Results

Table 3 reports the medians of the mean squared errors (MSEs) for the factor loading matrix $\mathbf{\Lambda}$ and the path coefficient matrix $\mathbf{B}$. Because $\mathbf{\Lambda}$ is not subject to L_2 regularization in RMDSEM, the MSEs of $\mathbf{\Lambda}$ are similar between MDSEM and RMDSEM across all conditions, confirming that the absence of shrinkage on factor loadings does not meaningfully affect loading recovery.

Turning to the MSE of $\mathbf{B}$, the results reveal a clear pattern across the three conditions. In Patterns A and B, where inter-factor correlations are weak to moderate, regularization introduces a small amount of bias in RMDSEM relative to MDSEM; in some conditions RMDSEM yields a slightly larger MSE of $\mathbf{B}$ than MDSEM (e.g., Pattern A, $n = 100$: 0.043 vs. 0.019). Nevertheless, the magnitude of these differences is negligible in absolute terms, suggesting that cross-validation-based tuning effectively controls the shrinkage effect and keeps the induced bias at a practically

Table 3. Medians of mean squared errors obtained in each of experimental conditions by MDSEM, RMDSEM, and SEM by maximum likelihood estimation (ML-SEM).

		MDSEM		RMDSEM		ML-SEM		RGSCA
pattern	n	$\mathbf{\Lambda}$	$\mathbf{B}$	$\mathbf{\Lambda}$	$\mathbf{B}$	$\mathbf{\Lambda}$	$\mathbf{B}$	$\mathbf{B}$
A	50	0.026	0.068	0.026	0.056	0.041	0.054	0.121
	100	0.015	0.019	0.015	0.043	0.018	0.022	0.123
	200	0.008	0.010	0.008	0.031	0.007	0.009	0.120
B	50	0.022	0.155	0.022	0.079	0.031	0.104	0.126
	100	0.012	0.044	0.012	0.054	0.014	0.036	0.131
	200	0.006	0.024	0.006	0.045	0.006	0.019	0.134
C	50	0.018	17.727	0.020	0.120	0.027	0.725	0.137
	100	0.010	41.156	0.010	0.110	0.011	0.484	0.134
	200	0.006	48.075	0.005	0.101	0.005	0.201	0.137

inconsequential level. Under Pattern C, where factors are strongly correlated and severe multicollinearity is expected, the MSE of $\mathbf{B}$ under MDSEM increases explosively with sample size (e.g., $n = 50$: 17.73; $n = 200$: 48.08), indicating that MDSEM's structural parameter estimation collapses under severe multicollinearity. This result indicates that, while MDSEM is not fundamentally flawed as a modeling framework, its structural parameter estimation deteriorates rapidly in the presence of inter-factor correlations. ML-SEM also shows substantially inflated MSEs under Pattern C (e.g., $n = 50$: 0.725), though less severely than MDSEM. In contrast, RMDSEM suppresses the MSE of $\mathbf{B}$ to a practically manageable level across all conditions (e.g., Pattern C, $n = 50$: 0.120; $n = 200$: 0.101), demonstrating the effectiveness of L_2 regularization on $\mathbf{B}$. It is also noteworthy that RMDSEM achieves smaller MSEs for $\mathbf{B}$ than RGSCA across all conditions (e.g., Pattern A, $n = 50$: 0.056 vs. 0.121), suggesting that the matrix-decomposition-based regularization framework adopted in RMDSEM is particularly well-suited for stabilizing structural parameter estimation in reflective models.

A similar pattern is observed for the medians of the standard errors of the path coefficients shown in Table 4. MDSEM exhibits consistently large standard errors across all conditions, with especially large values when the sample size is small or when factors are strongly correlated. ML-SEM yields smaller standard errors than MDSEM, but still shows inflated standard errors under Pattern C. In contrast, RMDSEM maintains small standard errors even in small-sample settings or when factors are highly correlated, and yields smaller standard errors than ML-SEM across all conditions. Among the four methods, RGSCA achieves the smallest standard errors across all conditions, which may be attributable to the uniqueness of factor scores in RGSCA suppressing multicollinearity in the estimation of $\mathbf{B}$ to some extent. Among the three methods that share factor-score indeterminacy—MDSEM, RMDSEM, and ML-SEM—RMDSEM achieves by far the most stable estimation, as evidenced by its consistently small standard errors relative to the other two.

The above results indicate that, when factors are strongly correlated, existing methods such as MDSEM and ML-SEM yield unstable path coefficient estimates due to multicollinearity, manifested as large bias and inflated standard errors. In contrast, RMDSEM achieves stable estimation

Table 4. Medians of standard errors of path coefficients obtained in each of experimental conditions by MDSEM, RMDSEM, ML-SEM, and RGSCA.

		MDSEM			RMDSEM		
pattern	n	b_{14}	b_{24}	b_{34}	b_{14}	b_{24}	b_{34}
A	50	6.855	7.113	9.551	0.139	0.123	0.158
	100	4.502	3.018	5.502	0.081	0.068	0.094
	200	1.523	0.946	1.936	0.079	0.072	0.088
B	50	7.131	5.506	8.263	0.103	0.109	0.135
	100	5.433	3.923	7.080	0.072	0.071	0.088
	200	3.882	2.409	4.625	0.079	0.084	0.106
C	50	7.060	6.541	9.001	0.095	0.102	0.126
	100	5.376	6.152	9.419	0.057	0.068	0.080
	200	5.245	5.938	10.709	0.042	0.050	0.062
		ML-SEM			RGSCA		
pattern	n	b_{14}	b_{24}	b_{34}	b_{14}	b_{24}	b_{34}
A	50	0.201	0.193	0.224	0.061	0.042	0.066
	100	0.149	0.139	0.150	0.060	0.038	0.066
	200	0.112	0.104	0.112	0.062	0.039	0.058
B	50	0.303	0.275	0.297	0.065	0.036	0.058
	100	0.183	0.189	0.199	0.067	0.030	0.053
	200	0.138	0.139	0.150	0.057	0.025	0.052
C	50	1.177	1.490	1.434	0.062	0.035	0.041
	100	0.789	0.860	0.920	0.058	0.031	0.037
	200	0.504	0.474	0.598	0.050	0.020	0.030

through L_2 regularization on $\mathbf{B}$, with smaller MSEs than RGSCA and smaller standard errors than both MDSEM and ML-SEM across all conditions.

Although RGSCA was included as a comparator in this simulation, it should be noted that RGSCA is a component-based framework originally designed for formative models, and the present simulation was conducted under a reflective model structure. Accordingly, the results for RGSCA may not generalize to settings for which RGSCA was originally intended, and the comparison should be interpreted with caution.

An additional advantage of MDSEM/RMDSEM over ML-SEM is that the former methods do not produce improper solutions. For example, under Pattern A, which yielded the highest frequency of improper solutions, negative uniqueness estimates were observed in ML-SEM in 60, 26, and 2 out of 100 replications for $n = 50,\ 100,$ and 200, respectively. Moreover, in the remaining conditions, ML-SEM produced improper solutions in approximately 20% of replications. That such a high frequency of improper solutions arises despite correct model specification represents a serious practical limitation of ML-SEM. In contrast, MDSEM/RMDSEM did not produce any improper solutions, as is evident from the fact that error variances are computed as $(\boldsymbol{\Psi}^{1/2})^2$.

4. Real Data Examples

4.1. American Customer Satisfaction Index (ACSI) Data

We present an application of the proposed method to the dataset analyzed by Hwang (2009) as an empirical example of RGSCA. Following Hwang (2009), we adopted a simple measurement structure with no cross-loadings, in which each observed variable loads on exactly one factor. This model specification was chosen to ensure direct comparability with the RGSCA results reported in Hwang (2009).

Table 5. Estimates of component loadings $\boldsymbol{\Lambda}$ and uniquenesses $\boldsymbol{\psi}$ obtained by GSCA, RGSCA, MDSEM, and RMDSEM applied to ACSI dataset. The GSCA and RGSCA results were directly reproduced from Hwang (2009).

		GSCA	RGSCA	MDSEM		RMDSEM	
		$\boldsymbol{\Lambda}$	$\boldsymbol{\Lambda}$	$\boldsymbol{\Lambda}$	$\boldsymbol{\psi}$	$\boldsymbol{\Lambda}$	$\boldsymbol{\psi}$
Customer Expectations (CE)	z_1	0.96	0.96	0.92	0.13	0.92	0.13
	z_2	0.97	0.97	0.96	0.07	0.96	0.07
	z_3	0.94	0.94	0.93	0.12	0.93	0.12
Perceived Quality (PQ)	z_4	0.97	0.98	0.97	0.04	0.97	0.04
	z_5	0.98	0.97	0.97	0.05	0.97	0.05
	z_6	0.97	0.97	0.94	0.10	0.94	0.10
Perceived Value (PV)	z_7	0.99	0.99	0.94	0.11	0.94	0.11
	z_8	0.99	0.99	0.99	0.00	0.99	0.00
Customer Satisfaction (CS)	z_9	0.98	0.98	0.99	0.00	1.00	0.00
	z_{10}	0.98	0.98	0.97	0.04	0.98	0.04
	z_{11}	0.95	0.95	0.92	0.13	0.92	0.13
Customer Complaints (CC)	z_{12}	1.00	1.00	0.79	0.37	0.80	0.35
Customer Loyalty (CL)	z_{13}	0.98	0.98	0.95	0.09	0.95	0.09
	z_{14}	0.98	0.98	0.98	0.03	0.98	0.03

Hwang (2009) reports the covariance matrix of the ACSI data, introduced in the first section of the article, and we applied the SEM procedures including RMDSEM to the covariance matrix $\mathbf{S}_{\mathbf{X}}^{ACSI}$. However, the hyperparameter tuning procedure described above and the construction of confidence intervals based on the bootstrapping require the original data matrix $\mathbf{X}^{ACSI}$. Therefore, in this application, we generated $\mathbf{X}^{ACSI}$ from 14-dimensional multivariate normal distribution $N_{14}(\mathbf{0}_{14}, \mathbf{S}_{\mathbf{X}}^{ACSI})$ with 152 observations as reported in Hwang (2009). Using 152×14 matrix $\mathbf{X}^{ACSI}$, we ran the cross-validation procedure presented in the previous section to tune the hyperparameters $\{\alpha,\ \lambda_B\}$ and applied MDSEM and RMDSEM with the selected values of the hyperparameters. The optimal hyperparameters were specified as $\{\alpha,\ \lambda_B\} = \{0.01, 0.10\}$, indicating that L_2 regularization of $\mathbf{B}$ is necessary. The GSCA and RGSCA results were directly reproduced from Hwang (2009).

Table 5 reports the estimated factor loadings and uniquenesses obtained by four SEM procedures, and Table 6 reports the estimated path coefficients and their standard errors. Consistent with the theoretical motivation of RMDSEM, the factor loading estimates are identical between MDSEM and RMDSEM, as $\boldsymbol{\Lambda}$ is not subject to regularization. In contrast, turning to the path coefficient estimates, MDSEM failed to produce reasonable coefficients for multiple paths, implying unreasonable causal relations among latent variables as reported in the first section. A similar tendency can also be observed in the GSCA results. The estimated inter-factor correlations under MDSEM support this interpretation: several factor pairs exhibit extremely high correlations (e.g., $r_{24} = .988$, $r_{12} = .947$, $r_{34} = .909$ where r_{ij} denotes the correlation between the i-th and j-th factors), indicating that severe multicollinearity is present in this dataset.

We interpret these inappropriate relations not as genuine causal effects, but as spurious relations induced by high correlations among variables, that is, by multicollinearity produced in the estimation process of the SEM procedures.

Such spurious relations are shown to be attenuated or removed via newly-introduced L_2 regularization with $\lambda_B = 0.01$, which was selected as optimal by the cross-validation

Table 6. Estimates of path coefficients and their standard errors obtained by GSCA, RGSCA, MDSEM, and RMDSEM. The GSCA and RGSCA results were directly reproduced from Hwang (2009).

	GSCA		RGSCA		MDSEM		RMDSEM	
	est.	s.e.	est.	s.e.	est.	s.e.	est.	s.e.
CE → PQ (b_1)	0.91	0.02	0.83	0.01	0.94	0.01	0.86	0.01
CE → PV (b_2)	−0.06	0.13	0.18	0.07	−0.31	0.18	0.14	0.06
CE → CS (b_3)	0.00	0.05	0.17	0.03	−0.19	0.07	0.15	0.03
PQ → PV (b_4)	0.88	0.14	0.59	0.08	1.15	0.18	0.66	0.06
PQ → CS (b_5)	0.75	0.05	0.50	0.03	0.97	0.07	0.52	0.03
PV → CS (b_6)	0.27	0.03	0.31	0.02	0.22	0.03	0.32	0.02
CS → CC (b_7)	−0.44	0.09	−0.41	0.07	−0.60	0.08	−0.53	0.07
CS → CL (b_8)	0.49	0.05	0.46	0.04	0.29	0.08	0.32	0.06
CC → CL (b_9)	−0.44	0.05	−0.42	0.04	−0.70	0.08	−0.61	0.06

procedure. Indeed, under RGSCA and RMDSEM, these unreasonable relations disappear, and signs of all coefficients are reasonably estimated. The RMDSEM path coefficient estimates and their sign pattern perfectly match those of RGSCA. Moreover, standard errors of the RMDSEM estimates are smaller than those of MDSEM, indicating that estimation is considerably stabilized by regularization as RGSCA does. Taken together, these results empirically demonstrate (i) that unreasonable paths arising from multicollinearity can occur in MDSEM as well as in GSCA, and (ii) that this problem can be mitigated by the proposed method with L_2 regularization. This example also highlights a practical advantage of regularization in structural modeling: beyond improving numerical conditioning, shrinkage can suppress overinterpretation of unstable sign flips that are artifacts of collinearity rather than substantive causal mechanisms.

4.2. Intelligence Test Data

Next, we present an application to the data of Holzinger & Swineford (1939). Because the observations × variables data matrix is available for this dataset, we can perform hyperparameter tuning by cross validation and construct confidence intervals using the bootstrap method.

To induce multicollinearity in a controlled manner, we modified the dataset as follows. We randomly generated a small perturbation from a uniform distribution $U[0.2, 0.3]$ and randomly flipped its sign. Then, we added this perturbation to the first observed variable Visual to create an additional variable Visual2 with small measurement error relative to Visual. This newly constructed variable was used as an indicator of the second factor. The first and second factors are expected to become highly correlated, and thus estimation of path coefficients is expected to become unstable due to multicollinearity. Applying the hyperparameter tuning procedure described above yielded $\{\alpha, \lambda_B\} = \{0.01, 0.10\}$ as the optimal set of values.

Under the above setting, the parameter estimates obtained from RMDSEM are reported in Tables (7) and (8). Consistent with the RMDSEM formulation, $\mathbf{\Lambda}$ is not regularized, and the loading estimates are therefore similar between MDSEM and RMDSEM.

In contrast, the results for the path coefficients differ significantly in RMDSEM, where the path coefficients are regularized with $\lambda_B = 0.10$. First, the path coefficient estimates obtained by MDSEM have extremely large absolute values, accompanied by correspondingly large standard errors. More seriously, the path from verbal to math is estimated to be considerably negative, yielding an unreasonable structural relation, as observed in the previous application example. That is, estimation of the path coefficients in MDSEM is strongly affected by the multicollinearity that was deliberately introduced into the data through preprocessing, leading to highly unstable estimates.

In contrast, the estimates obtained by RMDSEM exhibit dramatically smaller standard errors, indicating substantial stabilization of estimation. The corresponding confidence intervals are $[0.07, 0.23]$ and $[0.07, 0.19]$, respectively, and thus these paths are appropriately retained in the subsequent model modification process. In addition, all estimated coefficients are positive, allowing for a substantively reasonable interpretation of the causal relations among the factors.

These results demonstrate that RMDSEM effectively stabilizes structural parameter estimation under severe multicollinearity, preventing spurious path deletion and yielding interpretable causal structures that are not attainable with unregularized MDSEM.

Table 7. Estimates of factor loading ($\mathbf{\Lambda}$) and uniqueness ($\boldsymbol{\psi}$) obtained by MDSEM and RMDSEM applied to modified intelligence test dataset by Holzinger & Swineford (1939).

		MDSEM		RMDSEM	
		$\mathbf{\Lambda}$	$\boldsymbol{\psi}$	$\mathbf{\Lambda}$	$\boldsymbol{\psi}$
Vision	Visual	0.94	0.06	0.97	0.00
	Cubes	0.33	0.88	0.33	0.87
	Lozenges	0.44	0.78	0.45	0.78
Verbal	Visual2	0.58	0.54	0.59	0.54
	Paragraph	0.51	0.61	0.52	0.60
	Sentence	0.56	0.58	0.56	0.58
	Word	0.54	0.68	0.54	0.69
Math.	Addition	0.70	0.50	0.69	0.51
	Counting	0.72	0.46	0.72	0.46
	Discrimination	0.93	0.07	0.94	0.05

Table 8. Estimates of path coefficient and its standard error obtained by MDSEM and RMDSEM applied to modified intelligence test dataset by Holzinger & Swineford (1939).

	MDSEM		RMDSEM	
	est.	s.e.	est	s.e.
Vision → Math.	17.11	4.51	0.15	0.04
Verbal → Math.	−16.70	4.50	0.13	0.03

5. Discussion

This study proposed regularized matrix decomposition structural equation modeling (RMDSEM), an L_2-regularized extension of MDSEM designed to address multicollinearity in estimating path coefficients. By adding L_2 penalty exclusively to $\mathbf{B}$ within a matrix-decomposition framework, RMDSEM improves estimation stability while maintaining interpretability. We derived an iterative estimation algorithm with cross-validation-based tuning and demonstrated its performance through simulations and empirical applications.

Simulation results showed that when inter-factor correlations were weak, RMDSEM's estimates were found to be more biased than those of other methods, but all biases were small, indicating no substantial drawback of regularization. When correlations were moderate or strong, MDSEM and ML-SEM produced highly biased estimates with inflated standard errors, especially in small samples. RMDSEM effectively stabilized estimation with much smaller bias, yielding better performance than RGSCA while retaining MDSEM's advantages, such as guaranteed nonnegative uniquenesses.

In empirical analyses, RMDSEM produced interpretable path structures and reasonable coefficient signs even under strong multicollinearity, where MDSEM gave unstable and implausible estimates. Bootstrap standard errors were consistently and significantly smaller in RMDSEM, confirming that regularization improved numerical stability. Overall, these results indicate that RMDSEM achieves robustness against multicollinearity without compromising the theoretical coherence of the SEM framework.

It is worth noting that RMDSEM simultaneously estimates $\mathbf{S}_{XZ^{\#}}$ alongside the structural parameters $\boldsymbol{\Lambda}$ and $\mathbf{B}$, which differs from the conventional sequential approach in which factor scores and structural parameters are estimated in separate steps. This integrated estimation is a deliberate design feature of MDSEM: by optimizing a single unified criterion, the interdependence between the measurement and structural models is appropriately taken into account. Moreover, prior work has demonstrated that MDFA and MDSEM empirically yield solutions closely comparable to those of covariance-based estimators (Terada, 2025; Yamashita, 2024), providing empirical support for the validity of this approach.

Several limitations should be noted. First, the RMDSEM objective is nonconvex, so the algorithm finds local minima. Multiple random starts mitigate this, but global optimality is not guaranteed. Second, estimation quality depends on hyperparameter selection. The cross-validation procedure works well but is computationally intensive and may not optimize all inferential goals. Extending penalization to other parameters may further improve performance when instability arises elsewhere in the model. Finally, the method was tested only with continuous normal data and correctly specified models; its robustness to non-normality, categorical data, or model misspecification remains unexamined.

Future work may focus on four directions. First, extending RMDSEM to incorporate L_1 or elastic-net penalties offers a promising direction. Applying L_1 regularization to $\boldsymbol{\Lambda}$ would enable sparse estimation of factor loadings, facilitating exploratory measurement model search and variable selection—a capability that is not available in the current L_2-only formulation. Such an extension would be conceptually analogous to penalized SEM approaches for sparse cross-loading estimation (Jacobucci et al., 2016) and would broaden the practical applicability of RMDSEM in high-dimensional settings. Second, faster optimization and more efficient hyperparameter searches are desirable for large models. Third, applying RMDSEM to longitudinal, multilevel, or non-normal data will clarify its practical scope. Fourth, the asymptotic properties of RMDSEM remain to be established. Terada (2025) recently derived consistency and asymptotic normality of the MDFA estimator by formulating it as a semiparametric profile likelihood estimator, providing a theoretical foundation for the matrix decomposition approach to factor analysis. Since RMDSEM is an extension of MDSEM, which is itself an extension of MDFA, the framework developed by Terada (2025) may offer a promising starting point for deriving analogous theoretical results for MDSEM and RMDSEM. It should be noted, however, that the L_2 regularization introduced in RMDSEM is not intended to achieve asymptotic unbiasedness; rather, its purpose is to improve estimation stability through a bias–variance tradeoff, as is common with ridge-type estimators in general. Establishing formal asymptotic guarantees for RMDSEM, including the behavior of estimators as the regularization strength approaches zero, remains an important direction for future research.

Overall, RMDSEM provides a promising approach for stable and interpretable SEM estimation under multicollinearity while maintaining the generality and flexibility of the matrix decomposition framework.

Acknowledgement

The author is deeply grateful to the two anonymous reviewers for their careful reviews and constructive comments for improving the quality of the paper.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Use of Generative Artificial Intelligence

The authors used ChatGPT-5 and Claude Sonnet 4.6 for inspecting spelling and grammatical errors during the writing process, and preparing code for public release in the revision process, but not for generating or revising any scientific content. The authors have independently verified the final text and remain fully responsible for the accuracy and integrity of the work.

Code Availability

The R package `RMDSEM` and all code used to reproduce the simulation and empirical results reported in this paper are publicly available at

References

Adachi, K., & Trendafilov, N. T. (2018). Some mathematical properties of the matrix decomposition solution in factor analysis. *Psychometrika*, *83*, 407–424. https://doi.org/10.1007/s11336-017-9600-y

Yamashita, N. (2024). Matrix Decomposition Approach for Structural Equation Modeling as an Alternative to Covariance Structure Analysis and Its Theoretical Properties. *Structural Equation Modeling: A Multidisciplinary Journal*, *31*, 817–834. https://doi.org/10.1080/10705511.2024.2342381

Yanai, H., Takeuchi, K., & Takane, Y. (2011). *Projection Matrices, Generalized Inverse Matrices, and Singular Value Decomposition*. Springer New York.

Bentler, P. M. (1980). Multivariate analysis with latent variables: Causal modeling. *Annual Review of Psychology*, *31*, 419–456. https://doi.org/10.1146/annurev.ps.31.020180.002223

Bentler, P. M. (1986). Structural modeling and Psychometrika: An historical perspective on growth and achievements. *Psychometrika*, *51*, 35–51. https://doi.org/10.1007/BF02293997

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, *53*, 605–634. https://doi.org/10.1146/annurev.psych.53.100901.135239

Cho, G., Choi, J. Y., Sarstedt, M., & Hwang, H. (2025). Regularized structural equation modeling with both factors and components. *Structural Equation Modeling: A Multidisciplinary Journal*, *32*, 1057–1068. https://doi.org/10.1080/10705511.2025.2537946

Choi, J. Y., & Hwang, H. (2020). Bayesian generalized structured component analysis. *The British Journal of Mathematical and Statistical Psychology*, *73*, 347–373. https://doi.org/10.1111/bmsp.12166

Dijkstra, T. K., & Henseler, J. (2015). Consistent and asymptotically normal PLS estimators for linear structural equations. *Computational Statistics & Data Analysis*, *81*, 10–23. https://doi.org/10.1016/j.csda.2014.07.008

Efron, B. (1982). *The jackknife, the bootstrap and other resampling plans*. Society for industrial and applied mathematics.

Fornell, C., Johnson, M. D., Anderson, E. W., Cha, J., & Bryant, B. E. (1996). The American customer satisfaction index: Nature, purpose, and findings. *Journal of Marketing*, *60*, 7–18. https://doi.org/10.1177/002224299606000403

Holzinger, K., & Swineford, F. (1939). *A study in factor analysis: The stability of a bifactor solution*. Supplementary Educational Monograph.

Hoyle, R. H. (Ed.). (2023). *Handbook of structural equation modeling* (2nd ed.). Guilford press.

Hwang, H. (2009). Regularized generalized structured component analysis. *Psychometrika*, *74*, 517–530. https://doi.org/10.1007/s11336-009-9119-y

Hair, J. F., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2017). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage.

Hwang, H., & Takane, Y. (2004). Generalized structured component analysis. *Psychometrika*, *69*, 81–99. https://doi.org/10.1007/BF02295841

Hwang, H., & Takane, Y. (2014). *Generalized structured component analysis: A component-based approach to structural equation modeling*. CRC Press.

Hwang, H., Cho, G., Jung, K., Falk, C. F., Flake, J. K., Jin, M. J., & Lee, S. H. (2021). An approach to structural equation modeling with both factors and components: Integrated generalized structured component analysis. *Psychological Methods*, *26*, 273–294. https://doi.org/10.1037/met0000336

Hwang, H., Takane, Y., & Jung, K. (2017). Generalized structured component analysis with uniqueness terms for accommodating measurement error. *Frontiers in Psychology*, *8*, 2137. https://doi.org/10.3389/fpsyg.2017.02137

Jacobucci, R., Grimm, K. J., & McArdle, J. J. (2016). Regularized structural equation modeling. *Structural Equation Modeling: a Multidisciplinary Journal*, *23*, 555–566. https://doi.org/10.1080/10705511.2016.1154793

Jennrich, R. I. (2002). A simple general method for oblique rotation. *Psychometrika*, *67*, 7–19. https://doi.org/10.1007/BF02294706

Jennrich, R. I. (2004). Derivative free gradient projection algorithms for rotation. *Psychometrika*, *69*, 475–480. https://doi.org/10.1007/BF02295647

Kline, R. B. (2023). *Principles and practice of structural equation modeling*. Guilford publications.

Mulaik, S. A. (1986). Factor analysis and Psychometrika: Major developments. *Psychometrika*, *51*, 23–33. https://doi.org/10.1007/BF02293996

McArdle, J. J. (1980). Causal modeling applied to psychonomic systems simulation. *Behavior Research Methods & Instrumentation*, *12*, 193–209. https://doi.org/10.3758/BF03201598

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, *51*, 201–226. https://doi.org/10.1146/annurev.psych.51.1.201

Steiger, J. H. (1979). Factor indeterminacy in the 1930's and the 1970's some interesting parallels. *Psychometrika*, *44*, 157–167. https://doi.org/10.1007/BF02293967

Steiger, J. H. (1996). Coming full circle in the history of factor indeterminancy. *Multivariate Behavioral Research*, *31*, 617–630. https://doi.org/10.1207/s15327906mbr3104_14

Tenenhaus, M., Vinzi, V. E., Chatelin, Y. M., & Lauro, C. (2005). PLS path modeling. *Computational Statistics & Data Analysis*, *48*, 159–205. https://doi.org/10.1016/j.csda.2004.03.005

Terada, Y. (2025). *Statistical properties of matrix decomposition factor analysis*. arXiv Preprint, arXiv:2403.06968.

Yamashita, N. (2024). Exploratory Procedure for Component-Based Structural Equation Modeling for Simple Structure by Simultaneous Rotation. *Psychometrika*, *89*, 411–438.