

分类号 O212/18
UDC _____

密级 _____
编号 10741

兰州财经大学

LANZHOU UNIVERSITY OF FINANCE AND ECONOMICS

硕士学位论文

论文题目 基于贝叶斯协方差弹网的验证性因子分析研究

研究生姓名: 叶贇鑫

指导教师姓名、职称: 傅德印 教授

学科、专业名称: 统计学 数理统计学

研究方向: 复杂数据分析

提交日期: 2022年5月30日

独创性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名： 叶斌鑫

签字日期： 2022年5月30日

导师签名： 傅永印

签字日期： 2022.5.30

关于论文使用授权的说明

本人完全了解学校关于保留、使用学位论文的各项规定，同意（选择“同意”/“不同意”）以下事项：

- 1.学校有权保留本论文的复印件和磁盘，允许论文被查阅和借阅，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文；
- 2.学校有权将本人的学位论文提交至清华大学“中国学术期刊（光盘版）电子杂志社”用于出版和编入CNKI《中国知识资源总库》或其他同类数据库，传播本学位论文的全部或部分內容。

学位论文作者签名： 叶斌鑫

签字日期： 2022年5月30日

导师签名： 傅永印

签字日期： 2022.5.30

Confirmatory Factor Analysis Based on Bayesian Covariance Elastic Net

Candidate : Ye Yunxin

Supervisor: Fu Deyin

摘 要

验证性因子分析作为一种常用的测量模型，通常用于量表的开发。研究者根据已有的先验信息构建验证性因子分析的模型结构，并通过一些严格的假设对模型进行限制。而这些假设的存在，常常导致验证性因子分析不能对数据进行有效的拟合或拟合得到的模型结构不符合理论假设。目前，贝叶斯方法的兴起使得研究人员在使用验证性因子分析时更加灵活。通过对先验的选择，将正则化的方法引入。在放松了原有假设的同时增强了模型在分析实际数据时的实用性与稳健性。

根据前人的研究，在贝叶斯框架下对验证性因子分析模型结构的研究可以被视为一个贝叶斯变量选择问题。因此，为了进一步提升贝叶斯框架下，验证性因子分析的模型选择和参数估计能力，考虑通过引入贝叶斯弹网惩罚实现验证性因子分析模型中的残差协方差矩阵的模型选择，以此达到整体的模型选择的目的。本文主要研究为：首先，在贝叶斯弹网模型的基础上，通过加入分块 Gibbs 采样，实现对协方差矩阵和精度矩阵的识别，提出贝叶斯协方差弹网模型，并进一步拓展至贝叶斯自适应协方差弹网模型；其次，将其与验证性因子分析结合，通过贝叶斯协方差弹网模型实现残差协方差矩阵的模型选择与参数估计，达到修改验证性因子分析模型结构的目的；最后，通过数值模拟和实例分析评估该方法的有效性和实用性。

通过数值模拟与实例分析，将本文所提出的验证性因子分析方法与已有方法进行对比表明，所提出的方法对于识别验证性因子分析的模型结构具有良好的性能。

关键词：验证性因子分析 贝叶斯弹网 协方差矩阵估计 分块 Gibbs 采样 变量选择

Abstract

As a commonly used measurement model, confirmatory factor analysis is usually used in the development of scales. Researchers construct a confirmatory factor analysis model structure based on the existing prior information, and restrict the model through some strict assumptions. The existence of these assumptions often results in confirmatory factor analysis unable to effectively fit the data or the model structure obtained by the fit does not conform to the theoretical assumptions. Currently, the rise of Bayesian methods makes researchers more flexible when using confirmatory factor analysis. Through the choice of a prior, the regularization method is introduced. While relaxing the original assumptions, it enhances the practicality and robustness of the model when analyzing actual data.

According to previous studies, the study of confirmatory factor analysis model structure under the Bayesian framework can be regarded as a Bayesian variable selection problem. Therefore, in order to further improve the model selection and parameter estimation capabilities of confirmatory factor analysis under the Bayesian framework, considered to implement the model selection of the residual covariance matrix in the confirmatory factor analysis model by introducing the Bayesian elastic net penalty. The main research of this thesis is: First, on the basis of the Bayesian elastic net model, adding block Gibbs sampling to realize the identification of the covariance matrix and the precision matrix, proposes the Bayesian covariance elastic net model, and further expands it to Bayesian adaptive Covariance elastic net model; Secondly, combine it with confirmatory factor analysis, and realize the model selection and parameter estimation of the residual covariance matrix through the

Bayesian covariance elastic net model to achieve the purpose of modifying the structure of the confirmatory factor analysis model; Finally, the effectiveness and practicability of the method are evaluated through numerical simulation and empirical analysis.

Through numerical simulation and case analysis, the comparison between the confirmatory factor analysis method proposed in this thesis and the existing methods shows that the proposed method has good performance in identifying the model structure of confirmatory factor analysis.

Keywords : Confirmatory Factor Analysis ; Bayesian Elastic Net ; Covariance Matrix Estimation ; Block Gibbs Sampling ; Variable Selection

目 录

1 引言	1
1.1 研究背景.....	1
1.2 研究意义.....	2
1.3 国内外研究综述.....	3
1.3.1 验证性因子分析研究现状.....	3
1.3.2 正则化方法研究现状.....	4
1.3.3 文献述评.....	5
1.4 研究内容与结构.....	6
2 理论基础	7
2.1 贝叶斯图 Lasso 模型.....	7
2.2 验证性因子分析模型.....	10
2.2.1 贝叶斯框架下的验证性因子分析模型.....	10
2.2.2 贝叶斯协方差 Lasso 验证性因子分析模型.....	11
2.3 贝叶斯弹网模型.....	12
3 贝叶斯协方差弹网	14
3.1 协方差弹网先验.....	14
3.2 分块 Gibbs 采样.....	15
3.3 惩罚参数的选择.....	17
3.4 模型检验.....	18
3.5 模型拓展.....	18
4 贝叶斯协方差弹网验证性因子分析模型	20
4.1 贝叶斯协方差弹网验证性因子分析模型的构建.....	20
4.2 变量选择.....	22
4.3 模型检验.....	22
4.3.1 Gibbs 采样的收敛性检验.....	22
4.3.2 模型的拟合优度检验.....	23

4.4 模型拓展	23
5 数值模拟	25
5.1 模型设定	25
5.1.1 模拟 1	25
5.1.2 模拟 2	26
5.1.3 模拟 3	27
5.2 评估指标	27
5.3 结果分析	28
5.3.1 模拟 1 结果分析	28
5.3.2 模拟 2 结果分析	31
5.3.3 模拟 3 结果分析	32
6 实例分析	34
6.1 数据介绍	34
6.1.1 幽默风格问卷	34
6.1.2 肯塔基正念技能清单	35
6.2 结果分析	35
6.2.1 HSQ 结果分析	35
6.2.2 KIMS 结果分析	38
6.3 讨论	42
7 结论与展望	43
参考文献	44
附录	47
后记	49

1 引言

1.1 研究背景

验证性因子分析(Confirmatory Factor Analysis, CFA)作为结构方程模型(Structural Equation Model, SEM)中的测量模型,其目标是测量观察到的样本量或指标与潜在变量或因子之间的关系。作为用于操作测量模型的常用工具,CFA在心理学、教育、行为学和其他社会科学研究领域具有广泛的应用,关于CFA方法的研究也是国内外经久不衰的课题。不同于探索性因子分析(Exploratory Factor Analysis, EFA),CFA的一个基本特征是它的假设驱动性质。CFA要求研究人员对模型结构要有一个预先的指定。因此,研究人员必须充分了解所研究的问题的先验信息,即已有证据和理论,以了解数据中存在的潜在因子数量,以及其中的指标应与哪些因子有关等等。除了更强调理论和假设检验之外,CFA方法的研究中还提供了许多在EFA中不可用的其他分析可能性¹。对于CFA模型中载荷矩阵 Λ 根据已有的理论知识进行设定,当变量与潜在因子无关时其对应数值为0。残差方差协方差矩阵 ϵ 为对角矩阵。

然而,在使用CFA进行分析时由于其严格的假设限制,经常会导致理论模型不能有效的对实际数据进行拟合。对于拟合效果不佳的CFA模型,为了提高其拟合效果,传统方法通常使用post-hoc modification model(PMM)方法,通过修正指数不断地识别显著相关的残差项并令其相关来提高CFA的拟合效果。但是对于PMM,Pan等^[32](2017)指出对于CFA来说是一种“三心二意”的方法,其目的只是单纯的想要提高拟合优度的指标,同时在分析数据维度较高存在较多显著项的残差方差协方差矩阵项时,放松了参数固定为零这一假设可能会导致残差方差协方差矩阵不是一个稀疏矩阵。PMM方法的效率将会非常的低,且向正定的协方差矩阵中不断添加非对角元素可能会使其成为非正定。

近年来,得益于贝叶斯统计以及对应统计软件的快速发展,贝叶斯统计的复杂性和其对于统计软件的不可用性得到了解决,贝叶斯方法也逐渐开始在因子分析的研究中得到了应用。与传统的方法相比,贝叶斯分析有几个潜在的优

¹方法效应的评估、因子模型的稳定性或不不变性的检验。

势。这些优势包括在小样本中有更好的性能；缺少对于渐进理论的依赖¹；自动处理缺失数据；估计极大似然估计中不可行的复杂模型等等。虽然 CFA 可以针对潜在因子数量和潜在因子与指标的关系适当地指定测量模型，但由于存在交叉载荷固定为零的限制，极大似然分析通常会导致拟合效果不佳。同样，残差方差协方差矩阵非对角元素固定为零的限制也可能导致最后的拟合不佳。这些拟合效果不佳的情况在大量样本中以及需要大量指标的多维测量模型中更加严重。因此，贝叶斯分析通过允许通常固定为零的参数近似为零，避免了这些过度严格的假设。另外，在处理维度较高的数据时，将正则化方法加入到 CFA 中是使得协方差矩阵稀疏的一个好的也是可行的方案。Pan 等^[32](2017)使用贝叶斯协方差 Lasso 模型(Bayesian Covariance Lasso model)对 CFA 的残差协方差矩阵进行稀疏化处理，并提出了贝叶斯协方差 Lasso 验证性因子分析模型(Bayesian Covariance Lasso Confirmatory Factor Analysis model, BCLCFA)，通过贝叶斯正则化方法在保证协方差矩阵的正定性的情况下，一次性对所有显著的残差方差协方差矩阵项进行模型选择与参数估计。这一方法解决了放松 CFA 模型假设后矩阵不够稀疏所导致的计算量过大的问题，同时使得 CFA 模型在进行分析时能够有效的找出符合数据的模型结构，提高了 CFA 模型性能。

1.2 研究意义

随着心理学、教育、行为学和其他社会科学研究领域的发展与科学技术的进步，在使用 CFA 进行研究时经常需要对一些高维且复杂的数据集进行分析。而对于这些数据，BCLCFA 方法虽然可以在不严重改变因子结构的情况下，通过释放一些残差方差协方差矩阵非对角项进行模型的修改从而提高 CFA 的拟合优度减少结构模型中的偏差。但是，对于 Lasso 方法本身存在着一些缺陷如：(1)当现在变量的个数大于样本量时，Lasso 方法受限于样本的容量，所选择显著变量要少于样本量；(2)当一组变量之间相关性较高时，Lasso 经常只能对其中一个变量进行选择而忽略改组变量中的其余变量；(3)当变量个数大于样本容量且变量之间存在高相关时，基于 Lasso 方法的模型选择的结果并不理想。

基于上述问题，CFA 方法依旧存在着进一步提升的空间。理论上，对于

¹与极大似然方法不同，贝叶斯分析并不假设参数估计的分布在大样本理论上是正态的。

Lasso 方法存在的缺陷, Zou 和 Hastie^[42] (2005)提出的弹网模型改进了 Lasso 的性能。随后为了解决弹网模型中的“双缩问题”, Li 和 Lin^[24] (2010)提出了贝叶斯弹网模型(The Bayesian Elastic Net model)。因此, 为了能够更有效的对协方差矩阵进行模型选择与参数估计, 本文在 Wang^[39](2012)所提出的贝叶斯图 Lasso 模型(Bayesian Graphical Lasso model)的基础上, 考虑对协方差矩阵的非对角元素给定一个弹网先验分布, 将贝叶斯弹网扩展至贝叶斯协方差弹网模型(Bayesian Covariance Elastic net model, BCEN)。同时将 BCEN 模型与 CFA 进行结合, 对 CFA 模型的残差方差协方差矩阵进行模型选择, 补充现有的贝叶斯框架下的 CFA 方法的不足, 进一步提高 CFA 在应用研究时的可靠性。实际中, 通过 CFA 对所获得的数据进行分析时, 由于变量和潜在因子的增加, 都可能会导致模型存在着非常复杂模型结构。研究者在对数据进行分析时, 复杂的模型结构将会对数据的分析工作造成极大的困难。因此, 在因子结构不严重改变的情况下, 让模型简单化是一项十分重要的工作, 具有非常重要的现实意义。

1.3 国内外研究综述

1.3.1 验证性因子分析研究现状

因子分析作为一种降维技术, 其目的是考察一组变量之间的协方差或相关系数结构, 并用以解释这些变量与为数较少的因子(潜在因子)之间的关联。对于因子分析, 目前有两类方法: 探索性因子分析(EFA)和验证性因子分析(CFA)^[15,16]。得益于模型的灵活性, CFA 是研究者最常用的测量模型之一。

对于验证性因子分析, 当研究者使用设计的模型不能拟合数据时, 比较传统的做法是基于修正指数的 post-hoc modification model(PMM)来放松验证性因子分析的约束^[18,37]。Sörbom^[37](1989)与 MacCallum^[27](1995)指出如果使用得当, PMM 方法能够精确的发现模型当中的错误, 从而修改得到更加稳健的模型。同时, Schmitt^[36] (2011)指出研究人员使用修正指数来修改验证性因子分析其实是一种“探索性”的方法, 在 CFA 结果不理想时, 通过 EFA 进一步探索 CFA 模型也是合理的。然而, Pan 等^[32] (2017)认为 PMM 的理论在上其实是次优的, 当需要修改的次数过多或数据维度较高时, PMM 方法很难找出正确的模型结构。

随着贝叶斯统计的发展,贝叶斯框架下的 CFA 模型拥有了更大的灵活性。贝叶斯方法通过使用接近零的先验释放了模型的交叉载荷(cross-loading),提供了关于交叉载荷的信息,增强了模型解决实质性问题的能力。Lee^[22](1981)采用了分层贝叶斯方法进行验证性因子分析,并假设因子载荷矩阵中的自由参数是可以交换且服从正态分布的。Giudici 和 Stanghellini^[11](2001)通过允许残差的方差-协方差矩阵具有非零的对角元素来对因子分析模型进行推广,并将其命名为图因子分析模型。Jorgensen^[17](2017)发现对于贝叶斯 CFA,后验预测 p(PPP)值有时也不那么有效(除非将非目标参数固定为零)。同时,为了方便计算放松后的载荷矩阵或残差协方差矩阵。研究人员将贝叶斯正则化巧妙的运用在了因子分析当中,Muthén 和 Asparouhov^[30](2012)将结构方程模型推广至贝叶斯框架下,并给定岭回归先验实现了对因子载荷的正则化。Lu^[26]等(2016)认为贝叶斯框架下 SEM 中的因子载荷模式可以表述为一个贝叶斯变量选择问题,并使用了 spike-and-slab 先验进行建模。Pan 等^[32](2017)认为 CFA 模型中的残差协方差中包含着重要信息,而过往的研究通常会将其忽略,其将 Wang^[39](2012)提出的 Bayesian Graphical Lasso Model 应用在了对验证性因子分析的残差方差-协方差矩阵中非对角元素的正则化,并证明了贝叶斯方法在识别残差协方差矩阵中显著项时,相对于 PMM 具有明显的优越性。Chen 等^[4](2020)认为 EFA 和 CFA 之间可以形成互补,在此基础上结合 Pan 等^[32](2017)的 BCLCFA 模型提出了部分验证性因子分析。在 CFA 的模型评估方面,Hoofs 等^[13](2018)提出了贝叶斯近似均方根误差(Bayesian root mean square error of approximation)代替了 PPP 值,用于评估在大样本情形下贝叶斯 CFA 模型的可信度。

1.3.2 正则化方法研究现状

正则化方法是一种通过限制模型参数,达到简化模型、防止模型过拟合,使模型能够拥有更好的解释性的方法。随着高维复杂数据越来越常见,正则化的方法得到了更多的应用。而相比于传统的正则化方法,贝叶斯正则化方法的提出进一步提升了模型的稀疏能力与适应性。Hsiang^[14](1975)在贝叶斯的框架下提出了贝叶斯岭回归。Park 和 Casella^[33](2008)在 Tibshirani^[38](1996)提出的通过 Lasso 对线性回归模型进行收缩和变量选择的基础上研究了贝叶斯 Lasso 在线性回归模型中的应用。Hans^[12](2009)则是进一步发展了贝叶斯 Lasso 回归。O'Hara

和 Sillanpää^[13](2009)在做变量选择的时候加入了 spike-and-slab 先验。Li 和 Lin^[24](2010)提出了贝叶斯弹网解决了弹网的“双缩问题”。

然而，上述贝叶斯正则化方法在应用时并未考虑到对协方差矩阵的模型选择。对于协方差矩阵的研究，比较常见的是在图模型中的应用。Lauritzen^[21](1996)和 Edwards^[6](2012)介绍了高斯图模型的统计性质以及常用的模型选择和参数估计方法，其模型中的参数估计和模型选择相当于估计参数和识别精度矩阵中的零元。在对精度矩阵的估计和模型恢复的研究中，通过在图模型中施加 lasso 惩罚来达到目的(Yuan 和 Lin^[40], 2007 和 Friedman 等^[8], 2008)。在图 Lasso 的基础上，Wang^[39](2012)提出了贝叶斯图 lasso，并将其推广到贝叶斯自适应图 lasso。与图 lasso(Yuan 和 Lin^[40], 2007; Friedman 等^[8], 2008)、自适应图 lasso 以及 SCAD(Fan^[7]等, 2009)等相比，整体上表现出明显的优势。然而 lasso 方法存在着一定的缺陷，Zou 和 Hastie^[42](2005)在 lasso 的基础上增加了 L_2 范数提出了弹网(Elastic net)模型以解决 lasso 的不足¹。其中， L_1 范数提供了稀疏可解释性，而 L_2 范数提高了方法的灵敏度(Ryali 等^[34], 2012)。因此，在高斯图模型中加入弹网惩罚也是一个可行的方案。在研究图模型时，为了保持类内关系并减少类间关系，Liu 等^[25](2016)提出了约束弹网正则化模型。Bernardini 等^[31](2021)、Kovács 和 Ruckstuhl^[19](2021)在高斯图模型的背景下将图 lasso 模型推广至图弹网模型。

1.3.3 文献述评

通过对文献的梳理，现有文献对于 CFA 的研究主要有：1.通过贝叶斯方法放松传统 CFA 模型的假设，以应对传统 CFA 方法对实际数据的拟合效果不佳的问题，但该方法往往导致模型结构过于复杂；2.通过贝叶斯正则化的方法，在对贝叶斯框架下的 CFA 的因子载荷和残差协方差矩阵的估计中施加惩罚，在防止模型过拟合的同时对 CFA 模型结构进行修改，但对于变量较多的情况，依旧不能显著的降低模型的复杂性；3.将探索性的方法与验证性的方法结合，对 CFA 模型的结构进一步的探索。

¹ 如：高维情形下，样本量不足或者低维情形下，变量之间的高相关性导致预测性能较差。

1.4 研究内容与结构

本文的研究内容可分为以下几个部分：第一部分主要介绍目前 CFA 方法和正则化方法的研究现状以及本文主要的创新点，并介绍了研究内容相关的研究背景和文献综述及文献述评，并给出了主要的文章结构和内容跟框架；第 2 部分梳理了本文的理论基础和本文方法研究所涉及的理论基础知识；第 3 部分主要是在已有的贝叶斯图 Lasso 和贝叶斯自适应图 Lasso 模型的基础上，提出 BCEN 和贝叶斯自适应协方差弹网(Bayesian Adaptive Covariance Elastic Net, BACEN)模型；第 4 部分将提出的 BCEN 模型和 BACEN 模型与 CFA 方法结合，同时该部分也是本研究的主要创新；第 5 部分和第 6 部分分别是数值模拟和实证分析，主要通过模拟数据和实例数据检验所提出的 2 种方法的稳健性、有效性和实用性；第 7 部分是对于研究的总结和展望。本研究结构的流程框架图如下

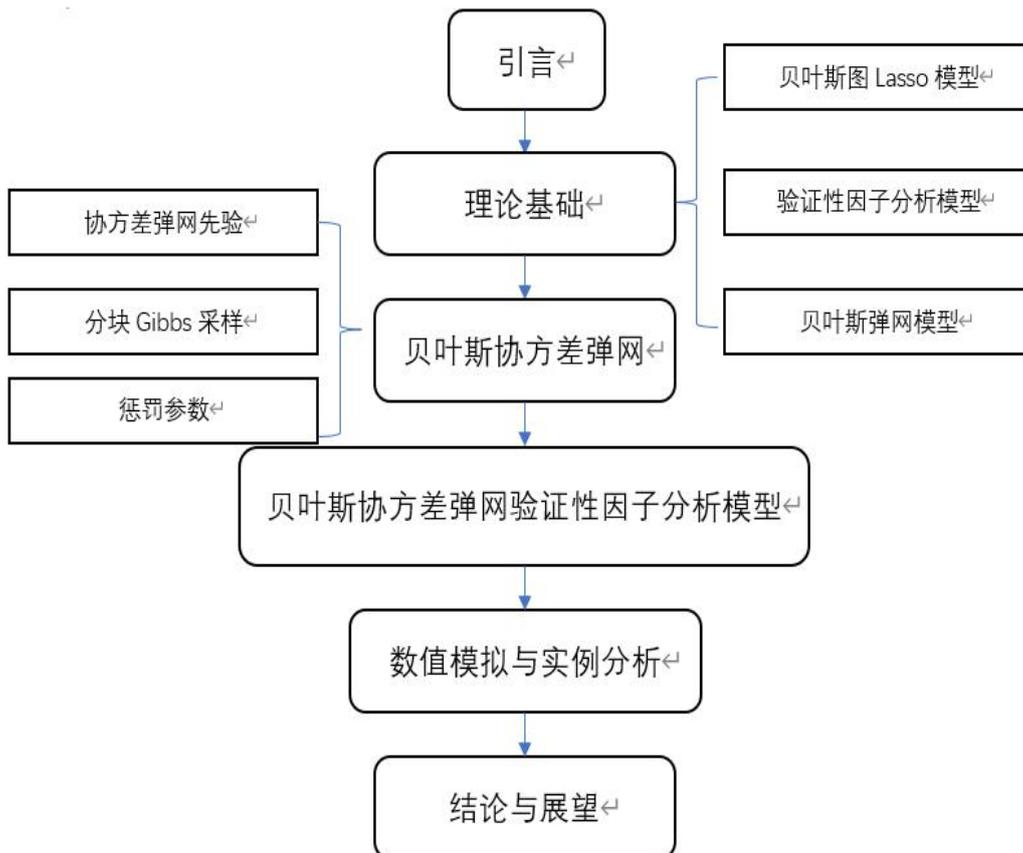


图 1.1 结构流程图

2. 理论基础

本章对贝叶斯图 Lasso 模型、贝叶斯协方差 Lasso 验证性因子分析以及贝叶斯弹网的相关理论知识的进行简要介绍。本章的主要目的是讨论贝叶斯图 Lasso 模型以及贝叶斯协方差弹网如何将其与验证性因子分析结合；对贝叶斯弹网进行简要介绍，为本文提出的贝叶斯协方差弹网以及贝叶斯协方差弹网验证性因子分析模型做准备。

2.1 贝叶斯图 Lasso 模型

在高斯图模型的研究中，研究者通过 Lasso 惩罚提升了其模型选择与参数估计的能力(Yuan 和 Lin^[40], 2007)。这种方法也被称为图 lasso，通过最大化惩罚对数似然实现参数估计

$$\log(\det \boldsymbol{\Theta}) - \text{tr}(\mathbf{S}\boldsymbol{\Theta}) - \rho \|\boldsymbol{\Theta}\|_1, \quad (2.1)$$

其中， $\boldsymbol{\Theta} = \boldsymbol{\Sigma}^{-1}$ 为精度矩阵， \mathbf{S} 为经验协方差矩阵， $\det \boldsymbol{\Theta}$ 为矩阵 $\boldsymbol{\Theta}$ 的行列式， $\text{tr}(\cdot)$ 为对应矩阵的迹， $\|\boldsymbol{\Theta}\|_1$ 为矩阵 $\boldsymbol{\Theta}$ 的 L1 范数， ρ 为惩罚参数。随后，Wang^[39](2012)在贝叶斯框架下进行了推广，通过其提出的分块 Gibbs 方法实现了图 Lasso 的完全贝叶斯处理，提出了贝叶斯图 Lasso 模型和贝叶斯自适应图 Lasso 模型。相较于频率派和其他贝叶斯方法，贝叶斯图 Lasso 有更好的矩阵估计和图形结构学习的性能。Wang^[39](2012)对精度矩阵 $\boldsymbol{\Theta}$ 的对角元素和非对角元素分别赋予了独立的指数先验和双指数先验，因此贝叶斯图 lasso 模型相当于以下模型的最大后验估计

$$\begin{aligned} p(\mathbf{y}_i | \boldsymbol{\Theta}) &= N(\mathbf{y}_i | 0, \boldsymbol{\Theta}^{-1}), \quad i = 1, 2, \dots, n \\ p(\boldsymbol{\Theta} | \lambda) &= C^{-1} \prod_{i < j} \{DE(\theta_{ij} | \lambda)\} \prod_{i=1}^p \left\{ \text{Exp}\left(\theta_{ii} | \frac{\lambda}{2}\right) \right\} I_{\boldsymbol{\Theta} \in M^+}, \end{aligned} \quad (2.2)$$

其中， $DE(x | \lambda)$ 是形式为 $p(x) = \lambda / 2 \exp(-\lambda |x|)$ 的双指数密度函数、 $\text{Exp}(x | \lambda)$ 是形式为 $p(x) = \lambda \exp(-\lambda x) I_{x > 0}$ 的指数密度函数、 $N(\mathbf{y} | \boldsymbol{\Sigma})$ 为多元正态分布、 C 是与惩罚参数 λ 无关的归一化常数。同时，为了使式(2.2)能够通过 Gibbs 方法进行采样，需将其转换成正态分布的尺度混合分布(Park 和 Casella^[33], 2008)。因

此，贝叶斯图 Lasso 模型的分层模型为

$$p(\boldsymbol{\theta} | \boldsymbol{\tau}, \lambda) = C_{\boldsymbol{\tau}}^{-1} \prod_{i < j} \left\{ \frac{1}{\sqrt{2\pi\tau_{ij}}} \exp\left(-\frac{\theta_{ij}^2}{2\tau_{ij}}\right) \right\} \prod_{i=1}^p \left\{ \frac{\lambda}{2} \exp\left(-\frac{\lambda}{2}\theta_{ii}\right) \right\} I_{\boldsymbol{\theta} \in M^+}, \quad (2.3)$$

$$p(\boldsymbol{\tau} | \lambda) \propto C_{\boldsymbol{\tau}} \prod_{i < j} \frac{\lambda^2}{2} \exp\left(-\frac{\lambda^2}{2}\tau_{ij}\right),$$

其中， $\boldsymbol{\theta} = \{\theta_{ij}\}_{i < j}$ 为精度矩阵 $\boldsymbol{\Theta}$ 的上三角元素、 $\boldsymbol{\tau} = \{\tau_{ij}\}_{i < j}$ 为潜在尺度参数、 $C_{\boldsymbol{\tau}}$ 为依赖于 $\boldsymbol{\tau}$ 的规范化常数。

由于贝叶斯框架下传统的 MCMC 方法很难对精度矩阵 $\boldsymbol{\Theta}$ 进行有效的估计。而将精度矩阵 $\boldsymbol{\Theta}$ 的条件分布适当的参数化之后，分块 Gibbs 采样将是一个十分有效的采样器(Wang^[39], 2012)。根据式(2.3)，贝叶斯图 Lasso 在使用分块 Gibbs 采样时，目标分布如下

$$p(\boldsymbol{\Theta}, \boldsymbol{\tau} | \mathbf{Y}, \lambda) \propto |\boldsymbol{\Theta}|^{\frac{n}{2}} \exp\left\{-tr\left(\frac{1}{2}\mathbf{S}\boldsymbol{\Theta}\right)\right\}$$

$$\times \prod_{i < j} \left\{ \tau_{ij}^{-\frac{1}{2}} \exp\left[-\frac{\theta_{ij}^2}{2\tau_{ij}}\right] \exp\left(-\frac{\lambda^2}{2}\tau_{ij}\right) \right\} \quad (2.4)$$

$$\times \prod_{i=1}^p \left\{ \exp\left(-\frac{\lambda}{2}\theta_{ii}\right) \right\} I_{\boldsymbol{\Theta} \in M^+},$$

其中，精度均值 $\boldsymbol{\Theta}$ 被约束为正定矩阵。同时，使用分块 Gibbs 采样时，需要对精度矩阵 $\boldsymbol{\Theta}$ 进行分块处理。令 $\mathbf{T} = (\tau_{ij})$ 为 $p \times p$ 的对称矩阵，对角线元素为 0。

得到分块矩阵 $\boldsymbol{\Theta}$ 、 \mathbf{S} 和 \mathbf{T}

$$\boldsymbol{\Theta} = \begin{pmatrix} \boldsymbol{\Theta}_{ii} & \boldsymbol{\theta}_i \\ \boldsymbol{\theta}_i' & \theta_{ii} \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} \mathbf{S}_{ii} & \mathbf{s}_i \\ \mathbf{s}_i' & s_{ii} \end{pmatrix}, \quad \mathbf{T} = \begin{pmatrix} \mathbf{T}_{ii} & \boldsymbol{\tau}_i \\ \boldsymbol{\tau}_i' & 0 \end{pmatrix}, \quad i=1, 2, \dots, p \quad (2.5)$$

其中， $\boldsymbol{\theta}_i = (\theta_{i,1}, \theta_{i,2}, \dots, \theta_{i,i-1}, \theta_{i,i+1}, \dots, \theta_{i,p})$ 为精度矩阵 $\boldsymbol{\Theta}$ 第 i 列的所有非对角元素所组成的向量， θ_{ii} 为精度矩阵 $\boldsymbol{\Theta}$ 对角线上的第 i 个元素， $\boldsymbol{\Theta}_{ii}$ 为除去精度矩阵 $\boldsymbol{\Theta}$ 的第 i 行和第 i 列元素后得到的 $(p-1) \times (p-1)$ 维矩阵。同理可得， s_{ii} 和 0 分别为矩阵 \mathbf{S} 和矩阵 \mathbf{T} 对角线上的第 i 个元素， \mathbf{s}_i 和 $\boldsymbol{\tau}_i$ 分别为矩阵 \mathbf{S} 和矩阵 \mathbf{T} 第 i 列的所有非对角元素所组成的向量， \mathbf{S}_{ii} 和 \mathbf{T}_{ii} 分别为去除矩阵 \mathbf{S} 和矩阵 \mathbf{T} 的第 i 行和第 i 列元素后得到的 $(p-1) \times (p-1)$ 维矩阵。

根据式(2.5)的 3 个分块矩阵， $\boldsymbol{\Theta}$ 第 i 列的条件分布为

$$p(\boldsymbol{\theta}_i, \theta_{ii} | \boldsymbol{\Theta}_{ii}, \mathbf{T}, \mathbf{Y}, \lambda) \propto (\theta_{ii} - \boldsymbol{\theta}'_i \boldsymbol{\Theta}_{ii}^{-1} \boldsymbol{\theta}_i)^{\frac{n}{2}} \times \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\theta}'_i \mathbf{D}^{-1} \boldsymbol{\tau} \boldsymbol{\theta}_i + 2 \mathbf{s}'_i \boldsymbol{\theta}_i + (s_{ii} + \lambda) \theta_{ii} \right] \right\}, \quad (2.6)$$

其中, $\mathbf{D}_\tau = \text{diag}(\tau_i)$ 。在式(2.6)的基础上, Wang^[39](2012)将变量进行转换, 令 $\boldsymbol{\beta} = \boldsymbol{\theta}_i$ 、 $\gamma = \theta_{ii} - \boldsymbol{\theta}'_i \boldsymbol{\Theta}_{ii}^{-1} \boldsymbol{\theta}_i$ 有

$$p(\boldsymbol{\beta}, \gamma | \boldsymbol{\Theta}_{ii}, \mathbf{T}, \mathbf{Y}, \lambda) \propto \gamma^{\frac{n}{2}} \exp \left(-\frac{s_{ii} + \lambda}{2} \gamma \right) \times \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\beta}' (\mathbf{D}_\tau^{-1} + (s_{ii} + \lambda) \boldsymbol{\Theta}_{ii}^{-1}) \boldsymbol{\beta} + 2 \mathbf{s}'_i \boldsymbol{\eta} \right] \right\}. \quad (2.7)$$

根据式(2.7), $(\gamma, \boldsymbol{\beta})$ 的条件分布为

$$\gamma, \boldsymbol{\beta} | \boldsymbol{\Theta}_{ii}, \mathbf{T}, \mathbf{Y}, \lambda \sim \text{Gamma} \left(\frac{n}{2} + 1, \frac{s_{ii} + \lambda}{2} \right) N(-\mathbf{C} \mathbf{s}_i, \mathbf{C}), \quad (2.8)$$

其中, $\mathbf{C} = (\mathbf{D}_\tau^{-1} + (s_{ii} + \lambda) \boldsymbol{\Theta}_{ii}^{-1})^{-1}$ 。此外, 对于潜在尺度参数 $\boldsymbol{\tau}$ 的更新, $1/\boldsymbol{\tau}$ 的条件后验分布为独立的逆高斯分布

$$\frac{1}{\boldsymbol{\tau}} \sim \text{IG} \left(\sqrt{\frac{\lambda^2}{\theta_{ij}^2}}, \lambda^2 \right), \quad (2.9)$$

对于惩罚参数 λ , Wang^[39](2012)给出了一个伽马先验 $\lambda \sim \text{Gamma}(r, s)$, 得到其条件后验分布

$$\lambda \sim \text{Gamma} \left(r + \frac{p(p+1)}{2}, s + \frac{\|\boldsymbol{\Theta}\|_1}{2} \right), \quad (2.10)$$

其中, r 和 s 为给定的常数。最后, Wang^[39](2012)通过对精度矩阵的所有元素赋予不同的惩罚参数, 将贝叶斯图 Lasso 模型拓展至贝叶斯自适应图 Lasso 模型, 进一步提升了模型的稳健性。贝叶斯自适应 Lasso 的后验形式为

$$p(\mathbf{y}_i | \boldsymbol{\Theta}) = N(\mathbf{y}_i | 0, \boldsymbol{\Theta}^{-1}), \quad i = 1, 2, \dots, n$$

$$p(\boldsymbol{\Theta} | \{\lambda_{ij}\}_{i \leq j}) = C_{\{\lambda_{ij}\}_{i \leq j}}^{-1} \prod_{i < j} \{DE(\theta_{ij} | \lambda_{ij})\} \prod_{i=1}^p \left\{ \text{Exp} \left(\theta_{ii} | \frac{\lambda_{ii}}{2} \right) \right\} I_{\boldsymbol{\Theta} \in M^+}, \quad (2.11)$$

$$p(\{\lambda_{ij}\}_{i < j} | \{\lambda_{ii}\}_{i=1}^p) \propto C_{\{\lambda_{ij}\}_{i < j}} \prod_{i < j} \text{Gamma}(r', s'),$$

其中, $C_{\{\lambda_{ij}\}_{i \leq j}}$ 为规范化常数且没有解析解、 $\{\lambda_{ii}\}_{i=1}^p$ 为对角元素的超参数。式(2.11)中的两项 $C_{\{\lambda_{ij}\}_{i \leq j}}$ 会在 λ_{ij} 的条件后验计算中抵消, 从而能够直接对 λ_{ij} 进行 Gibbs 采样。根据式(2.11), λ_{ij} 的条件后验为

$$\lambda_{ij} | \Theta \sim \text{Gamma}(1 + r', |\theta_{ij}| + s'). \quad (2.12)$$

2.2 验证性因子分析模型

2.2.1 贝叶斯框架下的验证性因子分析模型

对于收集到的 n 个样本 $\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_n$ ，且对于任意样本 $\mathbf{y}_k = (y_{k1}, y_{k2}, \dots, y_{kp})^T$ 具有以下模型形式

$$\mathbf{y}_k = \boldsymbol{\mu} + \mathbf{A}\boldsymbol{\omega}_k + \boldsymbol{\varepsilon}_k, \quad k = 1, 2, \dots, n \quad (2.13)$$

其中， $\boldsymbol{\mu}$ 为 $p \times 1$ 维的截距向量， \mathbf{A} 为 $p \times q$ 维的载荷矩阵， $\boldsymbol{\omega}_k$ 为 $q \times 1$ 维的潜在因子集合， $\boldsymbol{\varepsilon}_k$ 为 $p \times 1$ 维的残差向量。CFA 假设 $\boldsymbol{\omega}_k \sim N_q(\boldsymbol{\theta}, \boldsymbol{\Phi})$ ， $\boldsymbol{\varepsilon}_k \sim N_p(\boldsymbol{0}, \boldsymbol{\Psi})$ ，且矩阵 $\boldsymbol{\Phi}$ 为正定矩阵、矩阵 $\boldsymbol{\Psi}$ 为对角矩阵。CFA 中的协方差为

$$\boldsymbol{\Sigma}(\mathbf{y}_k) = \mathbf{A}\boldsymbol{\Phi}\mathbf{A}' + \boldsymbol{\Psi}, \quad (2.14)$$

而贝叶斯方法在估计参数时最重要的是对参数的后验分布的估计。通过给定模型参数的先验，再结合对应的似然函数即可得到所需估计的参数的后验形式。Lee^[22](2007)指出在贝叶斯框架下的结构方程模型中，给定数据 \mathbf{Y} 和模型 F 的参数 $\boldsymbol{\xi}$ 的先验分布 $p(\boldsymbol{\xi})$ ，得参数 $\boldsymbol{\xi}$ 的后验分布 $p(\boldsymbol{\xi} | \mathbf{Y}, F) = p(\mathbf{Y} | \boldsymbol{\xi}, F)p(\boldsymbol{\xi})$ 。同时，其对数后验分布的形式为

$$\log p(\boldsymbol{\xi} | \mathbf{Y}, F) = \log p(\mathbf{Y} | \boldsymbol{\xi}, F) + \log p(\boldsymbol{\xi}). \quad (2.15)$$

由于似然函数 $p(\mathbf{Y} | \boldsymbol{\xi}, F)$ 依赖于样本量，而 $p(\boldsymbol{\xi})$ 则是根据研究人员对于所收集的数据的认识给定。因此当样本量较大时，对数后验 $\log p(\boldsymbol{\xi} | \mathbf{Y}, F)$ 将会接近对数似然 $\log p(\mathbf{Y} | \boldsymbol{\xi}, F)$ ，此时贝叶斯方法与极大似然方法是渐进等价的；而当样本量较少时，先验信息通过先验分布 $p(\boldsymbol{\xi})$ 引入贝叶斯分析中，以此达到更好的估计效果(Lee, 2007)。对于 CFA 模型，参数 $\boldsymbol{\xi}$ 为式(1.1)中 \mathbf{A} 、 $\boldsymbol{\Phi}$ 和 $\boldsymbol{\Psi}$ 的未知元素的结构参数向量 (structural parameter vector)。而潜在因子 $\boldsymbol{\Omega} = (\omega_1, \omega_2, \dots, \omega_n)$ 通常假设为缺失数据，通过 Gibbs 采样对参数进行估计。

2.2.2 贝叶斯协方差 Lasso 验证性因子分析模型

对于贝叶斯框架下的 CFA 模型，Pan 等^[32](2017)认为，当给定合适的先验

分布 $p(\xi)$ ，其对数先验分布 $\log p(\xi)$ 能够充当惩罚函数的角色，实现对验证性因子分析的模型选择^[32]。因此，Pan 等^[32](2017)将贝叶斯图 Lasso(Wang^[39], 2012)模型与 CFA 结合，通过给定残差的精度矩阵 $\tilde{\Theta} = \Psi^{-1}$ 的非对角元素独立的双指数先验 $\lambda/2 \exp(-\lambda|\tilde{\theta}_{ij}|)$ ， $i < j$ ，和给定对角元素分布指数先验 $\lambda/2 \exp(-\frac{\lambda}{2}\tilde{\theta}_{ii})$ 实现了对残差协方差矩阵的模型选择与参数估计。根据式(2.2)和式(2.3)，精度矩阵 Θ 的上三角元素所组成的向量 $\theta = (\theta_{ij})_{i < j}$ 和对应的潜在尺度参数 $\tau = (\tau_{ij})_{i < j}$ 的条件概率密度为

$$p(\theta | \tau, \lambda) = B_{\tau}^{-1} \prod_{i < j} \left\{ \frac{1}{\sqrt{2\pi\tau_{ij}}} \exp \left[-\frac{\theta_{ij}}{2\tau_{ij}} \right] \right\} \\ \times \prod_{i=1}^p \left\{ \frac{\lambda}{2} \exp \left(-\frac{\lambda}{2} \theta_{ii} \right) \right\} I_{\theta \in M^+}, \quad (2.16)$$

$$p(\tau | \lambda) \propto B_{\tau} \prod_{i < j} \frac{\lambda^2}{2} \exp \left(-\frac{\lambda^2}{2} \tau_{ij} \right),$$

同时，得到条件分布 $p(\Theta | Y, \Omega, \xi, \tau, \lambda)$ ，并通过分块 Gibbs 采样。其形式为

$$p(\Theta | Y, \Omega, \xi, \tau, \lambda) \propto p(Y | \Omega, \xi, \Theta) p(\Theta | \tau, \lambda) \\ \propto |\Theta|^{\frac{n}{2}} \exp \left[-\frac{1}{2} \text{tr}(\tilde{S}\Theta) \right] \\ \times \prod_{i < j} \left\{ \exp \left[-\frac{\tilde{\theta}_{ij}^2}{2\tau_{ij}} \right] \right\} \\ \times \prod_{i=1}^p \left\{ \exp \left(-\frac{\lambda}{2} \theta_{ii} \right) \right\} I_{\theta \in M^+}, \quad (2.17)$$

其中， $\tilde{S} = \sum_{i=1}^n (y_i - \mu_i - \Lambda \omega_i)(y_i - \mu_i - \Lambda \omega_i)^T$ 。在通过式(2.17)采样得到残差协方差矩阵 $\Psi = \Theta^{-1}$ 后，同式(2.5)类似，对其进行分块处理

$$\Psi = \begin{pmatrix} \Psi_{ii} & \psi_i \\ \psi_i' & \psi_{ii} \end{pmatrix}, \quad (2.18)$$

因此，对于 BCLCFA 中的其他参数的条件分布为

$$\Lambda_j | Y, \Omega, \mu, \Lambda_{-j}, \Phi, \Psi \sim N \left(\left(\Psi_j^{*-1} \Omega \Omega^T + H_{0j}^{-1} \right)^{-1} \left(\Psi_j^{*-1} \Omega Y_j^{*T} + H_{0j}^{-1} \Lambda_{0j} \right), \left(\Psi_j^{*-1} \Omega \Omega^T + H_{0j}^{-1} \right)^{-1} \right), \quad (2.18)$$

$$\omega_i | Y, \mu, \Lambda, \Phi, \Psi \sim N \left(\left(\Phi^{-1} + \Lambda^T \Psi^{-1} \Lambda \right)^{-1} \Lambda^T \Psi^{-1} (y - \mu), \left(\Phi^{-1} + \Lambda^T \Psi^{-1} \Lambda \right)^{-1} \right), \quad (2.19)$$

$$\Phi | Y, \Omega, \mu, \Lambda, \Psi \sim IW \left(\Omega \Omega^T + R_0^{-1}, n + \rho_0 \right), \quad (2.20)$$

$$\boldsymbol{\mu} | Y, \boldsymbol{\Omega}, \boldsymbol{A}, \boldsymbol{\Phi}, \boldsymbol{\Psi} \sim N\left(\left(\boldsymbol{H}_{\boldsymbol{\mu}_0}^{-1} + n\boldsymbol{\Psi}^{-1}\right)^{-1}\left(\boldsymbol{\Psi}^{-1}\boldsymbol{V} + \boldsymbol{H}_{\boldsymbol{\mu}_0}^{-1}\boldsymbol{\mu}_0\right), \left(\boldsymbol{H}_{\boldsymbol{\mu}_0}^{-1} + n\boldsymbol{\Psi}^{-1}\right)^{-1}\right), \quad (2.21)$$

其中, $\boldsymbol{\Psi}_j^* = \boldsymbol{\Psi}_{jj} - \boldsymbol{\Psi}_j^T \boldsymbol{\Psi}_{jj}^{-1} \boldsymbol{\Psi}_j$, \boldsymbol{y}_{-j} 和 $\boldsymbol{\mu}_{-j}$ 分别为 \boldsymbol{y}_i 和 $\boldsymbol{\mu}$ 去除掉第 j 个元素后的向量, 矩阵 \boldsymbol{Y}_j^* 的元素 $y_{ij}^* = y_{ij} - \mu_j - \boldsymbol{\Psi}_j^T \boldsymbol{\Psi}_{jj}^{-1} (\boldsymbol{y}_{-ij} - \boldsymbol{\mu}_{-j} - \boldsymbol{A}_{-j} \boldsymbol{\omega})$, $\boldsymbol{V} = \sum_{i=1}^n (\boldsymbol{y}_i - \boldsymbol{A} \boldsymbol{\omega}_i)$ 。

2.3 贝叶斯弹网模型

对于一般的线性模型 $E(\boldsymbol{y} | \boldsymbol{X}, \boldsymbol{\beta}) = \boldsymbol{X}\boldsymbol{\beta}$, Li 和 Lin^[24](2010)在给定一个无信息先验 σ 的基础上, 假定其响应变量服从正态分布 $\boldsymbol{y} | \boldsymbol{X}, \boldsymbol{\beta} \sim N(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{I}_n)$ 。同时, 给定参数 $\boldsymbol{\beta}$ 的先验分布得到关于 σ 的条件分布

$$\pi(\boldsymbol{\beta} | \sigma) \propto \exp\left\{-\frac{1}{2\sigma^2}(\lambda_1 \|\boldsymbol{\beta}\|_1 + \lambda_2 \|\boldsymbol{\beta}\|_2^2)\right\}. \quad (2.22)$$

根据式(2.22), 进一步得到对应的分层模型

$$\begin{aligned} \boldsymbol{y} | \boldsymbol{\beta}, \sigma^2 &\sim N(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{I}_n), \\ \boldsymbol{\beta} | \sigma^2 &\sim \exp\left\{-\frac{1}{2\sigma^2}(\lambda_1 \|\boldsymbol{\beta}\|_1 + \lambda_2 \|\boldsymbol{\beta}\|_2^2)\right\}, \\ \sigma^2 &\sim \frac{1}{\sigma^2}, \end{aligned} \quad (2.23)$$

同时, 由于 $\boldsymbol{\beta}$ 的先验中包含了 $\|\boldsymbol{\beta}\|_1$, 导致式(2.23)所给出的分层模型难以通过 Gibbs 方法进行采样。因此, Li 和 Lin^[24](2010)同样采用了 Andrews 和 Mallows^[1](1974)所提出的恒等式

$$\frac{a}{2} \exp(-a|z|) = \int_0^\infty \frac{1}{\sqrt{2\pi s}} \exp\left(-\frac{1}{2} \frac{z^2}{s}\right) \frac{a^2}{2} \exp\left(-\frac{1}{2} a^2 s\right) ds. \quad (2.24)$$

将式(2.23)给出的分层模型进行转换, 得到与式(2.23)等价的分层模型

$$\begin{aligned} \boldsymbol{y} | \boldsymbol{\beta}, \sigma^2 &\sim N(\boldsymbol{X}\boldsymbol{\beta}, \sigma^2 \boldsymbol{I}_n), \\ \boldsymbol{\beta} | \boldsymbol{\tau}, \sigma^2 &\sim \prod_{j=1}^p N\left(0, \left(\frac{\lambda_2}{\sigma^2} \frac{\tau_j}{\tau_j - 1}\right)^{-1}\right), \\ \boldsymbol{\tau} | \sigma^2 &\sim \prod_{j=1}^p TG\left(\frac{1}{2}, \frac{8\lambda_2 \sigma^2}{\lambda_1^2}, (1, \infty)\right), \\ \sigma^2 &\sim \frac{1}{\sigma^2}, \end{aligned} \quad (2.25)$$

其中, $\boldsymbol{\tau} | \sigma^2$ 为截断伽马分布。最后, 由式(2.25)推导得到满条件后验分布

$$\begin{aligned}
\boldsymbol{\beta} | \mathbf{y}, \boldsymbol{\tau}, \sigma^2 &\sim N(\mathbf{A}^{-1} \mathbf{X}^T \mathbf{y}, \sigma^2 \mathbf{A}^{-1}), \\
(\boldsymbol{\tau} - \mathbf{I}_p) | \mathbf{y}, \boldsymbol{\beta}, \sigma^2 &\sim \prod_{j=1}^p GIG\left(\lambda = \frac{1}{2}, \varphi = \frac{\lambda_1^2}{4\lambda_2\sigma^2}, \chi = \frac{\lambda_2\beta_j^2}{\sigma^2}\right), \\
\sigma^2 | \mathbf{y}, \boldsymbol{\beta}, \boldsymbol{\tau} &\sim \left(\frac{1}{\sigma^2}\right)^{\frac{n}{2}+p+1} \left\{ \Gamma_U\left(\frac{1}{2}, \frac{\lambda_1^2}{8\sigma^2\lambda_2}\right) \right\}^{-p} \times \\
&\quad \exp\left[-\frac{1}{2\sigma^2} \left\{ \|\mathbf{y} - \mathbf{X}\boldsymbol{\beta}\|_2^2 + \lambda_2 \sum_{j=1}^p \frac{\tau_j}{\tau_j - 1} \beta_j^2 + \frac{\lambda_1^2}{4\lambda_2} \sum_{j=1}^p \tau_j \right\}\right],
\end{aligned} \tag{2.26}$$

其中， $\mathbf{A} = \mathbf{X}^T \mathbf{X} + \lambda_2 \text{diag}\left(\frac{\tau_1}{\tau_1 - 1}, \dots, \frac{\tau_p}{\tau_p - 1}\right)$ ， $GIG(\lambda, \varphi, \chi)$ 为广义逆高斯分布，

$\Gamma_U(\alpha, x)$ 为上不完全伽马函数。

3. 贝叶斯协方差弹网模型

对于贝叶斯框架下的 CFA，最大的问题是如何实现残差协方差矩阵的模型选择与参数估计。本章考虑在贝叶斯图 Lasso 模型的基础上，引入带 L_2 惩罚范数的贝叶斯弹网模型。通过分块 Gibbs 采样，在完全贝叶斯方法下实现对协方差矩阵的模型选择与参数估计。同时，本章还推广了一个自适应版本。

3.1 协方差弹网先验

对于图弹网模型，本文考虑对精度矩阵 $\boldsymbol{\Theta}$ 的所有非对角元素施加一个形式为 $p(x) = \lambda_1 / 2 \exp(-\lambda_1 |x| - \lambda_2 x^2)$ 的弹网先验(Li 和 Lin^[24], 2010)。因此，本文所考虑的协方差弹网具有以下形式的最大后验估计

$$\begin{aligned} p(\mathbf{y}_i | \boldsymbol{\Theta}) &= N(\mathbf{y}_i | 0, \boldsymbol{\Theta}^{-1}), \quad i = 1, 2, \dots, n \\ p(\boldsymbol{\Theta} | \lambda_1, \lambda_2) &= C^{-1} \prod_{i < j} \{EN(\theta_{ij} | \lambda_1, \lambda_2)\} \prod_{i=1}^p \left\{ \text{Exp}\left(\theta_{ii} | \frac{\lambda_1}{2}\right) \right\} I_{\boldsymbol{\Theta} \in M^+}, \end{aligned} \quad (3.1)$$

其中， $EN(x | \lambda_1, \lambda_2)$ 是形式为 $p(x) = \lambda_1 / 2 \exp(-\lambda_1 |x| - \lambda_2 x^2)$ 的弹网先验。为了将协方差弹网模型推广至贝叶斯框架并使其能够通过 Gibbs 采样方法对模型进行采样，需要将式(3.1)中的弹网先验进行转换。通常需要将 $EN(x | \lambda_1, \lambda_2)$ 替换成正态分布的尺度混合来表示(Andrews 和 Mallows^[1], 1974; Li 和 Lin^[24], 2010)。然而对 θ_{ij} 使用正态分布的尺度混合分布时，由于精度矩阵必须是正定矩阵这一限制， θ_{ij} 将不再独立于给定的尺度参数(Wang^[39], 2012)。对于精度矩阵 $\boldsymbol{\Theta}$ 的上三角元素所组成的向量 $\boldsymbol{\theta} = \{\theta_{ij}\}_{i < j}$ 与潜在尺度参数 $\boldsymbol{\tau} = \{\tau_{ij}\}_{i < j}$ 具有以下形式的条件概率密度函数

$$\begin{aligned} p(\boldsymbol{\theta} | \boldsymbol{\tau}, \lambda_1, \lambda_2) &= C_{\boldsymbol{\tau}}^{-1} \prod_{i < j} \left\{ \sqrt{\frac{2\lambda_2 \tau_{ij}}{2\pi(\tau_{ij} - 1)}} \exp\left[-\frac{\theta_{ij}^2}{2} \left(\frac{2\lambda_2 \tau_{ij}}{\tau_{ij} - 1}\right)\right] \right\} \\ &\quad \times \prod_{i=1}^p \left\{ \frac{\lambda_1}{2} \exp\left(-\frac{\lambda_1}{2} \theta_{ii}\right) \right\} I_{\boldsymbol{\theta} \in M^+}, \end{aligned} \quad (3.2)$$

$$p(\boldsymbol{\tau} | \lambda_1, \lambda_2) \propto C_{\boldsymbol{\tau}} \prod_{i < j} \tau_{ij}^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \frac{\lambda_1^2}{2\lambda_2} \tau_{ij}\right), \quad (3.3)$$

其中，规范化常数 $C_{\boldsymbol{\tau}}$ 依赖于 $\boldsymbol{\tau}$ 且没有解析表达式。同时，选取 $p(\boldsymbol{\Theta} | \boldsymbol{\tau})$ 和

$p(\boldsymbol{\tau} | \lambda_1, \lambda_2)$ 作为先验, 由于

$$\begin{aligned}
C_{\boldsymbol{\tau}} &= \int \prod_{i < j} \left\{ \sqrt{\frac{2\lambda_2 \tau_{ij}}{2\pi(\tau_{ij} - 1)}} \exp \left[-\frac{\theta_{ij}^2}{2} \left(2\lambda_2 \frac{\tau_{ij}}{\tau_{ij} - 1} \right) \right] \right\} \\
&\quad \times \prod_{i=1}^p \left\{ \frac{\lambda_1}{2} \exp \left(-\frac{\lambda_1}{2} \theta_{ii} \right) \right\} I_{\Omega \in M^+} d(\theta_{ij})_{i < j} \\
&< \int \prod_{i < j} \left\{ \sqrt{\frac{2\lambda_2 \tau_{ij}}{2\pi(\tau_{ij} - 1)}} \exp \left[-\frac{\theta_{ij}^2}{2} \left(2\lambda_2 \frac{\tau_{ij}}{\tau_{ij} - 1} \right) \right] \right\} \\
&\quad \times \prod_{i=1}^p \left\{ \frac{\lambda_1}{2} \exp \left(-\frac{\lambda_1}{2} \theta_{ii} \right) \right\} d(\theta_{ij})_{i < j} \\
&= 1,
\end{aligned} \tag{3.4}$$

虽然规范化常数 $C_{\boldsymbol{\tau}}$ 没有解析表达式, 但根据式(3.2)和式(3.3)进行分块 Gibbs 采样时可以将其抵消。因此, θ_{ij} 的边际分布依旧遵循式(3.1)的形式。

3.2 分块 Gibbs 采样

BCEN 在使用分块 Gibbs 进行采样时具有以下目标分布

$$\begin{aligned}
p(\boldsymbol{\Theta}, \boldsymbol{\tau} | \mathbf{Y}, \lambda_1, \lambda_2) &\propto |\boldsymbol{\Theta}|^{\frac{n}{2}} \exp \left\{ -tr \left(\frac{1}{2} \mathbf{S} \boldsymbol{\Theta} \right) \right\} \\
&\quad \times \prod_{i < j} \left\{ \sqrt{\frac{2\lambda_2 \tau_{ij}}{\tau_{ij} - 1}} \exp \left[-\frac{\theta_{ij}^2}{2} \left(\frac{2\lambda_2 \tau_{ij}}{\tau_{ij} - 1} \right) \right] \tau_{ij}^{-\frac{1}{2}} \exp \left(-\frac{1}{2} \frac{\lambda_1^2}{2\lambda_2} \tau_{ij} \right) \right\} \\
&\quad \times \prod_{i=1}^p \left\{ \exp \left(-\frac{\lambda_1}{2} \theta_{ii} \right) \right\} I_{\boldsymbol{\Theta} \in M^+},
\end{aligned} \tag{3.5}$$

其中, 矩阵 $\mathbf{S} = \mathbf{Y}^T \mathbf{Y} / n$ 。而 $\mathbf{Y} = (y_1, y_2, \dots, y_n)$ 是由 p 个变量和 n 个样本量组成的 $n \times p$ 维的数据矩阵。同时, 精度矩阵 $\boldsymbol{\Theta}$ 为正定矩阵、潜在尺度参数 $\boldsymbol{\tau}$ 则没有限制。为了能够使用分块 Gibbs 采样对精度矩阵进行估计, 需要对式(3.5)中的矩阵进行分块处理。令 $\mathbf{T} = (\tau_{ij})$ 为 $p \times p$ 的对称矩阵, 其对角线上的元素全部等于 0。矩阵 $\boldsymbol{\Theta}$ 、 \mathbf{S} 和 \mathbf{T} 的分块处理如下

$$\boldsymbol{\Theta} = \begin{pmatrix} \boldsymbol{\Theta}_{ii} & \boldsymbol{\theta}_i \\ \boldsymbol{\theta}'_i & \theta_{ii} \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} \mathbf{S}_{ii} & \mathbf{s}_i \\ \mathbf{s}'_i & s_{ii} \end{pmatrix}, \quad \mathbf{T} = \begin{pmatrix} \mathbf{T}_{ii} & \boldsymbol{\tau}_i \\ \boldsymbol{\tau}'_i & 0 \end{pmatrix}, \quad i=1, 2, \dots, p \tag{3.6}$$

其中, $\boldsymbol{\theta}_i = (\theta_{i,1}, \theta_{i,2}, \dots, \theta_{i,i-1}, \theta_{i,i+1}, \dots, \theta_{i,p})$ 为精度矩阵 $\boldsymbol{\Theta}$ 第 i 列的所有非对角元素所组成的向量, θ_{ii} 为精度矩阵 $\boldsymbol{\Theta}$ 对角线上的第 i 个元素, $\boldsymbol{\Theta}_{ii}$ 为除去精度矩阵 $\boldsymbol{\Theta}$ 的第 i 行和第 i 列元素后得到的 $(p-1) \times (p-1)$ 维矩阵。同理可得, \mathbf{s}_i 和 0 分别为矩

阵 \mathbf{S} 和矩阵 \mathbf{T} 对角线上的第 i 个元素, s_i 和 τ_i 分别为矩阵 \mathbf{S} 和矩阵 \mathbf{T} 第 i 列的所有非对角元素所组成的向量, \mathbf{S}_{ii} 和 \mathbf{T}_{ii} 分别为去除矩阵 \mathbf{S} 和矩阵 \mathbf{T} 的第 i 行和第 i 列元素后得到的 $(p-1) \times (p-1)$ 维矩阵。

根据式(3.5)的分布形式, $\boldsymbol{\theta}$ 第 i 列的条件分布具有以下形式

$$p(\theta_i, \theta_{ii} | \boldsymbol{\theta}_{ii}, \mathbf{T}, \mathbf{Y}, \lambda_1, \lambda_2) \propto (\theta_i - \boldsymbol{\theta}' \boldsymbol{\theta}_{ii}^{-1} \boldsymbol{\theta})^{\frac{n}{2}} \times \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\theta}' \mathbf{H}_\tau \boldsymbol{\theta} + 2s_i' \boldsymbol{\theta} + (s_{ii} + \lambda_1) \theta_i \right] \right\}, \quad (3.7)$$

其中, $\mathbf{H}_\tau = 2\lambda_2 \text{diag}[\boldsymbol{\tau}_i / (\boldsymbol{\tau}_i - 1)]$ 。为了方便计算, 需要对变量进行转变。令 $\boldsymbol{\eta} = \boldsymbol{\theta}$, $\kappa = \theta_i - \boldsymbol{\theta}' \boldsymbol{\theta}_{ii}^{-1} \boldsymbol{\theta}$ 。

$$p(\boldsymbol{\eta}, \kappa | \boldsymbol{\theta}_{ii}, \mathbf{T}, \mathbf{Y}, \lambda_1, \lambda_2) \propto \kappa^{\frac{n}{2}} \exp \left(-\frac{s_{ii} + \lambda_1}{2} \kappa \right) \times \exp \left\{ -\frac{1}{2} \left[\boldsymbol{\eta}' (\mathbf{H}_\tau + (s_{ii} + \lambda_1) \boldsymbol{\theta}_{ii}^{-1}) \boldsymbol{\eta} + 2s_i' \boldsymbol{\eta} \right] \right\}, \quad (3.8)$$

根据式(3.8)可以得到 $\boldsymbol{\eta}$ 和 κ 的条件分布

$$\boldsymbol{\eta} | \boldsymbol{\theta}_{ii}, \mathbf{T}, \mathbf{Y}, \lambda_1, \lambda_2 \sim N(-\mathbf{Q} \mathbf{s}_i, \mathbf{Q}), \quad (3.9)$$

$$\kappa | \boldsymbol{\theta}_{ii}, \mathbf{T}, \mathbf{Y}, \lambda_1, \lambda_2 \sim \text{Gamma} \left(\frac{n}{2} + 1, \frac{s_{ii} + \lambda_1}{2} \right), \quad (3.10)$$

其中, $\mathbf{Q} = \{(s_{ii} + \lambda_1) \boldsymbol{\theta}_{ii}^{-1} + \mathbf{H}_\tau\}^{-1}$ 。当式(3.9)和式(3.10)每生成一次新的模拟值就能得到 $\boldsymbol{\theta}_i = \boldsymbol{\eta}$ 、 $\boldsymbol{\theta}' = \boldsymbol{\eta}'$ 和 $\theta_i = \kappa + \boldsymbol{\theta}' \boldsymbol{\theta}_{ii}^{-1} \boldsymbol{\theta}$, 这意味着更新了某一次迭代中精度矩阵 $\boldsymbol{\theta}$ 的第 i 列, 重复 p 次即可完成一次对精度矩阵 $\boldsymbol{\theta}$ 的迭代。同时, 迭代完成一次就能得到协方差矩阵 $\boldsymbol{\Sigma} = \boldsymbol{\theta}^{-1}$ 。此外, $\kappa > 0$ 保证了精度矩阵 $\boldsymbol{\theta}$ 的正定性 (Wang^[39], 2012)。

对于潜在尺度参数 $\boldsymbol{\tau}$, 其迭代更新与贝叶斯弹网(Li 和 Lin^[24], 2010)类似。

$1/(\tau_{ij} - 1)$ 的条件后验分布独立且服从逆高斯分布, 即

$$\frac{1}{(\tau_{ij} - 1)} \sim IG \left(\frac{\lambda_1}{2\lambda_2 |\theta_{ij}|}, \frac{\lambda_1^2}{2\lambda_2} \right), \quad (3.11)$$

综上所述, 根据上式(3.5)-式(3.11), 进一步将 BCEN 模型的分块 Gibbs 采样总结为如下算法 1。

算法 1 分块 Gibbs 算法

输入: 对于第 t 次迭代, 给定当前值 $\Theta^{(t-1)} \in M^+$ 、 $\tau^{(t-1)}$;

过程:

1: **for** $i=1,2,\dots,p$, 进行下面的迭代, 直到收敛:

(a) 根据式(3.6), 对矩阵 Θ 、 S 和 T 进行分块处理。

(b) 根据式(3.9)和式(3.10), 采样得到 η 和 κ 。

(c) 更新得到 $\theta_i^{(t)} = \eta$ 、 $\theta_i^{\prime(t)} = \eta'$ 和 $\theta_{ii}^{(t)} = \kappa + \theta_i^{\prime(t)} \Theta_{ii}^{(t-1)} \theta_i^{(t)}$ 。

2: 对于 $i < j$ 时, 从 $a_{ij} \sim IG(\lambda_i / 2\lambda_2 | \theta_{ij}, \lambda_i^2 / 2\lambda_2)$ 采样得到 a_{ij} , 更新 $\tau_{ij}^{(t)} = (1 / a_{ij}) + 1$ 。

输出: $\Theta^{(t)} \in M^+$, $\tau^{(t)}$

3.3 惩罚参数的选择

对于协方差弹网模型中惩罚参数 λ_1 和 λ_2 的选择。常用的选择惩罚参数的方法是通过最大化数据边际似然的经验贝叶斯估计(Li 和 Lin^[24], 2010), 或者使用交叉验证对惩罚参数 λ_1 和 λ_2 进行选择(Zou 和 Hastie^[42], 2005; Bernardini^[3]等, 2021; Kovács 等^[9], 2021)。但是对于贝叶斯框架下的协方差弹网模型来说, 使用以上两种方法并不是完全贝叶斯方法。而在贝叶斯图 lasso 模型中已经实现了通过完全的贝叶斯分析方法对惩罚参数进行估计(Wang^[39], 2012)。根据式(3.1), 规范化常数 C 具有以下积分形式

$$C = \int_{\Theta \in M^+} \prod_{i < j} \{EN(\theta_{ij} | \lambda_1, \lambda_2)\} \prod_{i=1}^p \left\{ \text{Exp}\left(\theta_{ii} | \frac{\lambda_1}{2}\right) \right\} d\Theta, \quad (3.12)$$

可以注意到, 式(3.12)中规范化常数 C 并没有解析表达式且依赖于惩罚参数 λ_1 和 λ_2 。为了能够从满条件后验分布中对 λ_1 和 λ_2 进行采样可以考虑与式(3.2)和式(3.3)类似的方式将规范化常数 C 抵消(Wang^[39], 2012)。因此, 对于惩罚参数 λ_1 和 λ_2 给定以下先验

$$p(\lambda_1, \lambda_2) \propto C \prod_{i \leq j} \text{Gamma}(\alpha_{\lambda_1}, \beta_{\lambda_1}) \prod_{i < j} \text{Gamma}(\alpha_{\lambda_2}, \beta_{\lambda_2}), \quad (3.13)$$

对于式(3.13)中的超参数, 设置 $\alpha_{\lambda_1} = 1$ 、 $\alpha_{\lambda_2} = 0.01 * p$ 和 $\beta_{\lambda_1} = \beta_{\lambda_2} = 0.01$, 其中 p 为精度矩阵的维度。根据式(3.1)及式(3.13)可以得到条件后验

$$\begin{aligned} \lambda_1 | \lambda_2, \Theta &\sim \text{Gamma}\left(\alpha_{\lambda_1} + \frac{p(p+1)}{2}, \beta_{\lambda_1} + \frac{\|\Theta\|_1}{2}\right), \\ \lambda_2 | \lambda_1, \Theta &\sim \text{Gamma}\left(\alpha_{\lambda_2}, \beta_{\lambda_2} + \|\Theta\|_2^2\right), \end{aligned} \quad (3.14)$$

其中, $\theta = \{\theta_{ij}\}_{i < j}$ 为精度矩阵 Θ 上三角部分的全部非对角线元素所组成的向量。

3.4 模型检验

协方差弹网模型主要目标是对数据的协方差矩阵以及精度矩阵进行参数估计和变量选择。而其估计的结果可以考虑通过 Stein 损失函数进行评估。给定协方差矩阵 $\Sigma = \Theta^{-1}$, 根据 Laurent^[20] (2012), Stein 损失函数可以表示为

$$L(\hat{\Sigma}, \Sigma) = \text{tr}(\hat{\Sigma}\Sigma^{-1}) - \log \left\{ \det(\hat{\Sigma}\Sigma^{-1}) \right\} - p, \quad (3.15)$$

其中, $\hat{\Sigma}$ 为协方差矩阵 Σ 的估计。另外, Wang^[39](2012)指出对于图形结构学习性能的评估可以使用马修斯相关系数 (Matthews Correlation Coefficients, MCC)(Fan 等^[7], 2009)。其定义为

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}, \quad (3.16)$$

其中, TP 为真阳性数量, 即预测值与实际值均不为零的数量; TN 为真阴性数量, 即预测值与实际值均为零的数量; FP 为假阳性数量, 即预测值不为零、实际值为零的数量; FN 为假阴性数量, 即预测值为零、实际值不为零的数量。由于同时考虑了 TP 、 TN 、 FP 和 FN , MCC 通常被认为是一种平衡的方法。对于 MCC , 数值越大, 说明分类的效果越好, 即矩阵的还原效果越好。同时根据 Fan 等^[7](2009), 对于需要估计的参数 θ_{ij} , 如果 $\hat{\theta}_{ij} < 10^{-3}$ 则令 $\{\theta_{ij} = 0\}$ 。

3.5 模型拓展

在常见的变量选择方法中, 自适应 lasso 模型具有 SCAD 的 Oracle 性质 (Zou^[41], 2006)但是对于高维数据的处理缺乏稳定性、弹网模型能够很好的处理高维情况下的共线性问题(Zou 和 Hastie^[42], 2005)但不满足 Oracle 性质。而自适应弹网模型结合了上述模型的优点, 在处理共线问题方面优于其他算法, 大大提高了有限样本的性能(Zou 和 Zhang^[43], 2009)。本节考虑给予精度矩阵中的每个非对角元素 θ_{ij} 不同的惩罚参数 λ_1 和 λ_2 。将 BCEN 模型推广至 BACEN 模型。

根据式(3.1)和式(3.13), BCEN 模型具有和式(2.11)类似的处理规范化常数的方式, 同样 BACEN 模型给定精度矩阵 Θ 的非对角元素一个形式为

$p(x) = \lambda_1 / 2 \exp(-\lambda_1 |x| - \lambda_2 x^2)$ 的弹网先验 $EN(x | \lambda_1, \lambda_2)$

$$p(\mathbf{y}_i | \boldsymbol{\Theta}) = N(\mathbf{y}_i | 0, \boldsymbol{\Theta}^{-1}), \quad i = 1, 2, \dots, n$$

$$p(\boldsymbol{\Theta} | \boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \boldsymbol{\lambda}) = C_{\{\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \boldsymbol{\lambda}\}}^{-1} \prod_{i < j} \{EN(\theta_{ij} | \lambda_{1,ij}, \lambda_{2,ij})\} \prod_{i=1}^p \left\{ \text{Exp}\left(\theta_{ii} | \frac{\lambda_{ii}}{2}\right) \right\} I_{\boldsymbol{\Theta} \in M^+}, \quad (3.17)$$

$$p(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 | \boldsymbol{\lambda}) \propto C_{\{\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \boldsymbol{\lambda}\}} \prod_{i < j} \text{Gamma}(\alpha, \beta) \text{Gamma}(\alpha, \beta),$$

其中, $\boldsymbol{\lambda}_1 = \{\lambda_{1,ij}\}_{i < j}$ 、 $\boldsymbol{\lambda}_2 = \{\lambda_{2,ij}\}_{i < j}$ 和 $\boldsymbol{\lambda} = \{\lambda_{ii}\}_{i=1}^p$ 分别为对应精度矩阵的非对角元素和对角元素的惩罚参数与超参数, 而且对于所有对角线元素的超参数 $\{\lambda_{ii}\}_{i=1}^p$ 给定初始值为 1。 $p(\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 | \boldsymbol{\lambda})$ 为惩罚参数 λ_1 和 λ_2 的联合先验分布。根据式(3.17)可以推出惩罚参数的条件后验

$$\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2 | \boldsymbol{\Theta}, \boldsymbol{\lambda} \propto \prod_{i < j} \{EN(\theta_{ij} | \lambda_{1,ij}, \lambda_{2,ij})\} \text{Gamma}(\alpha, \beta) \text{Gamma}(\alpha, \beta) \times \prod_{i=1}^p \left\{ \text{Exp}\left(\theta_{ii} | \frac{\lambda_{ii}}{2}\right) \right\} I_{\boldsymbol{\Theta} \in M^+}, \quad (3.18)$$

其中, 规范化常数 $C_{\{\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \boldsymbol{\lambda}\}}$ 在计算条件后验时被抵消了。同时, 在式(3.18)中精度矩阵 $\boldsymbol{\Theta}$ 的每个非对角元素都被给予了不同的惩罚参数。因此, 对于任意的 $i < j$, 惩罚参数 $\lambda_{1,ij}$ 和 $\lambda_{2,ij}$ 具有不同形式伽马分布

$$\lambda_{1,ij} | \lambda_{2,ij}, \theta_{ij} \sim \text{Gamma}(\alpha + 1, \beta + |\theta_{ij}|),$$

$$\lambda_{2,ij} | \lambda_{1,ij}, \theta_{ij} \sim \text{Gamma}(\alpha, \beta + \theta_{ij}^2), \quad (3.19)$$

根据伽马分布的性质, 当 θ_{ij} 较小时将会得到一个较大的惩罚参数 λ_1 和 λ_2 , 反之亦然。但是, 为了降低先验分布的超参数对 θ_{ij} 的影响, 提升惩罚参数 λ_1 和 λ_2 对 θ_{ij} 的适应性, 式(3.19)中超参数必需足够的小(Wang^[39], 2012)。因此对于 BACEN 的超参数设定为 $\alpha_1 = 0.01$ 、 $\alpha_2 = 0.005 * p$ 和 $\beta_1 = \beta_2 = 10^{-4}$, 其中 p 为精度矩阵的维度。

4. 贝叶斯协方差弹网验证性因子分析模型

本章介绍了在贝叶斯框架下如何将所提出的 BCEN 和 BACEN 与 CFA 结合，实现对 CFA 的残差协方差矩阵的模型选择与参数估计。以此，达到修改 CFA 的模型结构、提升 CFA 方法在分析实际数据时的稳健性与实用性的目的。

4.1 贝叶斯协方差弹网验证性因子分析模型的构建

为了进一步提升 CFA 模型在应用时的可靠性，考虑将 BCEN 模型与 CFA 结合。根据式(3.1)，对精度矩阵 $\tilde{\Theta}$ 的非对角元素施加弹网先验 $EN(x|\lambda_1, \lambda_2)$ ，其先验分布为

$$p(\tilde{\Theta} | \lambda_1, \lambda_2) = D^{-1} \prod_{i < j} \left\{ EN(\tilde{\theta}_{ij} | \lambda_1, \lambda_2) \right\} \prod_{i=1}^p \left\{ EXP\left(\tilde{\theta}_{ii} | \frac{\lambda_1}{2}\right) \right\} I_{\Sigma \in M^+}, \quad (4.1)$$

其中， D 为规范化常数。此外，根据式(3.2)和式(3.3)将弹网先验替换为正态分布的尺度混合分布。对于精度矩阵 $\tilde{\Theta}$ 的上三角元素所组成的向量 $\tilde{\theta} = (\tilde{\theta}_{ij})_{i < j}$ 和潜在尺度参数 $\tau = (\tau_{ij})_{i < j}$ 具有以下形式的条件概率密度函数

$$\begin{aligned} p(\tilde{\theta} | \tau, \lambda_1, \lambda_2) &= \tilde{B}_\tau^{-1} \prod_{i < j} \left\{ \sqrt{\frac{2\lambda_2\tau_{ij}}{2\pi(\tau_{ij}-1)}} \exp\left[-\frac{\tilde{\theta}_{ij}^2}{2} \left(\frac{2\lambda_2\tau_{ij}}{\tau_{ij}-1}\right)\right] \right\} \\ &\quad \times \prod_{i=1}^p \left\{ \frac{\lambda_1}{2} \exp\left(-\frac{\lambda_1}{2} \tilde{\theta}_{ii}\right) \right\} I_{\tilde{\Theta} \in M^+}, \quad (4.2) \\ p(\tau | \lambda_1, \lambda_2) &\propto \tilde{B}_\tau \prod_{i < j} \tau_{ij}^{-\frac{1}{2}} \exp\left(-\frac{1}{2} \frac{\lambda_1^2}{2\lambda_2} \tau_{ij}\right). \end{aligned}$$

因此，对 CFA 模型残差的精度矩阵的条件分布 $p(\tilde{\Theta} | Y, \Omega, \mu, \Lambda, \Phi, \tau, \lambda_1, \lambda_2)$ 进行分块 Gibbs 采样时能够进行以下形式的分解(Pan 等^[32], 2017)

$$\begin{aligned} p(\tilde{\Theta} | Y, \Omega, \xi, \tau, \lambda_1, \lambda_2) &\propto p(Y | \Omega, \xi, \tilde{\Theta}) p(\tilde{\Theta} | \tau, \lambda_1, \lambda_2) \\ &\propto |\tilde{\Theta}|^{\frac{n}{2}} \exp\left[-\frac{1}{2} \text{tr}(\tilde{S}\tilde{\Theta})\right] \\ &\quad \times \prod_{i < j} \left\{ \sqrt{\frac{2\lambda_2\tau_{ij}}{2\pi(\tau_{ij}-1)}} \exp\left[-\frac{\tilde{\theta}_{ij}^2}{2} \left(\frac{2\lambda_2\tau_{ij}}{\tau_{ij}-1}\right)\right] \right\} \quad (4.3) \\ &\quad \times \prod_{i=1}^p \left\{ \frac{\lambda_1}{2} \exp\left(-\frac{\lambda_1}{2} \tilde{\theta}_{ii}\right) \right\} I_{\tilde{\Theta} \in M^+}, \end{aligned}$$

其中, $\tilde{\mathbf{S}} = \sum_{i=1}^n (\mathbf{y}_i - \boldsymbol{\mu}_i - \mathbf{A}\boldsymbol{\omega}_i)(\mathbf{y}_i - \boldsymbol{\mu}_i - \mathbf{A}\boldsymbol{\omega}_i)^T$ 。令 $\tilde{\boldsymbol{\eta}} = \tilde{\boldsymbol{\theta}}_i$, $\tilde{\kappa} = \tilde{\theta}_{ii} - \tilde{\boldsymbol{\theta}}_i' \tilde{\boldsymbol{\Theta}}_{ii}^{-1} \tilde{\boldsymbol{\theta}}_i$ 。根据 BCEN 模型中的分块 Gibbs 采样, 可以得到

$$\begin{aligned} \tilde{\boldsymbol{\eta}} | \tilde{\boldsymbol{\Theta}}_{ii}, \mathbf{Y}, \boldsymbol{\Omega}, \boldsymbol{\xi}, \boldsymbol{\tau}, \lambda_1, \lambda_2 &\sim N(-D\tilde{\mathbf{S}}, D), \\ \tilde{\kappa} | \tilde{\boldsymbol{\Theta}}_{ii}, \mathbf{Y}, \boldsymbol{\Omega}, \boldsymbol{\xi}, \boldsymbol{\tau}, \lambda_1, \lambda_2 &\sim \text{Gamma}\left(\frac{n}{2}+1, \frac{\tilde{\mathbf{S}}_{ii} + \lambda_1}{2}\right), \end{aligned} \quad (4.4)$$

其中, $\mathbf{D} = \{(\tilde{\mathbf{S}}_{ii} + \lambda_1)\tilde{\boldsymbol{\Theta}}_{ii}^{-1} + \mathbf{H}_\tau\}^{-1}$ 。同时, 根据式(4.4)的结果, 可以得到 $\tilde{\boldsymbol{\theta}}_i = \tilde{\boldsymbol{\eta}}$ 和 $\tilde{\theta}_{ii} = \tilde{\kappa} + \tilde{\boldsymbol{\theta}}_i' \tilde{\boldsymbol{\Theta}}_{ii}^{-1} \tilde{\boldsymbol{\theta}}_i$ 。最终估计得到协方差矩阵 $\boldsymbol{\Psi} = \boldsymbol{\Theta}^{-1}$ 。而对于 CFA 中的其他参数 $\boldsymbol{\mu}$ 、 \mathbf{A} 、 $\boldsymbol{\Phi}$ 和 $\boldsymbol{\omega}_i$ 的条件后验与 Lee(2007)和 Pan 等^[32](2017)相同

$$\boldsymbol{\mu} \sim N(\boldsymbol{\mu}_0, \mathbf{H}_{\mu_0}), \quad \mathbf{A}_j \sim N(\mathbf{A}_{0j}, \mathbf{H}_{0j}), \quad \boldsymbol{\Phi} \sim \text{Wishart}(\mathbf{R}_0, \rho_0), \quad (4.5)$$

其中, \mathbf{A}_j^T 为 \mathbf{A} 的第 j 行, $\boldsymbol{\mu}_0$ 、 \mathbf{A}_{0j} 和 ρ_0 以及正定矩阵 \mathbf{H}_{μ_0} 、 \mathbf{H}_{0j} 和 \mathbf{R}_0 为超参数, 其值根据先验信息给定。对于贝叶斯协方差弹网验证性因子分析(Bayesian covariance elastic net Confirmatory factor analysis, BCENCFA)模型中的潜在尺度参数 $\boldsymbol{\tau}$ 与惩罚参数 λ_1 和 λ_2 的先验分布分别与式(3.11)和式(3.13)相同。因此, 其满条件后验分布为

$$\boldsymbol{\omega}_i | \mathbf{Y}, \boldsymbol{\mu}, \mathbf{A}, \boldsymbol{\Phi}, \boldsymbol{\Psi} \sim N\left(\left(\boldsymbol{\Phi}^{-1} + \mathbf{A}^T \boldsymbol{\Psi} \mathbf{A}\right)^{-1} \mathbf{A}^T \boldsymbol{\Psi}^{-1} (\mathbf{y}_i - \boldsymbol{\mu}), \left(\boldsymbol{\Phi}^{-1} + \mathbf{A}^T \boldsymbol{\Psi} \mathbf{A}\right)^{-1}\right), \quad (4.6)$$

$$\boldsymbol{\mu} | \mathbf{Y}, \boldsymbol{\Omega}, \mathbf{A}, \boldsymbol{\Phi}, \boldsymbol{\Psi} \sim N\left(\left(\mathbf{H}_{\mu_0}^{-1} + n \boldsymbol{\Psi}^{-1}\right)^{-1} \left(\boldsymbol{\Psi}^{-1} V + \mathbf{H}_{\mu_0}^{-1} \boldsymbol{\mu}_0\right), \left(\mathbf{H}_{\mu_0}^{-1} + n \boldsymbol{\Psi}^{-1}\right)^{-1}\right), \quad (4.7)$$

$$\begin{aligned} \boldsymbol{\Psi}^{-1} | \mathbf{Y}, \boldsymbol{\Omega}, \boldsymbol{\mu}, \mathbf{A}, \boldsymbol{\Phi}, \boldsymbol{\tau}, \lambda_1, \lambda_2 &\sim \left| \tilde{\boldsymbol{\Theta}} \right|^{\frac{n}{2}} \exp\left[-\frac{1}{2} \text{tr}(\tilde{\mathbf{S}} \tilde{\boldsymbol{\Theta}})\right] \\ &\times \prod_{i < j} \left\{ \sqrt{\frac{2\lambda_2 \tau_{ij}}{2\pi(\tau_{ij} - 1)}} \exp\left[-\frac{\tilde{\theta}_{ij}^2}{2} \left(\frac{2\lambda_2 \tau_{ij}}{\tau_{ij} - 1}\right)\right] \right\} \\ &\times \prod_{i=1}^p \left\{ \frac{\lambda_1}{2} \exp\left(-\frac{\lambda_1}{2} \tilde{\theta}_{ii}\right) \right\} I_{\tilde{\boldsymbol{\Theta}} \in M^+}, \end{aligned} \quad (4.8)$$

$$\mathbf{A}_j | \mathbf{Y}, \boldsymbol{\Omega}, \boldsymbol{\mu}, \mathbf{A}_{-j}, \boldsymbol{\Phi}, \boldsymbol{\Psi} \sim N\left(\left(\boldsymbol{\Psi}_j^{*-1} \boldsymbol{\Omega} \boldsymbol{\Omega}^T + \mathbf{H}_{0j}^{-1}\right)^{-1} \left(\boldsymbol{\Psi}_j^{*-1} \boldsymbol{\Omega} \mathbf{Y}_j^{*T} + \mathbf{H}_{0j}^{-1} \mathbf{A}_{0j}\right), \left(\boldsymbol{\Psi}_j^{*-1} \boldsymbol{\Omega} \boldsymbol{\Omega}^T + \mathbf{H}_{0j}^{-1}\right)^{-1}\right), \quad (4.9)$$

$$\boldsymbol{\Phi} | \mathbf{Y}, \boldsymbol{\Omega}, \boldsymbol{\mu}, \mathbf{A}, \boldsymbol{\Psi} \sim IW\left(\boldsymbol{\Omega} \boldsymbol{\Omega}^T + \mathbf{R}_0^{-1}, n + \rho_0\right), \quad (4.10)$$

其中, $V = \sum_{i=1}^n (\mathbf{y}_i - \mathbf{A}\boldsymbol{\omega}_i)$, $\boldsymbol{\Psi}_j^* = \boldsymbol{\psi}_{jj} - \boldsymbol{\psi}_j^T \boldsymbol{\Psi}_{jj}^{-1} \boldsymbol{\psi}_j$, 令 \mathbf{y}_{-j} 和 $\boldsymbol{\mu}_{-j}$ 分别为 \mathbf{y}_i 和 $\boldsymbol{\mu}$ 去掉第 j 个元素后的向量, 矩阵 \mathbf{Y}_j^* 的元素 $y_{ij}^* = y_{ij} - \mu_j - \boldsymbol{\psi}_j^T \boldsymbol{\Psi}_{jj}^{-1} (\mathbf{y}_{-ij} - \boldsymbol{\mu}_{-j} - \mathbf{A}_{-j} \boldsymbol{\omega})$ 。

最后，根据满条件后验分布，BCENCFA 通过 Gibbs 采样实现参数估计。

4.2 变量选择

变量选择是弹网模型的一个重要功能。贝叶斯框架下，通常使用最高后验密度(Highest Posterior Density, HPD)区间对估计参数的不确定性进行描述(Chen 和 Shao, 1999)。其中 $100(1-\alpha)\%$ 的 HPD 区间定义为 $\mathbf{R}(\delta_\alpha) = \{x: p(x|\mathbf{Y}) \geq \delta_\alpha\}$ 。 $p(x|\mathbf{Y})$ 为后验分布、 δ_α 为对应分位数，即 $P(x \in \mathbf{R}(\delta_\alpha)) \geq 1-\alpha$ 。当利用 HPD 区间进行变量选择时，通过 MCMC 方法采样得到的样本的置信区间中包括了 0，则认为其与 0 没有显著区别，需要将该变量进行剔除。本文所考虑使用的置信区间为 0.95 水平得到 HPD 区间。

4.3 模型检验

4.3.1 Gibbs 采样的收敛性检验

使用 Gibbs 采样器进行采样时，为了保证采样结果的准确性，必须确保采样过程中马尔可夫链能够收敛。参照 Pan 等^[32](2017)，考虑通过潜在尺度缩减因子(potential scale reduction factor, PSRF, Gelman 和 Rubin^[10], 1992)判断 BCENCFA 中参数的收敛情况。

对于未知参数 ξ ，选定适当的初始值通过 Gibbs 采样得到 m 条马尔可夫链。令 ξ_{ij} 表示第 i 条马尔可夫链第 j 时刻的状态， $i=1,2,\dots,m$ ， $j=1,2,\dots,n$ 。有

$$\begin{aligned} B &= \frac{n}{m-1} \sum_{i=1}^m (\bar{\xi}_i - \bar{\xi})^2, \\ W &= \frac{1}{m} \sum_{i=1}^m S_i^2, \end{aligned} \quad (4.11)$$

其中， B 为链间方差、 W 为链内方差。 $\bar{\xi} = 1/m \sum_{i=1}^m \bar{\xi}_i$ 为总均值， $\bar{\xi}_i = 1/n \sum_{j=1}^n \xi_{ij}$ 和 $S_i^2 = 1/(n-1) \sum_{j=1}^n (\xi_{ij} - \bar{\xi}_i)^2$ 分别为链内均值和链内方差。因此，PSRF 统计量的形式为

$$\sqrt{\hat{R}} = \sqrt{\frac{n-1}{n} + \frac{1}{n} \frac{B}{W}}, \quad (4.12)$$

当所有参数的 $\sqrt{\hat{R}}$ 值小于 1.2 时, 则可认为整个马尔可夫链已经收敛, 可对参数进行估计。

4.3.2 模型的拟合优度检验

对于 CFA 模型的拟合状况, 通常采用后验预测 p 值(Posterior Predictive P-Value, P_{pp}, Gelman 等^[9], 1996)进行评估。Scheines 等^[35](1999)和 Lee^[23](2007)介绍了 P_{pp} 值对 SEM 的拟合优度的测量。Pan 等^[32](2017)将其作为一个补充统计量来评估单个 CFA 模型的拟合优度。当 P_{pp} 值在 0.5 附近时, 认为模型的拟合结果较好。

4.4 模型拓展

考虑到 BCENCFA 在进行模型选择与参数估计时, 对于残差的精度矩阵中的元素赋予了相同的惩罚参数 λ_1 和 λ_2 。而这无疑会在参数估计时造成一定程度的信息损失, 同时会缺乏对极端情况的认知。因此, 考虑将 BACEN 模型与 CFA 结合提出贝叶斯自适应协方差弹网验证性因子分析(Bayesian Adaptive Covariance Elastic Net Confirmatory Factor Analysis, BACENCFA)模型, 进一步提升 CFA 的模型选择与参数估计的能力。根据 3.5 节中提出的 BACEN 模型, 赋予 CFA 残差的精度矩阵中每个非对角元素 θ_{ij} 不同的弹网先验 $EN(\theta_{ij} | \lambda_{1,ij}, \lambda_{2,ij})$ 。根据式(3.17), BACENCFA 模型的满条件后验有

$$\begin{aligned}
\omega_i | Y, \mu, \Lambda, \Phi, \Psi &\sim N\left(\left(\Phi^{-1} + \Lambda^T \Psi \Lambda\right)^{-1} \Lambda^T \Psi^{-1} (y - \mu), \left(\Phi^{-1} + \Lambda^T \Psi \Lambda\right)^{-1}\right), \\
\mu | Y, \Omega, \Lambda, \Phi, \Psi &\sim N\left(\left(H_{\mu_0}^{-1} + n \Psi^{-1}\right)^{-1} \left(\Psi^{-1} V + H_{\mu_0}^{-1} \mu_0\right), \left(H_{\mu_0}^{-1} + n \Psi^{-1}\right)^{-1}\right), \\
\Psi^{-1} | Y, \Omega, \mu, \Lambda, \Phi, \tau, \lambda_1, \lambda_2, \lambda &\sim \left|\tilde{\Theta}\right|^{\frac{n}{2}} \exp\left[-\frac{1}{2} \text{tr}\left(\tilde{S} \tilde{\Theta}\right)\right] \\
&\quad \times \prod_{i < j} \left\{ \sqrt{\frac{2 \lambda_{2,ij} \tau_{ij}}{2 \pi (\tau_{ij} - 1)}} \exp\left[-\frac{\tilde{\theta}_{ij}^2}{2} \left(\frac{2 \lambda_{2,ij} \tau_{ij}}{\tau_{ij} - 1}\right)\right] \right\} \\
&\quad \times \prod_{i=1}^p \left\{ \frac{\lambda_{ii}}{2} \exp\left(-\frac{\lambda_{ii}}{2} \tilde{\theta}_{ii}\right) \right\} I_{\tilde{\Theta} \in M^+}, \\
\Lambda_j | Y, \Omega, \mu, \Lambda_{-j}, \Phi, \Psi &\sim N\left(\left(\Psi_j^{*-1} \Omega \Omega^T + H_{0_j}^{-1}\right)^{-1} \left(\Psi_j^{*-1} \Omega Y_j^{*T} + H_{0_j}^{-1} \Lambda_{0_j}\right), \left(\Psi_j^{*-1} \Omega \Omega^T + H_{0_j}^{-1}\right)^{-1}\right), \\
\Phi | Y, \Omega, \mu, \Lambda, \Psi &\sim IW\left(\Omega \Omega^T + R_0^{-1}, n + \rho_0\right),
\end{aligned} \tag{4.13}$$

其中，惩罚参数 λ_1 和 λ_2 的条件后验和满条件后验分布与式(3.18)和式(3.19)类似。对角元素的超参数 $\lambda = \{\lambda_{ii}\}_{i=1}^p$ 给定固定值为 $\lambda_{ii} = 1$ 。根据式(3.11)，自适应版本的潜在尺度参数 $\tau = \{\tau_{ij}\}_{i < j}$ 的条件后验分布为

$$\frac{1}{(\tau_{ij} - 1)} \sim IG\left(\frac{\lambda_{1,ij}}{2 \lambda_{2,ij} |\theta_{ij}|}, \frac{\lambda_{1,ij}^2}{2 \lambda_{2,ij}}\right). \tag{4.14}$$

通过式(4.13)和式(4.14)，实现了对 BACENCFA 的残差协方差矩阵 $\Psi = \Theta^{-1}$ 的估计，而其余参数的估计则与 BCENCFA 相同。

5. 数值模拟

为了检验所提出的两种 CFA 模型的参数估计与模型恢复的能力，本节考虑通过数值模拟将其与现有的 BCLCFA 方法进行对比，以检验模型的有效性。

5.1 模型设定

5.1.1 模拟 1

本小节指定了 3 个不同的模型，并通过这 3 个模型以检验 3 种 CFA 方法在不同维度下的参数恢复效果。由于模型的复杂度会对模型的可识别性产生影响，潜在因子数量越多越容易出现不收敛的情况(Marsh^[28], 1998)。因此，本节所考虑的 CFA 模型分别为 2 因子 10 变量、4 因子 20 变量和 4 因子 40 变量的模型。下面给出本节考虑的 3 个模型结构

模型 1 给定观察变量数 $p = 10$ ，潜在因子数 $q = 2$ 。该模型具有以下结构

$$A^T = \begin{pmatrix} 1 & \lambda_{2,1} & \lambda_{3,1} & \lambda_{4,1} & \lambda_{5,1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \lambda_{7,2} & \lambda_{8,2} & \lambda_{9,2} & \lambda_{10,2} \end{pmatrix}, \quad \Phi = \begin{pmatrix} \phi_{1,1} & \phi_{1,2} \\ \phi_{2,1} & \phi_{2,2} \end{pmatrix}$$

其中，1 和 0 为固定值，且

$$\Psi = \begin{pmatrix} \psi_{1,1} & 0 & 0 & 0 & 0 & \psi_{1,6} & 0 & 0 & 0 & 0 \\ 0 & \psi_{2,2} & 0 & 0 & 0 & 0 & \psi_{2,7} & 0 & 0 & 0 \\ 0 & 0 & \psi_{3,3} & 0 & \psi_{3,5} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \psi_{4,4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \psi_{5,3} & 0 & \psi_{5,5} & 0 & 0 & 0 & 0 & 0 \\ \psi_{6,1} & 0 & 0 & 0 & 0 & \psi_{6,6} & 0 & 0 & 0 & 0 \\ 0 & \psi_{7,2} & 0 & 0 & 0 & 0 & \psi_{7,7} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \psi_{8,8} & 0 & \psi_{8,10} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \psi_{9,9} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \psi_{10,8} & 0 & \psi_{10,10} \end{pmatrix}$$

对于该模型，其结构参数的真实值为： $\mu = (\mu_1, \mu_2, \dots, \mu_{10})^T = (0.5, 0.5, \dots, 0.5)^T$ ； $\lambda_{2,1} = \lambda_{7,2} = \lambda_{8,1} = 0.8$ ， $\lambda_{3,1} = \lambda_{4,1} = \lambda_{9,2} = 0.5$ ， $\lambda_{5,1} = \lambda_{10,2} = 0.3$ ； $\phi_{1,1} = \phi_{2,2} = 1$ ， $\phi_{1,2} = \phi_{2,1} = 0.3$ ； $\psi_{j,j} = 0.36$ ， $j = 1, 2, 6, 7, 8$ ， $\psi_{j,j} = 0.5$ ， $j = 3, 4, 5, 9, 10$ ，对于协方差矩阵非对角元素 $\psi_{1,6} = \psi_{6,1} = \psi_{2,7} = \psi_{7,2} = \psi_{3,5} = \psi_{5,3} = \psi_{8,10} = \psi_{10,8} = 0.3$ 。该模型参照了 Pan 等^[32](2017)的假设。

模型 2 给定观察变量数 $p = 20$ ，潜在因子数 $q = 4$ 。模型中的结构参数的真实值
 给定为： $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_{20})^T = (0.5, 0.5, \dots, 0.5)^T$ ； $\lambda_{2,1} = \lambda_{7,2} = \lambda_{12,3} = \lambda_{17,4} = 0.8$ ，
 $\lambda_{5,1} = \lambda_{10,2} = \lambda_{15,3} = \lambda_{20,4} = 0.4$ ，
 $\lambda_{3,1} = \lambda_{8,2} = \lambda_{13,3} = \lambda_{18,4} = \lambda_{4,1} = \lambda_{9,2} = \lambda_{14,3} = \lambda_{19,4} = 0.6$ ；
 $\phi_{j,j} = 1$ ， $j = 1, 2, 3, 4$ ， $\phi_{i,j} = 0.3$ ， $i \neq j$ ； $\psi_{j,j} = 0.6$ ， $j = 1, 2, \dots, 20$ 、对于协方差
 矩阵中的非对角元素 $\psi_{i,j} = \psi_{j,i} = 0.36$ ， $i = 3, 5, 6, 7, 10, 11$ ， $j = 6, 10, 9, 15, 13, 16$ 。
 其载荷矩阵结构为

$$A^T = \begin{pmatrix} 1 & \lambda_{2,1} & \dots & \lambda_{10,1} & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & \lambda_{12,2} & \dots & \lambda_{20,2} & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & \lambda_{22,3} & \dots & \lambda_{30,3} & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & \lambda_{32,4} & \dots & \lambda_{40,4} \end{pmatrix}$$

模型 3 给定观察变量数 $p = 40$ ，潜在因子数 $q = 4$ 。模型中的结构参数的真实值
 给定为： $\boldsymbol{\mu} = (\mu_1, \mu_2, \dots, \mu_{40})^T = (0.5, 0.5, \dots, 0.5)^T$ ；对于因子载荷 $\lambda_{i,1}$ 、 $\lambda_{i+10,2}$ 、
 $\lambda_{i+20,3}$ 和 $\lambda_{i+30,4}$ ，当下标 i 分别为 $i = 2, 7$ 、 $i = 3, 4, 9$ 、 $i = 5, 10$ 和 $i = 6, 8$ 时，对应的
 值给定 $\lambda_{i,1} = \lambda_{i+10,2} = \lambda_{i+20,3} = \lambda_{i+30,4}$ 分别 0.8、0.6、0.4 和 0.5； $\phi_{j,j} = 1$ ，
 $j = 1, 2, 3, 4$ 、 $\phi_{i,j} = 0.3$ ， $i \neq j$ ； $\psi_{j,j} = 0.6$ ， $j = 1, 2, \dots, 40$ 、对于协方差矩阵中的
 非 对 角 元 素 $\psi_{i,j} = \psi_{j,i} = 0.4$ ， $i = 6, 7, 8, 15, 19, 25, 28, 36, 37, 40$ ，
 $j = 5, 4, 5, 10, 11, 1, 25, 20, 33, 13$ 、 $\psi_{i,j} = \psi_{j,i} = 0.3$ ， $i = 4, 12, 20, 30, 32, 33, 34$ ，
 $j = 2, 10, 15, 4, 6, 30, 2$ 。其载荷矩阵结构为

$$A^T = \begin{pmatrix} 1 & \lambda_{2,1} & \dots & \lambda_{10,1} & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 1 & \lambda_{12,2} & \dots & \lambda_{20,2} & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & \lambda_{22,3} & \dots & \lambda_{30,3} & 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 1 & \lambda_{32,4} & \dots & \lambda_{40,4} \end{pmatrix}$$

5.1.2 模拟 2

本小节主要检验所提出的 2 种 CFA 方法对残差协方差矩阵中真正的非零项
 的识别能力。根据 Pan 等^[32](2017)的模拟研究，令变量 $p = 10$ 、潜在因子
 $q = 2$ ，其载荷矩阵 A 具有以下结构

$$A^T = \begin{pmatrix} 1 & \lambda_{2,1} & \lambda_{3,1} & \lambda_{4,1} & \lambda_{5,1} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & \lambda_{7,2} & \lambda_{8,2} & \lambda_{9,2} & \lambda_{10,2} \end{pmatrix}, \quad \Phi = \begin{pmatrix} \phi_{1,1} & \phi_{1,2} \\ \phi_{2,1} & \phi_{2,2} \end{pmatrix}$$

同时，令残差协方差矩阵的非对角元素 $\psi_{1,6}$ 和 $\psi_{2,7}$ 为非零元素，有

$$\Psi = \begin{pmatrix} \psi_{1,1} & 0 & 0 & 0 & 0 & \psi_{1,6} & 0 & 0 & 0 & 0 \\ 0 & \psi_{2,2} & 0 & 0 & 0 & 0 & \psi_{2,7} & 0 & 0 & 0 \\ 0 & 0 & \psi_{3,3} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \psi_{4,4} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \psi_{5,5} & 0 & 0 & 0 & 0 & 0 \\ \psi_{6,1} & 0 & 0 & 0 & 0 & \psi_{6,6} & 0 & 0 & 0 & 0 \\ 0 & \psi_{7,2} & 0 & 0 & 0 & 0 & \psi_{7,7} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \psi_{8,8} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \psi_{9,9} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \psi_{10,10} \end{pmatrix}$$

其中，因子载荷取值为{0.5,0.8}、因子协方差取值为{0.3,0.7}、残差协方差取值为{0.0,0.3,0.7}。

5.1.3 模拟 3

由于 CFA 通常在给定载荷矩阵时会进行严格的假设。而对于存在交叉载荷的情况，这些假设可能导致错误的模型结果。本小节根据 Pan 等^[32](2017)所设计的诊断工具对 BCENCFA 和 BACENCFA 进行检验，探究在模型被错误设定时两种方法的表现。根据 Pan 等^[32](2017)的模拟研究，在生成样本量为 n 的数据集时，载荷矩阵 A' 具有以下结构

$$A'^T = \begin{pmatrix} 1 & 0.8 & 0.8 & 0 & 0 & 0 \\ 0 & \lambda_{2,2} & 0 & 1 & 0.8 & 0.8 \end{pmatrix}$$

其中，变量 2 是因子 1 和 2 的交叉载荷，其载荷分别为 0.8 和 $\lambda_{2,2}$ 。同时， $\lambda_{2,2}$ 的取值分别为 $\{\lambda_{2,2} = 0.3, n = 250\}$ 、 $\{\lambda_{2,2} = 0.3, n = 500\}$ 、 $\{\lambda_{2,2} = 0.8, n = 250\}$ 和 $\{\lambda_{2,2} = 0.8, n = 500\}$ 。残差协方差矩阵 Ψ 则假设为对角矩阵。此外，在分析数据集时，使用没有交叉载荷的载荷矩阵 A ，使其具有以下结构

$$A^T = \begin{pmatrix} 1 & \lambda_{2,1} & \lambda_{3,1} & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \lambda_{5,2} & \lambda_{6,2} \end{pmatrix}$$

5.2 评估指标

对于以上 3 个模型，考虑重复进行 100 次模拟模拟试验，前两项模拟均生

成一组样本量 $n = 500$ 的数据集。对于每次模拟通过 Gibbs 采样进行 10000 次采样，并将前 5000 次采样结果视作燃烧期舍弃。同时，通过 5 项指标对模拟结果进行评估，比较不同的 CFA 方法的表现。这 5 项指标分别为偏差(bias)、标准误(Standard Error, SE)、均方根误差(Root Mean Squared Error, RMSE)、Stein 损失以及 MCC。其中，前 3 项指标分布计算估计值的偏差、估计均值的标准误以及估计值与真实值之间的均方根误差，用于分析模拟试验中对模型所有参数估计结果，而后 2 项指标用于比较不同方法对于模型的协方差矩阵的估计以及恢复能力。

5.3 结果分析

5.3.1 模拟 1 结果分析

对于本节设定的模型 1、模型 2 和模型 3，通过 bias、SE 和 RMSE 检验 3 种方法对真实值的估计。为了更加直观的展示模拟试验中 3 种方法的表现，根据模拟试验的结果，将模型设定中真实值不为 0 的模型参数作为横坐标，将计算得到的 3 项指标结果作为纵坐标，得到

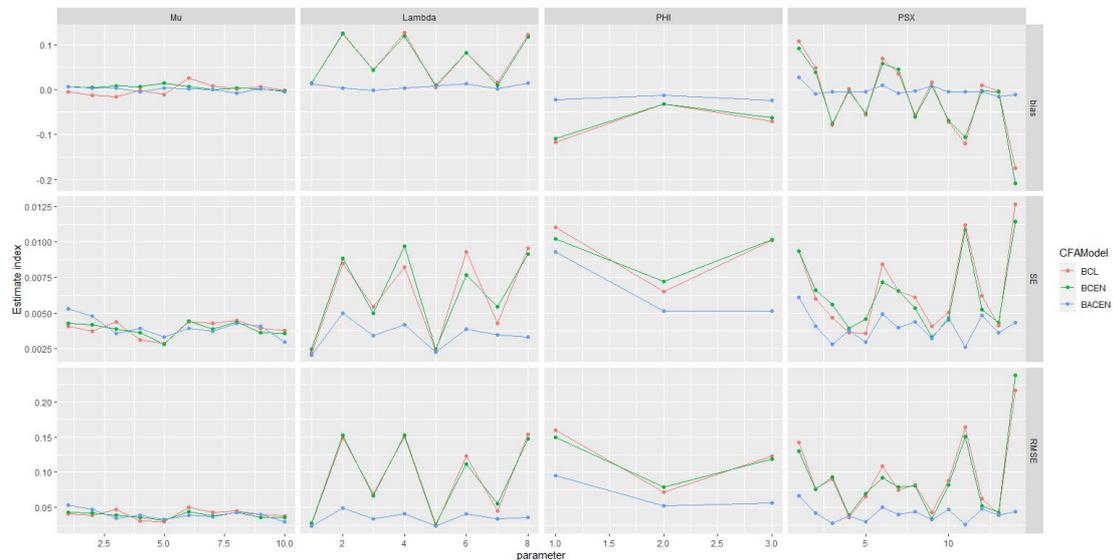


图 5.1 模型 1 估计结果

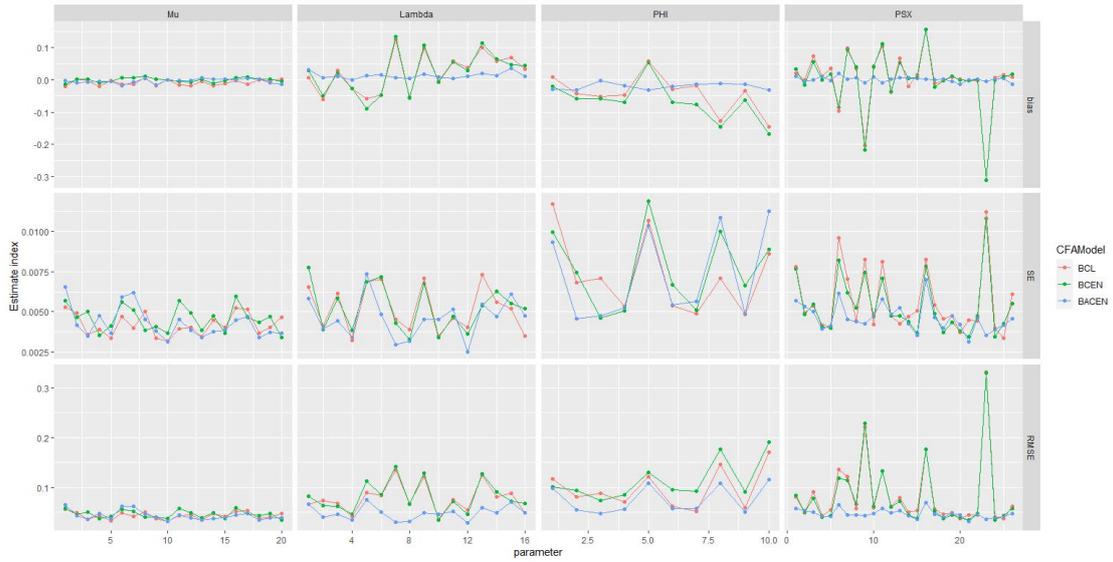


图 5.2 模型 2 估计结果

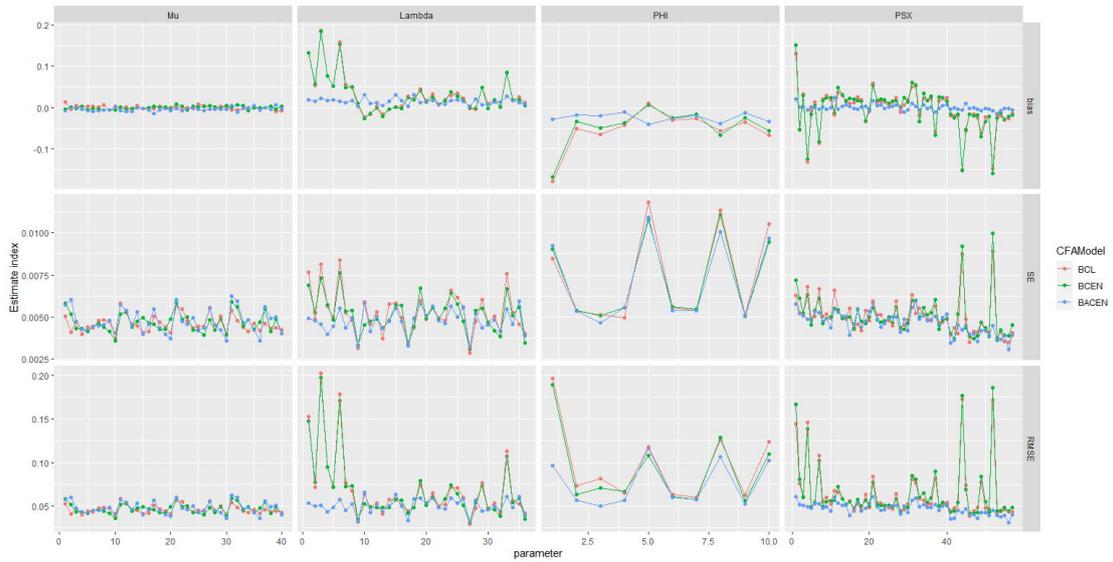


图 5.3 模型 3 估计结果

根据图 5.1-5.3 的结果可知，针对 CFA 模型中的截距项 μ ，3 种方法的估计结果非常接近，在设定的 3 个模型中的 3 种方法的估计结果并无显著差别，且估计的精度并没有随着变量增加发生变化；针对因子载荷 Λ ，两种非自适应的方法的估计结果非常相似，同时随着变量的增加 BCENCFA 的 SE 和 RMSE 结果要略微优于 BCLCFA，而 BACENCFA 在 3 个模型的模拟试验中均表现出了良好的性能，同时相对于另外两种方法，BACENCFA 更加稳定，且估计精度也更高；针对潜在因子 Ω ，对比模型 1 和模型 2 的结果，当潜在因子数增加时 3 种方法的估计精度均有所下降，而对比模型 2 和模型 3 的结果，变量的增加并

没有对参数估计产生较大的影响，同时在 3 种方法的对比中 BACENCFA 依旧估计精度最高的方法；针对残差协方差矩阵 Ψ ，两种非自适应的方法估计的结果同样非常相似且易出现较差的估计结果，BACENCFA 的估计结果相较于另外两种方法更加稳定，同时整体上精度更高。

上述的讨论仅仅针对了 3 个模型中非零的参数进行评估分析，而这并不能保证残差协方差矩阵整体的参数估计能够令人满意。因此，为了对模型整体的残差协方差矩阵的参数估计和变量选择的结果进行评估。考虑通过 Stein 损失和 MCC 进一步对残差协方差矩阵进行对比分析，比较 3 种 CFA 方法的性能。对于本节设定的 3 个模型的 100 模拟试验的 Stein 损失结果的中位数和对应的标准差如表 5.1 所示

表 5.1 Stein 损失结果

Stein 损失	M1	M2	M3
BCLCFA	1.0520(0.6752)	1.3789(0.5191)	1.0349(0.4803)
BCENCFA	1.1102(0.5137)	1.3467(0.4497)	1.0604(0.4849)
BACENCFA	0.0981(0.2815)	0.0952(0.0411)	0.3327(0.1665)

注：“M1”、“M2”和“M3”分别为模型 1、模型 2 和模型 3

对于模型的学习性能的评估，通过 MCC 进行检验。对于 MCC，得到的数值越大说明模型的还原效果越好。表 5.2 给出了 100 次模拟试验的 MCC 的均值

表 5.2 MCC 结果

MCC(%)	M1	M2	M3
BCLCFA	90.30	86.25	87.29
BCENCFA	90.71	86.41	87.60
BACENCFA	98.69	99.12	97.39

从表 5.1 可知，当残差协方差矩阵是稀疏矩阵时，BCLCFA 和 BCENCFA 在对矩阵的恢复上并没有显著的差别。根据 Stein 损失的标准差，BCENCFA 对模型的残差协方差矩阵的参数估计要优于 BCLCFA。而 BACENCFA 的 Stein 损失结果则是显著优于另外两种方法的结果。同时，从表 5.2 可知，BACENCFA 的分类效果要优于另外两种方法。

综上所述，对于模拟 1 的 3 个因子载荷不存在交叉载荷且残差协方差矩阵存在少量非对角元素的 CFA 模型，BACENCFA 表现出了良好的性能，而

BCLCFA 和 BCENCFA 并没有显著差异。

5.3.2 模拟 2 结果分析

为了检验 BCENCFA 和 BACENCFA 对于显著残差协方差的能力和不同情形下两种方法之间的差异，根据所设置的条件，每种方法分别进行了 12 次模拟试验。对于所设置的两个显著的残差协方差项，其识别情况如下表所示

表 5.3 显著残差协方差识别结果

因子载荷	因子协方差	残差协方差	BCLCFA		BCENCFA		BACENCFA	
			Rate1	Rate2	Rate1	Rate2	Rate1	Rate2
0.5	0.3	0.3	0.95	0.99	0.95	1	0.71	1
0.5	0.3	0.7	1	1	1	1	1	1
0.5	0.7	0.3	0.8	1	0.86	1	0.18	0.97
0.5	0.7	0.7	1	1	1	1	0.92	1
0.8	0.3	0.3	0.99	0.94	0.97	1	0.97	1
0.8	0.3	0.7	1	1	1	1	1	1
0.8	0.7	0.3	0.92	0.99	0.9	0.98	0.81	0.99
0.8	0.7	0.7	1	1	1	1	1	1
0.5	0.3	0	—	—	—	—	—	—
0.5	0.7	0	—	—	—	—	—	—
0.8	0.3	0	—	—	—	—	—	—
0.8	0.7	0	—	—	—	—	—	—

注：“Rate1”为 $\psi_{1,6}$ 的显著次数数/模拟次数；“Rate2”为 $\psi_{2,7}$ 的显著次数/模拟次数。

根据表 5.3 的结果可知，BCLCFA 和 BCENCFA 在残差协方差较小时的识别能力要较弱于残差协方差较大的情况。而 BACENCFA 对于残差协方差的识别能力在因子载荷与残差协方差较小时不能进行有效的识别。此外，对于整体残差协方差矩阵的估计结果通过 Stein 损失和 MCC 进行对比，如表 5.4 和表 5.5 所示

表 5.4 模拟 2 Stein 损失结果

因子载荷	因子协方差	残差协方差	BCLCFA	BCENCFA	BACENCFA
0.5	0.3	0.3	0.0860(0.0370)	0.0943(0.0309)	0.0847(0.0480)
0.5	0.3	0.7	0.1880(0.0557)	0.1938(0.0579)	0.0693(0.0478)
0.5	0.7	0.3	0.0691(0.0438)	0.0731(0.0441)	0.1519(0.1311)
0.5	0.7	0.7	0.1095(0.1374)	0.1041(0.1183)	0.0542(0.2135)
0.8	0.3	0.3	0.0579(0.0329)	0.0544(0.0302)	0.0479(0.0328)
0.8	0.3	0.7	0.1105(0.0604)	0.1156(0.0529)	0.0508(0.0591)

0.8	0.7	0.3	0.0367(0.0300)	0.0388(0.0351)	0.0470(0.0472)
0.8	0.7	0.7	0.0667(0.0498)	0.0613(0.0477)	0.0409(0.0212)
0.5	0.3	0	0.0685(0.0267)	0.0713(0.0274)	0.0585(0.0383)
0.5	0.7	0	0.0519(0.0279)	0.0454(0.0320)	0.0380(0.0424)
0.8	0.3	0	0.0479(0.0193)	0.0474(0.0190)	0.0451(0.0198)
0.8	0.7	0	0.0304(0.0158)	0.0266(0.0137)	0.0338(0.0182)

表 5.5 模拟 2MCC 结果

因子载荷	因子协方差	残差协方差	BCLCFA	BCENCFA	BACENCFA
0.5	0.3	0.3	0.9796	0.9815	0.9739
0.5	0.3	0.7	0.9667	0.9565	0.9877
0.5	0.7	0.3	0.9785	0.9785	0.9261
0.5	0.7	0.7	0.9357	0.9396	0.9862
0.8	0.3	0.3	0.9864	0.9893	0.9975
0.8	0.3	0.7	0.9778	0.974	0.9876
0.8	0.7	0.3	0.9916	0.989	0.983
0.8	0.7	0.7	0.9946	0.9913	0.9992
0.5	0.3	0	0.9807	0.9817	0.9961
0.5	0.7	0	0.9856	0.9858	0.999
0.8	0.3	0	0.9903	0.9903	0.999
0.8	0.7	0	0.9932	0.9951	0.9981

结合表 5.4 和表 5.5 的结果可知，当残差协方差较大时，BACENCFA 要优于另外两种方法。而当残差协方差矩阵为对角矩阵时，3 种方法并没有明显的差距。

5.3.3 模拟 3 结果分析

当 $\lambda_{2,2}$ 的值被错误的视作 0 时，将对同一潜在因子下的其他变量产生影响 (Pan 等^[32], 2017)。其往往导致该变量与同一潜在因子下的其余变量具有显著的残差协方差。因此，对模拟中所有残差协方差项出现显著的残差协方差项的次数进行统计，并计算显著项出现次数与总模拟次数的比值。其结果如图 5.4 所示

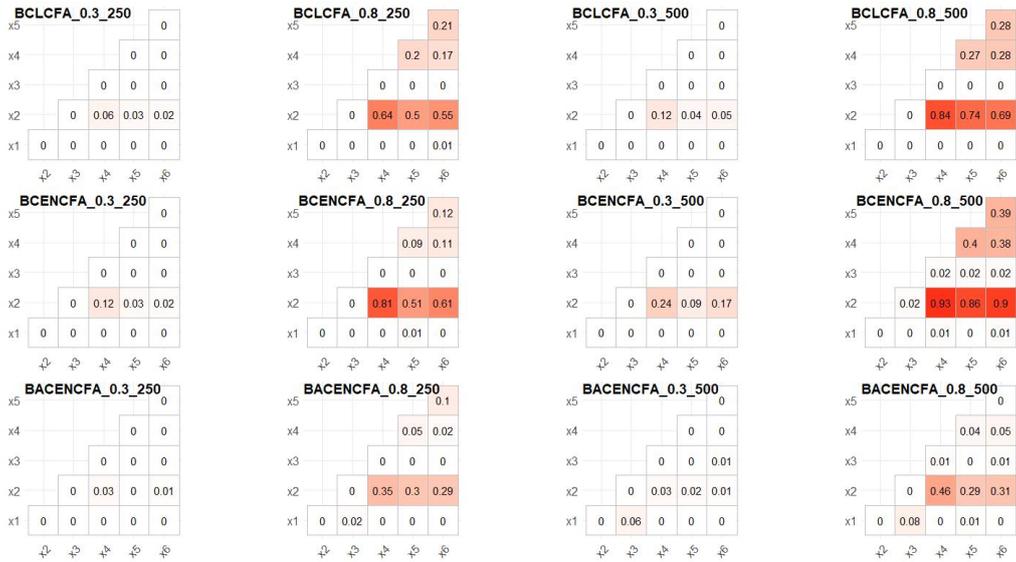


图 5.4 残差协方差矩阵识别结果

其中， $x_1 \sim x_6$ 分别对应变量 1 到变量 6。对比 3 种方法的模拟结果，显著项出现次数最多的是残差协方差项 $\psi_{2,4}$ 、 $\psi_{2,5}$ 和 $\psi_{2,6}$ ，其次是 $\psi_{4,5}$ 、 $\psi_{4,6}$ 和 $\psi_{5,6}$ 。同时，当 $\lambda_{2,2}$ 取值的较大或样本量增加时，显著残差协方差出现的频率也随之增加。在 3 种方法中，BCENCFA 在模拟时显著项出现的频率要高于 BCLCFA，而 BACENCFA 在模拟时显著项出现的频率要明显低于另外两种方法。

综上所述，当 $\lambda_{2,2}$ 的值被视为 0 处理时，其错误一定程度上会在残差协方差矩阵中有所表现。通过对所得到的残差协方差矩阵的观察，能够在一定程度上帮助研究者检验模型结构是否有误。

6 实例分析

本节将通过两个实例数据，进一步检验 BCENCFA 和 BACENCFA 方法在参数估计与模型选择方法的优越性和对真实数据的适应能力。

6.1 数据介绍

6.1.1 幽默风格问卷

人类在情感表达、社交和应对压力时，幽默一直扮演着重要的角色。而对于不同个体，人与人之间的幽默感和幽默风格也是各具特色。同时，幽默也是一个多维度的概念，它可以是不同类型心理特征的一部分(Martin 等^[29]，2003)。其中，幽默感是一种重要的心理特质。Martin 等^[29] (2003)指出幽默感能够被概念化为：1. 认知能力(a cognitive ability)；2. 审美反应(an aesthetic response)；3. 习惯性的行为模式(an habitual behavior pattern)；4. 与情绪相关的气质特征(an emotion-related temperament trait)；5. 一种态度(an attitude)；6. 应对策略或防御机制(a coping strategy or defense mechanism)。但幽默并不总代表着一个健康的心理行为，如讽刺他人时，就是一种咄咄逼人的幽默。对此，Martin 等^[29] (2003)介绍了用于测量幽默风格类型的幽默风格问卷(Humor Styles Questionnaire, HSQ)。HSQ 通过自我报告测量方法，从多个维度对幽默感进行测量。在衡量幽默时，其将幽默分成 4 个维度的不同幽默风格：1. 亲和幽默(Affiliative humor)；2. 自我提升的幽默(Self-enhancing humor)；3. 咄咄逼人的幽默(Aggressive humor)；4. 自我挫败的幽默(Self-defeating humor)。同时，在每个维度下设置了 8 个问题，通过这些问题对幽默风格进行评估。作为一个广泛使用的量表，其信度和效度得到过许多研究者的检验。Chen 等^[4](2020)使用了该问卷数据检验了所提出的部分验证性因子分析方法(Partially Confirmatory Factor Analysis, PCFA)的性能。此外，HSQ 的具体内容详见附表 1。

本节所使用 HSQ 量表的数据来源于开源网站 Open-Source Psychometrics Project(<https://openpsychometrics.org/>)。HSQ 数据包含了 32 个条目，每个条目都假设只对应了一个幽默风格。量表采用 5 点李克特量表评分方式，其中“1”表

示完全不同意，“2”表示略微不同意，“3”表示中立，“4”表示略微同意，“5”表示完全同意。同时，HSQ共有1071条有效数据，对于并未选择答案的条目视为缺失值，用“-1”表示。

对于HSQ量表实例分析，选取年龄在40岁以下的男性与女性认为自己的答案准确性在90以上的数据。此外，由于所提出的两个方法并未考虑缺失数据的情况且缺失数据同样会对3种CFA方法的模型选择和参数估计产生影响。因此，在删除缺失数据后，将剩余的503条数据作为本次实例分析的样本。最后，根据Martin等^[29](2003)建议，对部分条目进行处理。

6.1.2 肯塔基正念技能清单

肯塔基正念技能清单(Kentucky Inventory of Mindfulness Skills, KIMS)由Baer等^[2](2004)所提出，其目的是对正念进行测量。KIMS总共包含了39个条目，Baer等(2004)将其分为4个维度，分别为：1.观察(Observe)；2.描述(Describe)；3.有意识的行动(Act With Awareness)；4.不加判断的接受(Accept Without Judgment)。对于KIMS的具体内容，详见附表2。

对于所使用的KIMS数据同样来源于开源网站Open-Source Psychometrics Project。量表采用5点李克特量表评分方式，其中“1”表示完全不同意，“2”表示略微不同意，“3”表示中立，“4”表示略微同意，“5”表示完全同意。对于未被选择的题目则由“0”表示。

对于KIMS的实例分析，在去除所有缺失值(包括性别选项中的缺失值)后，将剩余的517条数据作为本次实例分析的样本。并根据Baer等(2004)的建议，对部分问题的得分进行处理。

6.2 结果分析

6.2.1 HSQ 结果分析

为了进一步减少模型误差，对所用的实例数据参照Chen等^[4](2020)中C-step的形式放松因子载荷，并对比3种方法的模型选择与参数估计能力。对于HSQ的估计结果的拟合优度通过PPp值进行评估。其中，BCLCFA的PPp值为

0.4934、BCENCFA 的 PPp 值为 0.5126、BACENCFA 的 PPp 值为 0.4236。相较于两种非自适应的方法，BACENCFA 的拟合优度较低。此外，对于 3 种方法的 Gibbs 采样的收敛性通过 PSRF 检验，得到

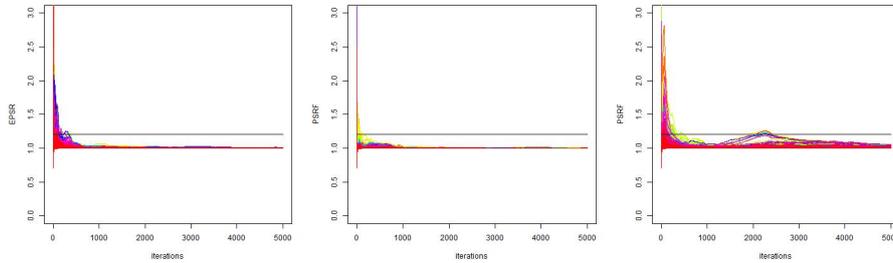


图 6.1 PSRF 收敛性诊断图

如图 6.1 所示的收敛性诊断图从左到右分别为 BCLCFA、BCENCFA 和 BACENCFA 的 PSRF 结果。由图 6.1 可以看出，3 种方法在迭代 5000 次后均能达到收敛，且 BCLCFA 和 BCENCFA 的收敛速度非常快，而 BACENCFA 的收敛则要相对差一些。结合 PPp 值和图 6.1 结果，认为 3 种方法对于 HSQ 的模型拟合结果是良好的。对于 HSQ 数据在 3 种方法下的因子载荷估计结果如表 6.1 所示

表 6.1 因子载荷估计结果

	BCLCFA				BCENCFA				BACENCFA			
	F1	F2	F3	F4	F1	F2	F3	F4	F1	F2	F3	F4
1	0.652				0.659				0.654			
5	0.593				0.599				0.591			
9	0.528				0.526				0.526			
13	0.575				0.578				0.578			
17	0.816				0.815				0.816			
21	0.717				0.717				0.772		-0.183	
25	0.554				0.558				0.570			
29	0.459				0.458				0.546			-0.174
2		0.665				0.662				0.663		
6		0.477				0.498				0.497		
10		0.758				0.712				0.761		
14		0.714				0.733				0.713		
18		0.743				0.688				0.741		
22		0.483				0.492				0.468		
26		0.679				0.704				0.698		
30		0.386				0.388				0.394		

3		0.607		0.625		0.573	0.170
7		0.530		0.533		0.556	
11		0.492		0.496		0.532	
15		0.691		0.688		0.674	
19	0.278	0.430	0.267	0.433	0.270	0.452	
23		0.560		0.563		0.540	
27		0.560		0.554		0.564	
31		0.663		0.669		0.684	
4		0.603		0.604		0.606	
8		0.758		0.754		0.803	
12		0.653		0.654		0.645	
16		0.613		0.612		0.609	
20		0.724		0.718		0.771	
24		0.391		0.42	-0.192	0.512	
28		0.371		0.398	0.221	0.180	0.274
32		0.718		0.728		0.714	

注：“F1”为 Affiliative humor；“F2”为 Self-enhancing humor；“F3”为 Aggressive humor；“F4”为 Self-defeating humor。

表 6.1 展示了 HSQ 数据在 3 种方法下的因子载荷估计的标准化结果。根据表 6.1 结果可知，BCLCFA 和 BCENCFA 的因子载荷结构是相同的，这 2 种方法均只有 1 个显著的交叉载荷，即增加了潜在因子 F1 和条目 19 之间的交叉载荷。而 BACENCFA 则产生了 7 个显著的交叉载荷，但 3 种方法并未出现数值过大的交叉载荷。表 6.2 给出了 HSQ 数据在 3 种方法下的残差协方差矩阵的标准化估计均值、标准差以及 HPD 区间

表 6.2 残差协方差估计结果

	BCLCFA		BCENCFA		BACENCFA	
	est	HPD	est	HPD	est	HPD
(6,5)	0.223(0.040)	(0.143,0.297)	0.219(0.041)	(0.138,0.298)	0.219(0.041)	(0.135,0.297)
(13,6)	0.101(0.044)	(0.012,0.188)	0.097(0.045)	(0.009,0.186)		
(16,15)	0.161(0.052)	(0.053,0.256)	0.171(0.053)	(0.063,0.274)		
(18,10)			0.284(0.050)	(0.180,0.378)		
(19,6)	0.098(0.043)	(0.015,0.180)	0.101(0.042)	(0.017,0.184)		
(19,11)	0.186(0.048)	(0.097,0.281)	0.184(0.047)	(0.090,0.274)		
(20,8)	0.307(0.053)	(0.201,0.407)	0.319(0.052)	(0.212,0.418)		
(22,16)	0.131(0.049)	(0.030,0.225)	0.147(0.047)	(0.055,0.239)		
(24,4)	0.154(0.045)	(0.063,0.237)	0.134(0.046)	(0.037,0.227)		
(25,13)	0.259(0.046)	(0.165,0.346)	0.256(0.046)	(0.167,0.343)	0.272(0.044)	(0.185,0.353)
(25,24)	-0.128(0.045)	(-0.218,-0.042)	-0.129(0.044)	(-0.217,-0.045)		
(28,3)	0.160(0.047)	(0.068,0.249)				
(28,26)	0.147(0.046)	(0.051,0.232)				
(28,27)	0.143(0.046)	(0.053,0.230)	0.145(0.047)	(0.057,0.239)		

(29,5)	0.217(0.043)	(0.131,0.299)	0.215(0.045)	(0.134,0.305)	0.217(0.046)	(0.123,0.303)
(29,22)	0.080(0.045)	(-0.011,0.168)	0.085(0.046)	(-0.006,0.177)		
(29,24)	-0.129(0.045)	(-0.216,-0.041)	-0.139(0.042)	(-0.219,-0.059)		
(30,6)	0.312(0.040)	(0.233,0.387)	0.305(0.041)	(0.221,0.381)	0.315(0.041)	(0.233,0.396)
(30,13)	0.128(0.045)	(0.040,0.214)	0.124(0.045)	(0.034,0.210)		
(31,22)	0.119(0.049)	(0.024,0.216)				
(32,24)	0.134(0.048)	(0.041,0.229)				

注：第一列为残差协方差非对角元素对应坐标；“est”为对应估计值。

根据表 6.2 的残差协方差非对角元素估计结果可知，对于 HSQ 数据，BCLCFA 方法增加了 20 个显著的残差协方差项、BCENCFA 增加了 18 个显著的残差协方差项、BACENCFA 仅增加了 4 个显著的残差协方差项。

由于交叉载荷和残差协方差都对模型的拟合度产生影响(Chen 等^[4], 2020)。结合表 6.1 和表 6.2 的结果，对于载荷矩阵和残差协方差矩阵，BACENCFA 方法通过放松假设而增加的自由参数最少。而 BCENCFA 方法在估计得到的因子载荷结构与 BCLCFA 方法相同的情况下，得到了更加稀疏的残差协方差矩阵。因此，对于 HSQ 数据，BACENCFA 方法在拟合的时拥有着更加简洁的模型结构。

6.2.2 KIMS 结果分析

KIMS 同样放松其因子载荷 (Chen 等^[4], 2020)，并对比 3 种方法的模型选择与参数估计结果。根据 KIMS 数据的拟合结果，BCLCFA 的 PPp 值为 0.5028、BCENCFA 的 PPp 值为 0.5162、BACENCFA 的 PPp 值为 0.4824。同时，图 6.2 给出了模型 PSRF 的检验结果

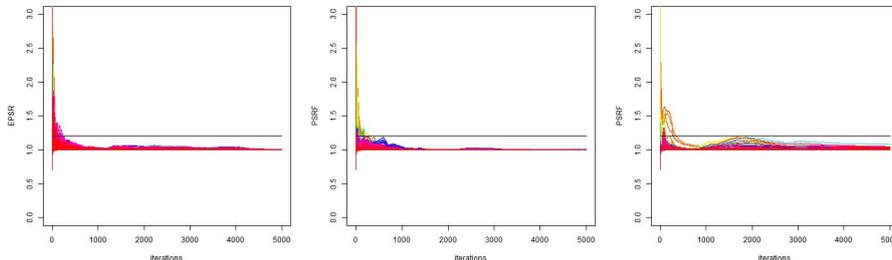


图 6.2 PSRF 收敛性诊断图

如图 6.2 所示的收敛性诊断图从左到右分别为 BCLCFA、BCENCFA 和 BACENCFA 的 PSRF 结果。由图 6.2 可以看出，3 种方法在迭代 5000 次后均能达到收敛，且 BCLCFA 和 BCENCFA 的收敛速度非常快，同样 BACENCFA 的

收敛则要相对差一些。结合 PPp 值和图 6.2 结果，认为 3 种方法对于 KIMS 的模型拟合结果是良好的。对于 KIMS 数据在 3 种方法下的因子载荷估计结果如表 6.3 所示

表 6.3 因子载荷估计结果

	BCLCFA				BCENCFA				BACENCFA			
	F1	F2	F3	F4	F1	F2	F3	F4	F1	F2	F3	F4
1	0.507				0.504				0.484			
5	0.594				0.606				0.568			
9	0.586				0.633				0.590			
13	0.569				0.564				0.534			
17	0.583				0.569				0.552			
21	0.731				0.737				0.721			
25	0.594				0.596				0.644			
29	0.561				0.586				0.685			
30	0.566				0.531				0.545			
33	0.525				0.572				0.659			
37	0.546				0.540				0.546			
39	0.483				0.456				0.463			
2		0.782				0.753				0.771		
6		0.763				0.759				0.778		
10		0.693			0.339	0.546			0.353	0.560		
14		0.745				0.788				0.750		
18		0.790				0.827				0.783		
22		0.682				0.684				0.628		0.148
26		0.722				0.694				0.714		
34		0.581				0.563				0.629		
3			0.699				0.680				0.699	
7			0.582				0.576				0.568	
11			0.383				0.381		0.190		0.339	
15			0.355				0.354				0.343	
19			0.363				0.369				0.329	
23			0.700				0.698				0.703	
27			0.587				0.596			-0.156	0.671	
31			0.484				0.479				0.477	
35			0.755				0.753				0.743	
38			0.513				0.540				0.467	
4				0.638				0.626				0.666
8				0.241	-0.293			0.277	-0.329			0.299
12				0.798				0.795				0.809
16				0.878				0.876				0.873
20				0.756				0.759				0.770
24				0.564				0.560				0.582
28				0.898				0.897				0.891
32				0.896				0.893				0.889
36				0.719				0.725				0.737

注：“F1”为 Observe items；“F2”为 Describe items；“F3”为 Act With Awareness items；“F4”为 Accept Without Judgment items。

根据表 6.3 的因子载荷与交叉载荷估计结果，BCLCFA 没有交叉载荷、BCENCFA 有两项交叉载荷且均与 $F1$ 有关，而 BACENCFA 则有 5 项交叉载荷。此外，对与 KIMS 数据的残差协方差估计结果的标准化估计均值、标准差以及 HPD 区间在表 6.4 给出

表 6.4 残差协方差估计结果

	BCLCFA		BCENCFA		BACENCFA	
	est	HPD	est	HPD	est	HPD
(5,1)	0.352(0.040)	(0.272,0.428)	0.359(0.042)	(0.276,0.437)	0.380(0.041)	(0.297,0.457)
(9,1)					0.094(0.043)	(0.005,0.172)
(9,5)	0.140(0.042)	(0.057,0.224)			0.199(0.042)	(0.116,0.277)
(9,8)	-0.153(0.039)	(-0.228,-0.073)	-0.107(0.042)	(-0.193,-0.031)		
(11,7)	-0.187(0.046)	(-0.277,-0.098)	-0.198(0.047)	(-0.289,-0.101)	-0.218(0.047)	(-0.304,-0.126)
(11,9)	0.110(0.041)	(0.029,0.190)	0.113(0.044)	(0.026,0.198)		
(12,4)	0.245(0.043)	(0.159,0.326)	0.245(0.043)	(0.160,0.329)		
(13,7)			0.122(0.042)	(0.038,0.204)	0.112(0.043)	(0.031,0.197)
(13,9)	0.278(0.041)	(0.197,0.357)	0.279(0.042)	(0.192,0.355)	0.311(0.042)	(0.225,0.389)
(15,7)	0.151(0.047)	(0.060,0.243)	0.174(0.045)	(0.083,0.256)	0.184(0.048)	(0.090,0.278)
(15,10)	0.068(0.043)	(-0.021,0.149)	0.095(0.044)	(0.012,0.185)		
(15,11)					-0.066(0.045)	(-0.149,0.027)
(15,13)	0.171(0.043)	(0.083,0.249)	0.197(0.042)	(0.118,0.279)	0.163(0.040)	(0.085,0.241)
(18,14)	0.268(0.051)	(0.161,0.360)			0.285(0.048)	(0.189,0.374)
(19,4)	-0.028(0.033)	(-0.094,0.033)	-0.025(0.033)	(-0.091,0.037)		
(19,7)	0.444(0.036)	(0.368,0.511)	0.448(0.038)	(0.372,0.523)	0.471(0.037)	(0.393,0.537)
(19,11)	-0.181(0.043)	(-0.267,-0.098)	-0.188(0.042)	(-0.263,-0.099)	-0.207(0.043)	(-0.287,-0.120)
(19,13)	0.105(0.036)	(0.031,0.175)	0.152(0.041)	(0.069,0.231)	0.140(0.041)	(0.061,0.221)
(19,15)	0.165(0.046)	(0.079,0.264)	0.166(0.042)	(0.082,0.244)	0.200(0.044)	(0.112,0.286)
(20,8)	0.224(0.042)	(0.143,0.304)	0.187(0.045)	(0.096,0.273)		
(20,17)	-0.113(0.044)	(-0.198,-0.026)	-0.133(0.045)	(-0.221,-0.048)		
(21,9)					0.176(0.049)	(0.076,0.269)
(21,13)	0.224(0.047)	(0.131,0.313)	0.232(0.047)	(0.138,0.322)	0.280(0.045)	(0.190,0.365)
(21,15)	0.099(0.048)	(0.005,0.200)	0.109(0.047)	(0.016,0.202)		
(23,14)	0.185(0.045)	(0.098,0.275)	0.218(0.051)	(0.116,0.318)		
(23,19)			-0.044(0.043)	(-0.129,0.039)		
(24,8)	0.155(0.043)	(0.071,0.239)	0.157(0.045)	(0.066,0.245)	0.153(0.044)	(0.070,0.241)
(24,20)	0.188(0.043)	(0.103,0.272)	0.191(0.046)	(0.102,0.280)		
(27,11)	0.081(0.048)	(-0.015,0.174)	0.076(0.046)	(-0.018,0.162)		
(27,15)	-0.201(0.046)	(-0.288,-0.111)	-0.216(0.044)	(-0.306,-0.131)	-0.217(0.044)	(-0.299,-0.126)
(29,3)	-0.118(0.047)	(-0.207,-0.016)	-0.107(0.046)	(-0.209,-0.028)		
(29,5)					-0.090(0.045)	(-0.181,-0.005)
(29,9)					-0.185(0.050)	(-0.284,-0.090)
(29,10)	0.175(0.050)	(0.077,0.273)	0.121(0.046)	(0.031,0.210)		
(29,25)	0.210(0.047)	(0.114,0.296)	0.222(0.046)	(0.131,0.310)		
(30,2)			0.139(0.046)	(0.048,0.226)		
(30,8)			-0.076(0.042)	(-0.162,0.003)		
(30,17)					0.091(0.047)	(-0.004,0.179)

(30,26)			0.152(0.046)	(0.056,0.238)	0.125(0.048)	(0.026,0.214)
(31,19)					0.074(0.039)	(-0.002,0.149)
(33,10)	0.232(0.047)	(0.137,0.321)	0.178(0.046)	(0.084,0.264)		
(33,25)	0.161(0.045)	(0.073,0.246)	0.166(0.045)	(0.079,0.257)		
(33,29)	0.197(0.047)	(0.106,0.290)	0.192(0.049)	(0.087,0.281)		
(33,30)					-0.157(0.049)	(-0.250,-0.056)
(34,6)	0.160(0.050)	(0.066,0.259)	0.171(0.045)	(0.082,0.256)		
(34,19)	0.109(0.038)	(0.037,0.187)	0.118(0.038)	(0.044,0.193)		
(34,26)	0.137(0.045)	(0.047,0.223)	0.174(0.045)	(0.088,0.257)		
(36,4)	0.185(0.040)	(0.106,0.263)	0.173(0.043)	(0.089,0.259)		
(36,20)	0.122(0.040)	(0.042,0.196)				
(36,35)	0.220(0.046)	(0.123,0.305)	0.218(0.048)	(0.121,0.312)	0.217(0.050)	(0.122,0.317)
(37,2)	0.156(0.043)	(0.071,0.238)	0.191(0.043)	(0.107,0.272)		
(37,8)	-0.160(0.043)	(-0.244,-0.074)	-0.145(0.042)	(-0.224,-0.053)	-0.119(0.041)	(-0.196,-0.036)
(37,17)					0.047(0.046)	(-0.048,0.130)
(37,24)	-0.105(0.042)	(-0.188,-0.022)	-0.098(0.038)	(-0.172,-0.025)		
(37,26)	0.174(0.044)	(0.086,0.260)	0.236(0.044)	(0.150,0.319)	0.200(0.048)	(0.108,0.296)
(37,30)	0.250(0.045)	(0.160,0.334)	0.303(0.043)	(0.217,0.387)	0.313(0.045)	(0.222,0.401)
(37,33)					-0.075(0.050)	(-0.169,0.021)
(37,34)			0.088(0.040)	(0.008,0.167)		
(38,7)	0.444(0.037)	(0.367,0.513)	0.427(0.042)	(0.341,0.505)	0.478(0.035)	(0.406,0.542)
(38,11)	-0.179(0.046)	(-0.270,-0.091)	-0.209(0.045)	(-0.295,-0.120)	-0.227(0.044)	(-0.313,-0.140)
(38,13)	0.035(0.038)	(-0.043,0.107)	0.094(0.042)	(0.013,0.175)	0.081(0.042)	(0.001,0.165)
(38,15)	0.191(0.048)	(0.090,0.284)	0.208(0.045)	(0.121,0.301)	0.229(0.048)	(0.138,0.322)
(38,19)	0.589(0.031)	(0.528,0.647)	0.588(0.031)	(0.527,0.646)	0.608(0.028)	(0.547,0.656)
(38,23)	-0.100(0.041)	(-0.185,-0.025)	-0.147(0.046)	(-0.238,-0.055)		
(38,24)	-0.030(0.036)	(-0.102,0.038)	-0.041(0.034)	(-0.108,0.024)		
(38,27)			-0.142(0.039)	(-0.215,-0.064)		
(38,31)					0.071(0.040)	(-0.008,0.150)
(38,34)	0.073(0.039)	(-0.005,0.145)	0.079(0.040)	(0.001,0.158)		
(39,26)	0.195(0.045)	(0.101,0.276)	0.238(0.047)	(0.145,0.328)	0.217(0.047)	(0.120,0.305)
(39,30)	0.186(0.047)	(0.091,0.276)	0.201(0.046)	(0.107,0.289)	0.215(0.045)	(0.122,0.296)
(39,33)					-0.189(0.048)	(-0.278,-0.091)
(39,34)	0.122(0.044)	(0.041,0.211)	0.148(0.042)	(0.064,0.230)		
(39,37)	0.264(0.044)	(0.171,0.345)	0.288(0.042)	(0.202,0.366)	0.298(0.043)	(0.213,0.382)

表 6.4 结果显示，BCLCFA 有 54 项(7.29%)显著残差协方差、BCENCFA 有 58 项(7.83%)显著残差协方差、BACENCFA 有 40 项(5.40%)显著残差协方差。而根据 Pan 等(2017)研究，若显著残差协方差项小于残差协方差项总数的 10%，则认为协方差矩阵足够稀疏，即模型结果是合适的。同时，若残差协方差项大于 10%或某个残差协方差项的绝对值大于 0.5，则认为需要调整整个 CFA 方案。在表 6.4 的估计结果中， $\psi_{38,19}$ 在 3 个方法的估计结果都是大于 0.5。因此，对于 KIMS 量表的结构可能需要进行一些修整。

6.3 讨论

本章通过开源网站的实例数据对幽默风格问卷和肯塔基正念技能清单进行验证性因子分析，发现这两个量表均存在交叉载荷和残差协方差。同时，对比了 BCLCFA、BCENCFA 和 BACENCFA 在实际数据研究中的表现。

对比 3 种方法的拟合结果，3 种方法均没有改变原量表的结构。其中，BCLCFA 和 BCENCFA 的拟合结果较好，模型收敛速度更快，添加的交叉载荷少。但显著残差协方差项较多，模型整体结构更加复杂，可能会对后续分析造成困难。而 BACENCFA 方法得到的模型结构更加清晰，残差协方差矩阵更加稀疏，但模型拟合结果不如前两种方法，且模型收敛速度较慢，可能需要增加样本量或修改先验信息进行改进。

7 结论与展望

本文旨在进一步提升贝叶斯框架下 CFA 的模型选择与参数估计的稳健性与灵活性。研究主要由三部分组成，第一部分是在贝叶斯图 Lasso 模型(Wang^[39], 2012)的基础上，通过对模型精度矩阵的非对角元素施加弹网先验，提出 BCEN 模型，通过分块 Gibbs 采样方法实现模型的参数估计，并进一步的推广至 BACEN 模型；第二部分是在 BCLCFA 方法(Pan 等^[32], 2017)的基础上将 BCEN 和 BACEN 模型与 CFA 结合，提出了 BCENCFA 和 BACENCFA 方法，这 2 种方法同样能够在不改变原有因子结构的基础上通过残差协方差实现 CFA 模型的模型调节，找出被忽视的模型结构；第三部分是通过数值模拟和实证分析比较所提出的 2 种 CFA 方法和 BCLCFA 方法进行比较。根据数值模拟和实证分析结果可知，当固定因子载荷，仅针对残差协方差矩阵进行模型选择与参数估计时，BACENCFA 方法是一个更好的选择；当选择同时放松交叉载荷和残差协方差时，BACENCFA 虽然得到了更少的残差协方差，但也存在更多的显著的交叉载荷项，因此在实际问题的研究中还需要根据研究需求来对方法进行选择。

本文所提出的方法仅仅是对贝叶斯 CFA 的初步研究。研究过程中依旧存在不少未曾讨论的方向。包括：研究仅仅针对连续数据进行研究，并未考虑不同类型数据下 BCENCFA 方法的解决方案；许多在回归分析背景下得到检验的先验理论能够通过分块 Gibbs 采样方法实现对协方差矩阵的参数估计和模型选择(Wang^[39], 2012)，这为 CFA 方法的研究提供了更多的选择；对于 CFA 中交叉载荷的估计仅使用了 Chen 等^[4](2020)提出的通过 Lasso 方法进行参数估计与变量选择，而改变估计交叉载荷时的先验，进一步提升 CFA 模型灵活性也是一个值得研究的问题。

参考文献

- [1] Andrews D F, Mallows C L. Scale mixtures of normal distributions[J]. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1974, 36(1): 99-102.
- [2] Baer R A, Smith G T, Allen K B. Assessment of mindfulness by self-report: The Kentucky Inventory of Mindfulness Skills[J]. *Assessment*, 2004, 11(3): 191-206.
- [3] Bernardini D, Paterlini S, Taufer E. New estimation approaches for graphical models with elastic net penalty[J]. *arXiv preprint arXiv:2102.01053*, 2021.
- [4] Chen J, Guo Z, Zhang L, et al. A partially confirmatory approach to scale development with the Bayesian Lasso[J]. *Psychological Methods*, 2020, 26(2), 210-235.
- [5] Chen M H, Shao Q M. Monte Carlo estimation of Bayesian credible and HPD intervals[J]. *Journal of Computational and Graphical Statistics*, 1999, 8(1): 69-92.
- [6] Edwards D. *Introduction to graphical modelling*[M]. Springer Science & Business Media, 2012.
- [7] Fan J, Feng Y, Wu Y. Network exploration via the adaptive LASSO and SCAD penalties[J]. *The Annals of Applied Statistics*, 2009, 3(2): 521-541.
- [8] Friedman J, Hastie T, Tibshirani R. Sparse inverse covariance estimation with the graphical lasso[J]. *Biostatistics*, 2008, 9(3): 432-441.
- [9] Gelman A, Meng X L, Stern H. Posterior predictive assessment of model fitness via realized discrepancies[J]. *Statistica Sinica*, 1996: 733-760.
- [10] Gelman A, Rubin D B. Inference from iterative simulation using multiple sequences[J]. *Statistical Science*, 1992, 7(4): 457-472.
- [11] Giudici P, Stanghellini E. Bayesian inference for graphical factor analysis models[J]. *Psychometrika*, 2001, 66(4): 577-591.
- [12] Hans C. Bayesian lasso regression[J]. *Biometrika*, 2009, 96(4): 835-845.
- [13] Hoofs H, van de Schoot R, Jansen N W H, et al. Evaluating model fit in Bayesian confirmatory factor analysis with large samples: Simulation study introducing the BRMSEA[J]. *Educational and Psychological Measurement*, 2018, 78(4): 537-568.
- [14] Hsiang T C. A Bayesian view on ridge regression[J]. *Journal of the Royal Statistical Society: Series D (The Statistician)*, 1975, 24(4): 267-268.
- [15] Jennrich R I, Sampson P F. Rotation for simple loadings[J]. *Psychometrika*, 1966, 31(3): 313-323.

- [16]Jöreskog K G. A general approach to confirmatory maximum likelihood factor analysis[J]. *Psychometrika*, 1969, 34(2): 183-202.
- [17]Jorgensen T D, Garnier-Villarreal M, Pornprasertmanit S, et al. Small-variance priors can prevent detecting important misspecifications in Bayesian confirmatory factor analysis[C]//The Annual Meeting of the Psychometric Society. Springer, Cham, 2017: 255-263.
- [18]Kaplan D. Evaluating and modifying covariance structure models: A review and recommendation[J]. *Multivariate Behavioral Research*, 1990, 25(2): 137-155.
- [19]Kovács S, Ruckstuhl T, Obrist H, et al. Graphical Elastic Net and Target Matrices: Fast Algorithms and Software for Sparse Precision Matrix Estimation[J]. arXiv preprint arXiv:2101.02148, 2021.
- [20]Laurent S, Rombouts J V K, Violante F. On the forecasting accuracy of multivariate GARCH models[J]. *Journal of Applied Econometrics*, 2012, 27(6): 934-955.
- [21]Lauritzen S L. *Graphical models*[M]. Clarendon Press, 1996.
- [22]Lee S Y. A Bayesian approach to confirmatory factor analysis[J]. *Psychometrika*, 1981, 46(2): 153-160.
- [23]Lee S Y. *Structural equation modeling: A Bayesian approach*[M]. John Wiley & Sons, 2007.
- [24]Li Q, Lin N. The Bayesian elastic net[J]. *Bayesian Analysis*, 2010, 5(1): 151-170.
- [25]Liu B, Jing L, Yu J, et al. Robust graph learning via constrained elastic-net regularization[J]. *Neurocomputing*, 2016, 171: 299-312.
- [26]Lu Z H, Chow S M, Loken E. Bayesian factor analysis as a variable-selection problem: Alternative priors and consequences[J]. *Multivariate Behavioral Research*, 2016, 51(4): 519-539.
- [27]MacCallum R C. Model specification: Procedures, strategies, and related issues[J]. 1995, 16-36.
- [28]Marsh H W, Hau K T, Balla J R, et al. Is more ever too much? The number of indicators per factor in confirmatory factor analysis[J]. *Multivariate Behavioral Research*, 1998, 33(2): 181-220.
- [29]Martin R A, Puhlik-Doris P, Larsen G, et al. Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire[J]. *Journal of Research in Personality*, 2003, 37(1): 48-75.
- [30]Muthén B, Asparouhov T. Bayesian structural equation modeling: a more flexible

- representation of substantive theory[J]. *Psychological Methods*, 2012, 17(3): 313-335.
- [31] O'Hara R B, Sillanpää M J. A review of Bayesian variable selection methods: what, how and which[J]. *Bayesian Analysis*, 2009, 4(1): 85-117.
- [32] Pan J, Ip E H, Dubé L. An alternative to post hoc model modification in confirmatory factor analysis: The Bayesian lasso[J]. *Psychological Methods*, 2017, 22(4): 687-704.
- [33] Park T, Casella G. The bayesian lasso[J]. *Journal of the American Statistical Association*, 2008, 103(482): 681-686.
- [34] Ryali S, Chen T, Supekar K, et al. Estimation of functional connectivity in fMRI data using stability selection-based sparse partial correlation with elastic net penalty[J]. *NeuroImage*, 2012, 59(4): 3852-3861.
- [35] Scheines R, Hoijtink H, Boomsma A. Bayesian estimation and testing of structural equation models[J]. *Psychometrika*, 1999, 64(1): 37-52.
- [36] Schmitt T A. Current methodological considerations in exploratory and confirmatory factor analysis[J]. *Journal of Psychoeducational Assessment*, 2011, 29(4): 304-321.
- [37] Sörbom D. Model modification[J]. *Psychometrika*, 1989, 54(3): 371-384.
- [38] Tibshirani R. Regression shrinkage and selection via the lasso[J]. *Journal of the Royal Statistical Society: Series B (Methodological)*, 1996, 58(1): 267-288.
- [39] Wang H. Bayesian graphical lasso models and efficient posterior computation[J]. *Bayesian Analysis*, 2012, 7(4): 867-886.
- [40] Yuan M, Lin Y. Model selection and estimation in the Gaussian graphical model[J]. *Biometrika*, 2007, 94(1): 19-35.
- [41] Zou H. The adaptive lasso and its oracle properties[J]. *Journal of the American Statistical Association*, 2006, 101(476): 1418-1429.
- [42] Zou H, Hastie T. Regularization and variable selection via the elastic net[J]. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 2005, 67(2): 301-320.
- [43] Zou H, Zhang H H. On the adaptive elastic-net with a diverging number of parameters[J]. *Annals of Statistics*, 2009, 37(4): 1733-1751.

附录

附表 1 幽默风格问卷

内容		
1	I usually don't laugh or joke around much with other people.*	F1
2	If I am feeling depressed, I can usually cheer myself up with humor.	F2
3	If someone makes a mistake, I will often tease them about it.	F3
4	I let people laugh at me or make fun at my expense more than I should.	F4
5	I don't have to work very hard at making other people laugh—I seem to be a naturally humorous person.	F1
6	Even when I'm by myself, I'm often amused by the absurdities of life.	F2
7	People are never offended or hurt by my sense of humor.*	F3
8	I will often get carried away in putting myself down if it makes my family or friends laugh.	F4
9	I rarely make other people laugh by telling funny stories about myself.*	F1
10	If I am feeling upset or unhappy I usually try to think of something funny about the situation to make myself feel better.	F2
11	When telling jokes or saying funny things, I am usually not very concerned about how other people are taking it.	F3
12	I often try to make people like or accept me more by saying something funny about my own weaknesses, blunders, or faults.	F4
13	I laugh and joke a lot with my closest friends.	F1
14	My humorous outlook on life keeps me from getting overly upset or depressed about things.	F2
15	I do not like it when people use humor as a way of criticizing or putting someone down.*	F3
16	I don't often say funny things to put myself down.*	F4
17	I usually don't like to tell jokes or amuse people.*	F1
18	If I'm by myself and I'm feeling unhappy, I make an effort to think of something funny to cheer myself up.	F2
19	Sometimes I think of something that is so funny that I can't stop myself from saying it, even if it is not appropriate for the situation.	F3
20	I often go overboard in putting myself down when I am making jokes or trying to be funny.	F4
21	I enjoy making people laugh.	F1
22	If I am feeling sad or upset, I usually lose my sense of humor.*	F2
23	I never participate in laughing at others even if all my friends are doing it.*	F3
24	When I am with friends or family, I often seem to be the one that other people make fun of or joke about.	F4
25	I don't often joke around with my friends.*	F1
26	It is my experience that thinking about some amusing aspect of a situation is often a very effective way of coping with problems.	F2
27	If I don't like someone, I often use humor or teasing to put them down.	F3
28	If I am having problems or feeling unhappy, I often cover it up by joking around, so that even my closest friends don't know how I really feel.	F4
29	I usually can't think of witty things to say when I'm with other people.*	F1
30	I don't need to be with other people to feel amused—I can usually find things to laugh about even when I'm by myself.	F2
31	Even if something is really funny to me, I will not laugh or joke about it if someone will be offended.*	F3
32	Letting others laugh at me is my way of keeping my friends and family in good spirits.	F4

注：“F1”为 Affiliative humor；“F2”为 Self-enhancing humor；“F3”为 Aggressive humor；“F4”为 Self-defeating humor。“*”表示需要反向计算得分。

附表 2 肯塔基正念技能清单

内容		
1	I notice changes in my body, such as whether my breathing slows down or speeds up.	F1
2	I'm good at finding the words to describe my feelings.	F2
3	When I do things, my mind wanders off and I'm easily distracted.*	F3
4	I criticize myself for having irrational or inappropriate emotions.*	F4
5	I pay attention to whether my muscles are tense or relaxed.	F1
6	I can easily put my beliefs, opinions, and expectations into words.	F2
7	When I'm doing something, I'm only focused on what I'm doing, nothing else.	F3
8	I tend to evaluate whether my perceptions are right or wrong.*	F4
9	When I'm walking, I deliberately notice the sensations of my body moving.	F1
10	I'm good at thinking of words to express my perceptions, such as how things taste, smell, or sound.	F2
11	I drive on "automatic pilot" without paying attention to what I'm doing.*	F3
12	I tell myself that I shouldn't be feeling the way I'm feeling.*	F4
13	When I take a shower or bath, I stay alert to the sensations of water on my body.	F1
14	It's hard for me to find the words to describe what I'm thinking.*	F2
15	When I'm reading, I focus all my attention on what I'm reading.	F3
16	I believe some of my thoughts are abnormal or bad and I shouldn't think that way.*	F4
17	I notice how foods and drinks affect my thoughts, bodily sensations, and emotions.	F1
18	I have trouble thinking of the right words to express how I feel about things.*	F2
19	When I do things, I get totally wrapped up in them and don't think about anything else.	F3
20	I make judgments about whether my thoughts are good or bad.*	F4
21	I pay attention to sensations, such as the wind in my hair or sun on my face.	F1
22	When I have a sensation in my body, it's difficult for me to describe it because I can't find the right words.*	F2
23	I don't pay attention to what I'm doing because I'm daydreaming, worrying, or otherwise distracted.*	F3
24	I tend to make judgments about how worthwhile or worthless my experiences are.*	F4
25	I pay attention to sounds, such as clocks ticking, birds chirping, or cars passing.	F1
26	Even when I'm feeling terribly upset, I can find a way to put it into words.	F2
27	When I'm doing chores, such as cleaning or laundry, I tend to daydream or think of other things.*	F3
28	I tell myself that I shouldn't be thinking the way I'm thinking.*	F4
29	I notice the smells and aromas of things.	F1
30	I intentionally stay aware of my feelings.	F1
31	I tend to do several things at once rather than focusing on one thing at a time.*	F3
32	I think some of my emotions are bad or inappropriate and I shouldn't feel them.*	F4
33	I notice visual elements in art or nature, such as colors, shapes, textures, or patterns of light and shadow.	F1
34	My natural tendency is to put my experiences into words.	F2
35	When I'm working on something, part of my mind is occupied with other topics, such as what I'll be doing later, or things I'd rather be doing.*	F3
36	I disapprove of myself when I have irrational ideas.*	F4
37	I pay attention to how my emotions affect my thoughts and behavior.	F1
38	I get completely absorbed in what I'm doing, so that all my attention is focused on it.	F3
39	I notice when my moods begin to change.	F1

注：“F1”为 Observe items；“F2”为 Describe items；“F3”为 Act With Awareness items；“F4”为 Accept Without Judgment items。“*”表示需要反向计算得分。

后 记

时光匆匆，转眼已是三年时光。随着这篇论文的完成，三年的研究生学习也将画上句号。回顾这三年的点点滴滴，也算是过的充实。三年的求学历程，有欢乐，有艰辛，感谢这三年来给予我支持和帮助的老师、同学们。

首先，感谢我的导师，对我这三年的学习和生活提供指导和帮助，让我有了明确的目标，也让我更加坚定的在学习的道路上不断前进。

其次，感谢学校的各位老师。通过参加老师们所设的讨论班，让我能够在课程学习之余，学习了更多的新知识。也让我对一些新领域有了更加深入的了解。

再者，感谢我的同门，在日常学习中给我提供了必要的协助。同时要感谢的还有我的同学们，感谢他们让我的学习生活不是那么的枯燥乏味。

最后，感谢我的父母，他们的支持和认可是我不断追求与学习的动力。三年的学习，虽然过程不是一帆风顺，但一路走来，也算是有所收获。