

# 高维贝叶斯向量自回归模型对中国宏观经济变量预测能力的比较研究

李玥 侯成瀚

(湖南大学经济管理研究中心, 湖南省长沙市, 410000)

**摘要:** 随着中国宏观经济调控体系的逐步建立健全, 宏观经济预测的相关研究已能够作为国家制定经济战略、调控市场时的重要参考。近些年大数据的出现使得研究人员能应用更多的信息, 并且将经济活动视为随时间变化的过程来进行实时预测, 这在改善预测结果的同时也给预测工作带来了挑战。为了在这种日益复杂的条件下提高我国宏观经济预测的精确度、探究能够提高精确度的主要因素, 本文通过理论和实证分析, 采用误差协方差中具有特定Kronecker结构的高维贝叶斯向量自回归(BVAR)模型, 对我国20个主要宏观经济金融变量进行了拟合和预测研究。文中采用的BVAR模型一方面能够通过Kronecker结构拆分协方差矩阵, 从而加快大数据下运算高维模型中大量待估参数的速度; 另一方面能假定回归方程的方差是时变的, 并且允许误差协方差中单独或同时存在非高斯、共同随机波动和序列相关的创新结构。在大数据下对我国宏观经济数据的拟合及预测结果表明: 拥有三种创新协方差结构的模型都能在一定程度上提高中国宏观经济预测的精确度, 即令协方差具有时变因子能更好地捕捉中国宏观数据的性质。综合来看, 模型协方差结构中同时具有t分布以及共同随机波动分布是提高中国宏观经济预测精确度的主要因素, 这在点预测和密度预测上都能提供较大改进。同时, 相较于小型模型, 含有更多变量信息的高维BVAR能更好地考虑变量间的联合信息, 从而改善中国宏观经济预测的效果。目前通过具有不同时变误差协方差结构的高维BVAR对中国宏观经济进行系统性预测和研究的文献较少, 因此本文填补了这方面的空白, 有助于在大数据环境下提高我国宏观预测精确度, 为日后研究我国宏观经济趋势时的模型选择提供了新思路。

**关键词:** 中国宏观经济预测; 高维BVAR; 重尾分布; 共同随机波动; 序列相关

**中图分类号:** F064.1

**文献标识码:** A

## 引言

宏观预测可以对由经济和金融系统发展引起的可变行为的变化作出预期判断, 从而预测未来的经济趋势。同时随着科学技术和信息技术的不断发展, 政府和企业的统计数据量和公开数据量不断增加, 大数据进入了人们的视野。Bań bura 等(2010)研究发现将大量数据集放入模型可以改善预测结果<sup>[1]</sup>。因为研究人员能够集合大数据中的多种信息, 并且可以将经济活动视为变化和发展过程进行研究和预测。这从很大程度上可以改善我们对宏观经济的预测和理解。但这种难以使用传统方法处理的大规模数据集带来了许多处理高维数据的实际问题<sup>[2][3]</sup>。例如, 模型中有许多潜在的解释变量, 但是我们无法区分这些变量中哪些是重要的, 或者数据中没有足够信息来对参数进行准确的估计。另一方面, 一旦当我们将经济活动视为变化的过程, 模型中参数则应随时间变化, 并且误差也应包含随时间的变化。

为了克服这些问题, 贝叶斯分析方法进入了人们的视野。贝叶斯方法能够使用大量变量从而减少遗漏变量偏差, 同时充分结合整体信息、经验信息和样本信息等已知条件, 基于后验分布进行统计推测和判断<sup>[4]</sup>。贝叶斯计量的一个主要应用领域是用于估计高维模型, 相比于其他方法, 其不仅可以减少由于样本量较小造成的统计误差, 而且可以将变量的统计特性用作参数的先验信息, 以此克服过度参数化的缺点。高维贝叶斯向量自回归(Bayesian Vector Autoregression 简记为BVAR)基于贝叶斯方法, 把经济活动看做变化的过程, 将模型中各类参数等设置为随时间波动的变量, 通过设定不同的先验分布克服了在经典统计建模中确定参数值的困难。尤其是对于观测变量行为在经济和金融体系发展进程中的波动, 贝叶斯分析方法被当成了一种有效的分析途径<sup>[5]</sup>。

在估计高维模型时, 计算贝叶斯模型中的大量待估参数会耗费较多时间, 因此在处理大数据方面具有一定困难。基于上述考虑, 本文中介绍的高维贝叶斯向量自回归模型的误差协方差中引入了Kronecker结构, 该结构能够缩减协方差矩阵大小, 从而极大加快运算速度,

解决在处理大数据时的困难。为了提高宏观预测结果的精确度,解决高维模型过度参数化的弊端,我们分别介绍了模型中不同先验分布的原理和推导,同时引入各类误差协方差结构并对参数设置进行总结说明。我们通过中国的宏观经济数据对这些不同假设不同结构的模型在经济预测中进行比较,判断大数据量是否能够改善模型的预测结果以及哪些因素能更好地提高中国宏观经济预测的精确度。

## 1. 文献综述

### 1.1 向量自回归模型的发展

Sims 在 1980 年的文章中提出了向量自回归模型(Vector Autoregression 简记为 VAR)。他将 VAR 从一个单纯的统计学模型扩展成一种一般架构,该架构可以通过经济理论模型来辨别变量之间的动态变化关系。自此,一套基于联立方程组的区别于传统的宏观计量手段创立形成。VAR 模型把研究问题中的所有变量都视为内生变量,通过每个内生变量与其滞后值的关系来构建模型,这样便可以将单一变量自回归模型拓展到由多元时间序列变量组合而成的向量自回归模型。

最初的实际应用中,VAR 的大小通常在大约 3 到 10 个变量之间,这种小型模型可能会造成变量的遗漏偏差,从而对结构分析和预测产生不利影响<sup>[6]</sup>。例如 Christiano 等人(1997)在其研究中指出货币紧缩带来的价格变动是由于省略了诸如商品价格指数之类的前瞻性变量而产生的假象<sup>[7]</sup>。在对于研究多个经济体、分类信息等更大型的变量集的应用时,若仍对模型大小做限制是不恰当的。

由于数据量和计算量的迅速增加,如何对多变量模型进行更快、更准确地运算已成为亟待解决的问题。Doan 等人(1984)的研究首先提出,相对于单变量方程,基于贝叶斯方法的标准向量自回归估计可以更好地对高维模型进行计算<sup>[8]</sup>。通过应用贝叶斯收缩能够处理大型 VAR,同时组合先验信息并提高预测效率与性能,这能够解决传统计量模型中由于没有充足的信息来支撑参数高精度估计的问题。随后,Koop 等人(2009)使用了三个宏观变量来显示货币政策冲击的影响,并提出因子增强的向量自回归模型和含有随时间变化的参数扩展模型,同时呈现了贝叶斯推理怎样在这些模型中实行<sup>[9]</sup>。在最近的开创性工作中 Banbura 等人(2010)提出了一种具有超过 20 个因变量的大型高维标准贝叶斯向量自回归模型,并证明了其性能较小型贝叶斯系统更胜一筹。文章中对货币冲击影响的分析表明,基于 20 个变量的向量自回归产生的结果在进一步扩大模型规模时仍保持稳健,其结果能产生可靠的脉冲响应,但是使用包含 20 个关键宏观经济指标的系统已经获得了这种性能。

考虑到结合大量数据有助于进行更准确的预测,在近几年的研究中,为了解决高维贝叶斯模型中处理大量待估参数的问题,研究者们开始尝试通过给高维贝叶斯向量自回归模型赋予不同的先验,同时改变模型中误差部分的协方差结构来提高模型的预测效率和准确性。

### 1.2 宏观数据预测

传统计量经济学中对宏观经济的预测主要借助以宏观经济理论为基础构造的理论模型,利用统计样本进行回归从而估计内含参数,借此分析各个变量的数量相关联系,最后达到对研究变量进行预测的目的。自 VAR 模型被提出以后,使用 VAR 进行宏观预测成为普遍的研究方法<sup>[10][11]</sup>。在预测宏观趋势并进行宏观经济数据的结构分析时,经济学家们广泛使用了向量自回归等多元时间序列模型。使用此类模型进行策略分析时,它们不会着重强调变量间的经济理论逻辑,不会自动产生因果关联解释,但会提供一组通过动态统计得到的各个经济变量存在彼此依存关系的证据,这在一定程度上有助于评估因果假设。

在宏观经济指标方面,一些经济学家使用 BVAR 模型结合大数据预测了 GDP、CPI、通胀和利率等指标。比如 Christine 等人(2006)通过对 BVAR 施加不同先验,应用大面板时间序列数据对美国的 IP 和 CPI 进行预测,证明在广泛先验选择下 BVAR 的预测结果表现出色<sup>[12]</sup>;Carriero 等人(2016)在预测时使用了美国经济中 14 个变量,结果证明 BVAR 模型能极大提高运算速度,并可以通过利用大数据集中包含的信息得到更好的密度预测结果<sup>[13]</sup>。如

Ramos (2003)、Berg 等 (2015) 还使用了 BVAR 模型来预测股票、债券、各类贸易等金融市场活动, 同无限制 VAR 及单变量模型产生的基准预测结果相比, 由 BVAR 得出的样本外预测结果显著较好, 即这些研究结果指出 BVAR 的直接预估效果表现明显胜于其他应用模型<sup>[14~16]</sup>。

### 1.3 大数据下的模型估计

大型数据集的出现使得宏观经济学模型中的待估参数数量激增, 同时有些参数需要随时间发生改变, 这为模型的估计带来巨大挑战。以下将从三个方面对大数据下 VAR 模型估计存在的问题及解决方案进行说明。

第一个是过度参数化问题。VAR 中通常需包含尽可能多的变量, 但这也会带来过度参数化问题。尤其在多元回归中, 过度参数化问题被放大, 此时每个 VAR 方程都包含所有因变量的滞后和外生变量, 并且对于月度数据, 滞后长度还通常较长。在大型模型下则可以通过贝叶斯先验信息克服该问题。在处理大数据时通过先验收缩将系数缩小到合理的值 (通常为零), 即反映了许多潜在的解释变量不重要的信念, 以减轻过拟合现象, 这能够用于解决由于数据信息不足带来的麻烦<sup>[17][18]</sup>。使用明尼苏达先验是最常用的解决方法之一, 该方法能够收缩模型系数矩阵, 反映出经济学家对变量滞后的想法<sup>[19][20]</sup>。而使用自然共轭先验的贝叶斯 VAR, 无论样本量有多大, 其协方差矩阵总可得到正定矩阵。随机搜索变量选择先验 (Stochastic Search Variable Selection 简记为 SSVS) 也是一种较为有效的先验信息。

第二个问题则是许多模型具有时变协方差。传统向量自回归模型参数的非时变性和同方差假设影响了模型在经济学实证预测中的精确度。多数时候, 令误差协方差矩阵随时间变化更为重要<sup>[21]</sup>。许多证据表明, 大部分宏观经济变量的性质会随时间变化, 例如一个经济体在不同时期宏观经济波动性有所不同, 这意味着模型中应该有时变误差。说明现有的传统 VAR 模型在估计时具有一定局限性, 并且显示了在构建模型时设定协方差参数随时间变化的合理性。下面我们对三种误差协方差分布假设进行初步说明, 许多文献已经证明拥有这些结构的 BVAR 能够比标准 BVAR 更好地预测宏观经济。

目前对 BVAR 模型误差协方差的扩展研究主要存在下面三种。第一种是误差协方差的边际似然函数为多元 t 分布的 BVAR 模型。这类模型对自由度参数的先验分布更敏感, 增加了预测极端事件的可能性, 并且可以帮助改善预测性能<sup>[22]</sup>。第二种扩展为协方差中具有共同随机波动结构 (BVAR-CSV)。这类模型假定波动率随时间游走不定, 故而可以减少误差。它可以被应用于大数据集的预测, 提高预测性能。第三种是具有移动平均结构的 BVAR 模型 (BVAR-MA)。这类模型假设各回归式的方差是序列相关的, 因此能够进行更精确地预测并提高分析性能。例如, Athanasopoulos 等人 (2016) 创建了一个结构化的 VAR-MA 模型并通过该途径分析研究了加拿大施行的货币政策, 结果表明这类模型能够有效解决尤其是在小型开放经济体中实证货币文献内常见的异常现象, 包括价格、产量和汇率的难题<sup>[23]</sup>。

还有一个问题是, 在大数据下模型中的变量即便较少也会产生许多参数, 若此时估计传统 VAR 矩阵将会带来巨大计算量, 这会大大降低计算效率。因此估计高维模型时, 我们可以通过在误差协方差矩阵结构中添加 Kronecker 结构, 将矩阵拆分, 从而减少计算操作, 大大提高计算效率。

### 1.4 我国宏观数据预测研究

国外已有的研究证明, 高维贝叶斯向量自回归模型 (BVAR) 在实证中应用广泛且能够有效地研究各个方面的宏观经济现象, 例如宏观或稍小地域范围的经济周期、国家或地方政策对经济造成的影响以及大规模的经济震荡。同时可以应用于宏观预测, 为政府提供更多决策参考。从国际研究的进展来看, 尤其是应用美国数据进行的研究, 关于动态因素模型的理论 and 实证研究已经相当深入。不过中国在这方面的相关探索仍缺乏系统性讨论。

国内较早便开始使用线性和非线性模型对宏观经济预测方面进行了研究。李仲莲等人 (1990) 使用受控自回归模型, 选择 1971 年至 1987 年的统计数据作为初始样本, 并预测未来几年的货币流通, 结果的最大误差为 11%<sup>[24]</sup>。目前国内在使用 BVAR 做预测方面的文献较

少,使用高维 BVAR 进行研究的文献则更少。赵昕东等人(2002)通过 VAR 模型的脉冲响应函数同时利用货币政策工具变量建立了货币冲击模型,该模型将一个标准差大小的冲击添加进 VAR 的干扰项中,因此这种冲击可以经由 VAR 模型相互作用的动态构造对其余全部变量产生影响<sup>[25]</sup>。该文章分析和预测了货币供给量、贷款及实际利率对我国 GDP 的冲击反应,最终得到的响应较显著。张劲帆等人(2018)依据贝叶斯方法构建了混频向量自回归模型,通过中国宏观经济指标对模型进行了验证说明。结果表明模型在允许多变量、不同频率数据共存条件下提高了模型估计的自由度,从而显著改善了预测精度<sup>[26]</sup>。

当下,中国的经济学家也开始逐步应用大数据做预测。张崇等(2012)收集到一些关键词在网络中的搜索量作为研究样本,以探究它们同 CPI 是否存在一定关联。他们通过搜索量构建了具有一定转折点预测能力的搜索指数模型,结果能够得到先于国家统计局一个月左右的预测数据<sup>[27]</sup>。张伟等人(2020)利用动态因子模型构造预测指数,使用混频抽样数据模型结合对应大数据技术对中国宏观经济总量进行实时预测检测,结果表明大数据能大幅度提升模型预测的精确度<sup>[28]</sup>。范辉等人(2011)用我国 1985 至 2008 这 24 年内的宏观数据证明了对于国内 GDP 和 CPI 增长速率的预测而言,高维贝叶斯向量自回归和向量自回归移动平均模型获得的预测结果都大大优于标准向量自回归下获取的结果<sup>[29]</sup>。

目前,传统低维、固定方差的贝叶斯自回归模型难以在当前数据量激增的背景下满足研究要求。但是中国应用高维贝叶斯向量自回归模型做预测方面的研究尚不深入,多是一些用基本模型对中国部分产业或地区进行的研究,几乎没有人系统性地研究过在大量数据集下运用高维 BVAR 估计存在的问题以及应对方法,并且对它们的预测结果、尤其是密度预测方面进行比较。运用多种具有时变协方差结构的高维 BVAR 模型对中国宏观经济同时进行点预测和密度预测的文献更是少之又少。因此本文的研究将填补一部分国内应用高维贝叶斯向量自回归做宏观预测的空白,为日后经济学家进行宏观预测时的模型选择提供较为全面的参考,从而全方位地有效提高最终我国宏观预测的精确程度。

## 2. 数据描述

如前所述,已有文献表明拥有 20 个关键宏观经济指标的 BVAR 模型已然能较好地进行预测,因此本文中我们参考了 Banbura 等(2010)和 Chan(2018)文献中选择的数据集,选取了中国相同或同类型的 20 个变量。不同的是上述文章中所采用的都是季度数据,而本文中均为月度数据。所有数据都是按照时间顺序收集得到的,数据来源包括中国国家统计局、EPS DATA 平台、CSMAR 中国经济金融研究数据库及 Wind 统计机构。月度时间序列数据的起止范围为 2005 年 1 月到 2020 年 11 月。数据类别如表 2.1 所示。

在正式描述之前,最开始要先处理一些仅具有累计值的数据。先把这些累计值数据转变成每月当期值,接着再同其他数据一起进行描述性统计分析。表 2.1 集中展现了各个变量的数据样本在还未进行季节调整前的汇总统计信息。最终计算时,我们需要统一每个变量的起止时间,因此每个变量的观察个数应该相同。

表 2.1 变量描述性统计(季节调整前)

变量	数量	均值	标准差	最小值	最大值
工业生产指数	191	10.48	4.69	-1.1	20.9
消费者物价指数	191	102.62	1.90	98.2	108.7
国家财政收入(亿元)	191	9821.10	4836.11	2056.6	23621.1
社会消费品零售总额(亿元)	191	19130.89	10023.23	4663.3	39514.2
原油产量(万吨)	191	1634.19	98.12	1381.4	1832.3
制造业采购经理指数	191	51.65	2.84	35.7	59.2
商品住宅施工面积(万平方米)	191	13231.80	4523.78	4458.51	29256.4
商品房销售额(亿元)	191	7429.79	4897.81	582.39	20719.3

					4
进出口总额（千美元）	191	2.84e+08	9.25e+07	8.42e+07	4.61e+08
实际利用外商直接投资金额（百万美元）	191	8948.70	2896.74	3874	20225
上证指数	191	2730.27	840.61	1060.74	5954.77
汇率（人民币/100美元）	191	684.38	60.27	610.43	827.65
一年期国债收益率	191	2.58	0.41	1	4.12
十年期国债收益率	191	3.55	0.48	2.51	4.67
外汇储备（亿美元）	191	26625.16	9602.99	6236.46	39932.13
消费者信心指数	191	109.41	7.53	97	126.6
货币和准货币（M2）供应量（亿元）	191	1045568	581391.10	257708.50	2172000
债券市场月成交总金额（元）	191	8.04e+12	8.79e+12	7.52e+10	2.72e+13
宏观经济景气指数	191	182370.30	169209.80	3287.03	635636
工业生产者出厂价格指数	191	101.59	2.37	94.2	106.89

尽管同一个变量在同一年内的数值有较大变动，但绝大部分数据在每年内的变动趋势很相似，会在某几个月份偏高，或者在某几个月份中偏低。这是由于部分宏观经济变量受到年内季节变动的影响，随时间的推进产生周期性相似的变化，当这种影响过大时可能会令人难以辨别出时间序列短期自身的基本变动趋势。因此如果想要掌握整体经济运转的变化，我们务必要对数据进行季节性调整来消除这些影响因素，以便获得更平稳的数据序列。

我们选择用 Eviews 软件中提供的、也是目前最常用的 X-12 方法对所有数据进行季节调整。该方法中包括了对交易日、节假日影响的调节以及对各种极端值的处理。经过季节调整后的数据能更好地去除季节趋势，变为更加平稳的序列，从而增强数据的可比性，便于识别数据变化从而进行观察评估。因为对时间序列数据取对数以后易于获取比较平稳的序列，并且可以消除异方差，因此我们把所有季节调整后的结果记录下来后，计算每个数据的自然对数值并得出前后两个月之间的差值，记作  $\Delta \log x = \log x_t - \log x_{t-1}$ 。鉴于差值过小，可以将所有差值扩大 100 倍。对于某些同比增长指数数据中含有负数（如工业生产指数），我们在其原来的值上加 100，再进行相同处理。最终我们获得了所需 20 个变量的增长率数据。

我们主要希望预测的是中国的工业生产指数（IP）增长率和消费者物价指数（CPI）。通常来说，标志性的宏观经济变量有 GDP 和通货膨胀率，但在我国这两个数据的统计频率没有月度统计。很多文献都发现，收集更多类型的数据能够利用更大量的数据一定程度上提升预测的精准度。因此为了获得大量的宏观数据，参考经济学家普遍使用的替代方法，我们选择用工业生产指数来代替 GDP，用消费者物价指数的同比增长来描述通货膨胀率。

宏观数据的方差会随时间的改变而产生波动。为了更直观地探讨我国宏观数据方差的时变性，在提出模型之前，我们首先观察主要预测变量的方差变化情况。我们将 10 个工业生产指数增长率的数据作为一组，每组都有 9 个数据与前一组重叠。即第一组为数据 1 至 10，然后按顺序从组中删除第一个数据再添加该组外后面的第一个数据，即第二组包含数据 2 至 11，以此类推……最后依次计算每组的方差。对消费者物价指数数据执行相同的操作。最终计算出的 IP 和 CPI 各自方差随时间的波动显示如图 2.1。

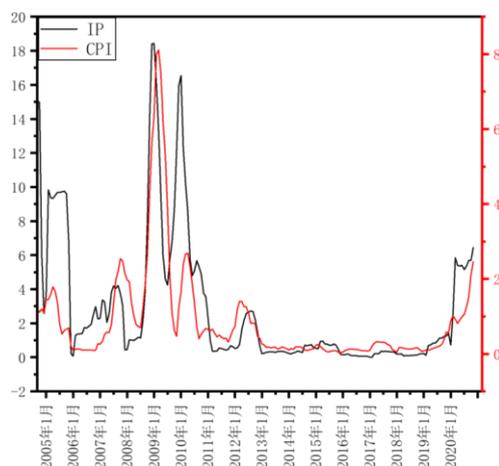


图 2.1 IP 与 CPI 的方差变化

通过观察 IP 和 CPI 增长率方差的变化，我们可以发现它们的方差不是静态不变的，而是会随时间推移产生随机波动。这个现象说明在预测时使用具有恒定方差的模型可能已经无法满足实际需求，而是需要非静态方差模型来捕捉这种波动。对于不同的变量，方差不仅在变化，而且变化规律也不尽相同，因此单一模型可能无法完全捕捉这些变化。所以，在预测不同经济指标时我们要根据变量的特性更换更适合的模型结构。

### 3. 计量模型的构建

#### 3.1 贝叶斯向量自回归模型及结构

贝叶斯分析方法是基于贝叶斯定理发展起来的一种计量方法，该方法通常将未知参数的先验信息与样本信息相结合，再通过贝叶斯定理得出后验分布，最后根据后验信息推断未知参数。经济学家可以通过该方法系统阐述和解决统计问题，也可有效避免样本信息不足的问题。一个完整的贝叶斯包括分析数据、构造概率模型、假设先验信息以及最后的决策。

向量自回归模型（VAR）构建的基本思路是源于数据的统计性质，在事先不设定任何约束条件的情况下，构造时将模型中所有当期变量对各自的若干滞后变量进行回归，也就是说，回归式中包含了每一个内生变量及其滞后值。那么此时单一变量自回归模型被拓展到向量自回归模型，该模型由多元时间序列变量组成。这本小节中，我们首先对本研究中的基本模型展开介绍。

首先，令  $\mathbf{y}_t = (y_{1t}, \dots, y_{nt})'$  表示时间  $t$  时  $n \times 1$  维的向量，其中包含  $n$  个变量，则基本的 VAR( $p$ ) 模型如下：

$$\mathbf{y}_t = \mathbf{b}_0 + \mathbf{B}_1 \mathbf{y}_{t-1} + \dots + \mathbf{B}_p \mathbf{y}_{t-p} + \mathbf{u}_t \quad (3.1)$$

其中  $\mathbf{b}_0$  是一个  $n \times 1$  维的截距矩阵， $\mathbf{B}_1, \dots, \mathbf{B}_p$  是由系数组成的  $n \times n$  维矩阵，换句话说，VAR( $p$ ) 是一个多方程回归模型。 $\mathbf{u}_t$  服从正态分布  $N(0, \Sigma)$  并且被假定为独立均匀分布。当存在  $T$  个时期时，将公式展开即

$$\begin{pmatrix} \mathbf{y}_1' \\ \mathbf{y}_2' \\ \vdots \\ \mathbf{y}_T' \end{pmatrix} = \begin{pmatrix} 1 & \mathbf{y}_0' & \Lambda & \mathbf{y}_{1-p}' \\ 1 & \mathbf{y}_1' & \Lambda & \mathbf{y}_{2-p}' \\ \vdots & \vdots & \vdots & \vdots \\ 1 & \mathbf{y}_{T-1}' & \Lambda & \mathbf{y}_{T-p}' \end{pmatrix} \begin{pmatrix} \mathbf{b}_0' \\ \mathbf{B}_1' \\ \vdots \\ \mathbf{B}_p' \end{pmatrix} + \begin{pmatrix} \mathbf{u}_1' \\ \mathbf{u}_2' \\ \vdots \\ \mathbf{u}_T' \end{pmatrix} \quad (3.2)$$

当总计有  $T$  个时期时，则共有  $T$  个等式，各个等式都有  $k = np + 1$  个回归量，因此 VAR 模型系数的总数可能远远大于统计到的观测值总数。

接着，令将因变量堆叠为  $T \times n$  维的向量矩阵  $\mathbf{Y}$ ，则第  $t$  行为  $\mathbf{y}_t'$ ；令  $\mathbf{X}$  为  $T \times k$  维的回归量矩阵，则第  $t$  行为  $\mathbf{x}_t'$ ，即  $\mathbf{x}_t' = (1, \mathbf{y}_{t-1}', \dots, \mathbf{y}_{t-p}')$ ；令  $\mathbf{B} = (\mathbf{b}_0, \mathbf{B}_1, \dots, \mathbf{B}_p)'$  表示一个  $k \times n$  维的 VAR 系数矩阵。将它们结合起来得到的 VAR( $p$ ) 模型如下：

$$\mathbf{Y} = \mathbf{XB} + \mathbf{U} \quad (3.3)$$

这里的  $\mathbf{U}$  是一个  $T \times n$  维的矩阵, 服从正态分布

$$\text{vec}(\mathbf{U}) \sim N(\mathbf{0}, \mathbf{\Sigma} \otimes \mathbf{I}_T) \quad (3.4)$$

其中含有一个灵活的 Kronecker 误差协方差结构  $\mathbf{\Sigma} \otimes \mathbf{I}_T$ ,  $\otimes$  是 Kronecker 乘积,  $\mathbf{\Sigma}$  是一个  $n \times n$  维的协方差矩阵,  $\mathbf{I}_T$  是维度为  $T$  的单位矩阵。

为了解决大数据下应用 VAR 进行宏观预测时可能会影响结果的几个困难, 我们选择采用基于特定 Kronecker 结构的高维贝叶斯向量自回归模型, 该模型的协方差矩阵中可以包含非高斯、共同随机波动和序列相关的创新结构。通过组合不同的内容可以使模型分别包含多种协方差结构, 从而应对不同数据下的预测要求。

这里我们生成了一个更具有普适性的协方差结构, 即将  $\mathbf{I}_T$  变为  $\mathbf{\Omega}$ , 则

$$\text{vec}(\mathbf{U}) \sim N(\mathbf{0}, \mathbf{\Sigma} \otimes \mathbf{\Omega}) \quad (3.5)$$

### 3.2 高维 BVAR 的先验概率

在第 1 章中我们对大数据下 VAR 模型估计存在的问题及常用解决方案进行了简单说明。在通过大数据集预测宏观经济活动指标时, VAR 模型中需要解决的主要有过度参数化问题和如何令协方差具有时变性质。在这一小节里, 我们首先说明如何解决过度参数化的问题。

在预测时, 如果使用传统的经典 VAR 模型, 即使试图预估的目标变量较少, VAR 的参数也会非常多, 所需观测值的数量仍十分巨大以至于难以收集到, 当 VAR 规模较大时更是如此。这时若没有信息先验或正则化, 过多的 VAR 参数会导致模型的预测能力降低。为了解决这一问题, BVAR 选择使用含有非数据信息的先验概率, 大部分先验的共同出发点是对模型系数进行收缩估计, 这使得诸如脉冲响应和预测的特征等将被趋于精确估计。因此在本节里, 我们将讨论使用高维 BVAR 进行研究时常用的几种先验分布: 明尼苏达先验、自然共轭先验以及随机搜索变量选择先验 (Stochastic Search Variable Selection 简记为 SSVS)。这些先验的共同目标是将一个不受限制的 VAR 收缩为简约的 VAR, 从而减轻因大量复杂的宏观经济数据而导致的过度参数化和过度拟合的影响。这些先验概率在如何实现这一目标、能否有解析解、是否能简化 MCMC 算法来估计后验分布以及是否能够更便捷地应用于更灵活的 VAR 模型方面有所差别。

#### 3.2.1 明尼苏达先验

明尼苏达大学或明尼阿波利斯联邦储备银行的研究人员 Doan, Litterman 和 Sims(1984) 以及 Litterman (1986) 首先提出了在小型 VAR 情况下的收缩先验<sup>[8][30]</sup>。他们提出了一种易于被其他研究人员标准化和复制, 并能够反映许多研究情况的先验分布, 这便是明尼苏达先验。该文章得出一个近似值, 认为一个好的标准公共先验可能是这些参数先验的加权平均值, 并将先验分布指定为向量自回归系数的多元正态分布, 这可以极大简化先验的设定, 能够解决向量自回归模型中系数过多的问题。在该设定下, 除了每个回归式中因变量的第一阶滞后项系数期望值是 1 以外, 其他所有项系数的先验均值都是零。因此先验值在其均值附近被严格限制, 意味着每个方程都采用随机波动的形式。实证表明, 明尼苏达先验可以直接应用于大型 VAR。

对于一个简单的包含  $n$  个变量的 VAR(p) 模型而言有

$$y_{it} = \sum_{r=1}^n \sum_{\tau=1}^T b_{i,r,\tau} y_{r,t-\tau} + u_{it}$$

其中  $b_{i,r,\tau}$  代表第  $i$  个方程中变量  $y_r$  的  $\tau$  阶滞后项系数, 且服从正态分布  $N(\delta_{i,r,\tau}, S_{i,r,\tau}^2)$ 。则此时明尼苏达先验的基本假设有: (1) 正态性。模型中的随机扰动向量  $\mathbf{u}_t$  遵从多元正态分布  $N(\mathbf{0}, \mathbf{\Sigma})$ ; (2) 独立性。协方差矩阵  $\mathbf{\Sigma}$  与模型系数矩阵相互独立; (3) 误差协方差矩阵  $\mathbf{\Sigma}$  的先验分布假设为扩散先验分布; (4) 除与方程左边变量本身相关的一阶滞后变量系数期望为 1 外, 其他变量的系数期望为 0; (5) 等式中变量系数的先验协方差矩阵

和回归方程误差方差的标准 OLS 估计有关。

注意，任何等式里面 VAR 回归右侧的解释变量都由三类参数构成：截距、与变量自身滞后相关的系数以及变量与其他变量滞后相关的系数。其中，对于增长率类型的数据（例如 IP 增长率或 CPI 增长率等），其先验均值通常被设为零。该先验的平均值确保了 VAR 系数的收缩，减少了过度拟合的风险，并反映了设置该先验的前提，即增长率数据通常不会反复出现。对于货币供应或消费水平之类的水平数据，明尼苏达先验使用先验均值来表示出这类数据随机游走分布方面的偏向。在实际应用中，第一个滞后系数的先验均值也可以被设为非 1 值，即变量能表现出一定程度的持久性，而不是单位根。

### 3.2.2 自然共轭先验

对于正态线性回归模型的参数，当其后验概率分布与先验概率分布属于同种时，二者被称作共轭分布，而先验概率分布为似然函数的共轭先验。此时代数运算变得较为方便，因为这能够令具有不同参数的先验概率分布和后验概率分布拥有同样的形式。一方面合符人的直观验看，可以直接给出后验分布的解析解；另一方面在反复计算时能够形成一个先验链。

为了描述其特定形式，我们首先要做以下说明。设有一个  $n \times n$  维的矩阵  $\Phi$ ，当它的分布满足

$$f(\Phi; \nu, \mathbf{S}) = \frac{|\mathbf{S}|^{\nu/2}}{2^{n\nu/2} \Gamma_n(\nu/2)} |\Phi|^{-\frac{\nu+n+1}{2}} \exp\left[-\frac{1}{2} \text{tr}(\mathbf{S}\Phi^{-1})\right]$$

时，则称该矩阵服从 inverse-Wishart 分布，即  $\Phi \sim IW(\nu, \mathbf{S})$ 。其中  $\nu > 0$  是形状参数， $\mathbf{S}$  为规模矩阵， $\Gamma_n$  为多项伽马方程。当  $\nu > m+1$  时， $E\Phi = \mathbf{S}/(\nu - m - 1)$ 。

接下来再设有一个  $m \times n$  维的随机矩阵  $\mathbf{W}$ ，则当给定  $\Phi$  后有  $(\text{vec}(\mathbf{W}) | \Phi) \sim N(\text{vec}(\mathbf{M}), \Phi \otimes \mathbf{P})$  且  $\Phi \sim IW(\nu, \mathbf{S})$  时，我们称  $\mathbf{W}$  和  $\Phi$  服从 normal-inverse-Wishart 分布，写作  $(\mathbf{W}, \Phi) \sim NIW(\mathbf{M}, \mathbf{P}, \nu, \mathbf{S})$ 。具体分布函数为

$$f(\mathbf{W}, \Phi; \mathbf{M}, \mathbf{P}, \nu, \mathbf{S}) \propto |\Phi|^{-\frac{\nu+m+n+1}{2}} \exp\left[-\frac{1}{2} \text{tr}(\Phi^{-1}(\mathbf{W} - \mathbf{M})' \mathbf{P}^{-1}(\mathbf{W} - \mathbf{M}))\right] \exp\left[-\frac{1}{2} \text{tr}(\Phi^{-1} \mathbf{S})\right]$$

上式推导的过程为

$$\begin{aligned} [\text{vec}(\mathbf{W} - \mathbf{M})]' (\Phi \otimes \mathbf{P})^{-1} \text{vec}(\mathbf{W} - \mathbf{M}) &= [\text{vec}(\mathbf{W} - \mathbf{M})]' (\Phi^{-1} \otimes \mathbf{P}^{-1}) \text{vec}(\mathbf{W} - \mathbf{M}) \\ &= [\text{vec}(\mathbf{W} - \mathbf{M})]' \text{vec}[\mathbf{P}^{-1}(\mathbf{W} - \mathbf{M})\Phi^{-1}] \\ &= \text{tr}(\Phi^{-1}(\mathbf{W} - \mathbf{M})' \mathbf{P}^{-1}(\mathbf{W} - \mathbf{M})) \end{aligned}$$

则根据定义可知  $(\text{vec}(\mathbf{W}), \Phi)$  的联合密度函数

$$\begin{aligned} f(\mathbf{W}, \Phi) &\propto |\Phi|^{-\frac{\nu+n+1}{2}} \exp\left[-\frac{1}{2} \text{tr}(\Phi^{-1} \mathbf{S})\right] |\Phi \otimes \mathbf{P}|^{-\frac{1}{2}} \\ &\quad \exp\left\{-\frac{1}{2} [\text{vec}(\mathbf{W} - \mathbf{M})]' (\Phi \otimes \mathbf{P})^{-1} \text{vec}(\mathbf{W} - \mathbf{M})\right\} \\ &= |\Phi|^{-\frac{\nu+m+n+1}{2}} \exp\left[-\frac{1}{2} \text{tr}(\Phi^{-1} \mathbf{S})\right] \exp\left\{-\frac{1}{2} \text{tr}[\Phi^{-1}(\mathbf{W} - \mathbf{M})' \mathbf{P}^{-1}(\mathbf{W} - \mathbf{M})]\right\} \end{aligned}$$

即该密度函数分布与  $f(\mathbf{W}, \Phi; \mathbf{M}, \mathbf{P}, \nu, \mathbf{S})$  相同。

### 3.2.3 随机搜索变量选择先验 (SSVS)

这一节我们将介绍另一个比较常用的 VAR 系数的收缩先验——随机搜索变量选择先验 (SSVS)，这个先验最早由 George 等 (2008) 提出。SSVS 的基本思想是为模型中应该包含的参数分配常用的先验方差，同时令不相关的参数先验方差趋近于零。这样便可以估算出相关参数，并且使得无关变量对预测和脉冲响应不产生显著影响。

SSVS 先验的基础概念来源于独立正态先验和 inverse-Wishart 先验,但它没有统一为系数设定约束,而是在系数的正态先验上引入了层次结构,其主要思想是以基于数据的方式将系数中的元素分为两组:第一组的系数被强烈收缩至 0,而在第二组中这些系数则没有收缩。也就是说,模型名称中“变量选择”意味着被选择的部分是由于令第一组中的系数趋近于 0 而迫使第二组的变量被选择。这个区分每次在 MCMC 抽样时的迭代中都是随机完成的,因此也被称为“随机搜索”。

### 3.3 高维 BVAR 的误差协方差分布

在第 1 章中提到,大数据下 BVAR 的预测需要解决的第二个问题是时变协方差。尽管目前具有标准误差假设的高维 BVAR 模型在实证方面取得了一定的成功,尤其是在应用美国数据时,同方差、高斯假设与序列独立假设下的实证预测都得到了相对理想的结果。但通过对实际数据的应用和观察,我们可以发现许多特征在数据中经常出现,如时变波动性、非高斯误差分布、共同随机波动性等等,这些特性显然可以在宏观经济时间序列建模时被考虑进去从而帮助提高模型性能。这一小节中我们将介绍三种模型里用到的协方差结构。

#### 3.3.1 非高斯误差结构

近期研究发现,当 VAR 模型拥有像类似 t 分布这种具有重尾误差分布时,其在拟合实际数据和预测未来经济走向上得到的结果比具有传统高斯误差分布的模型更好<sup>[31]</sup>。

由于许多分布可以写成高斯分布的比例混合,因此 (3.5) 式的框架可以容纳许多非高斯分布。为了实现这一目标,我们可以假设 VAR 的误差为 t 分布且独立同分布,则  $\mathbf{\Omega} = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_T)$ , 并且在给定  $\nu$  的条件下每个  $\lambda_i$  都服从反伽马分布即  $(\lambda_i | \nu) \sim \text{IG}(\nu/2, \nu/2)$ 。因此,在这种情况下  $\mathbf{u}_i$  的边际形式是一个均值向量为  $\mathbf{0}$ 、自由度参数为  $\nu$  的多元 t 分布,而  $\mathbf{\Sigma}$  为规模矩阵<sup>[32]</sup>。此时  $\mathbf{\Omega}^{-1}$  是一个对角矩阵,即  $\mathbf{\Omega}^{-1} = \text{diag}(\lambda_1^{-1}, \lambda_2^{-1}, \dots, \lambda_T^{-1})$ 。

#### 3.3.2 共同随机波动分布

方差能反应数据的波动性,是对宏观数据研究时需要考虑的最有用的特征之一。显然,在实际经济活动中,不同时间区间中经济数据的变动趋势都并非一成不变。当受到政策变化、经济冲击等情况的影响,宏观经济数据的波动率会发生改变,此时具有同方差误差的模型则无法描述数据的这类特征,从而导致模型结果的偏颇,因此一些经济学家提出了新的拓展模型。参考 Carriero、Clark 和 Marcellino (2016) 提出的模型:他们在异方差创新结构的误差协方差中增加了共同随机波动率。该模型的误差协方差矩阵增加了一个共同的时变因子,它可以解释为整体宏观经济的波动。

具体而言,该模型的误差协方差分布被规定为:  $\mathbf{u}_i \sim N(\mathbf{0}, e^{h_i} \mathbf{\Sigma})$ , 共同波动的时变因子即  $e^{h_i}$ , 其对数波动率遵循一个固定的 AR(1) 过程:

$$h_i = \rho h_{i-1} + \varepsilon_i^h$$

其中  $\varepsilon_i^h \sim N(0, \sigma_h^2)$ ,  $|\rho| < 1$ , 注意这里假设了该 AR(1) 过程的无条件均值零。则根据前文的分解方式,共同随机波动分布中误差协方差矩阵  $\mathbf{\Omega} = \text{diag}(e^{h_1}, e^{h_2}, \dots, e^{h_T})$ , 该矩阵也为对角矩阵。但这种方差分布设定有一个缺点,那就是所有的方差都由一个单一因素缩放,也因此这些方差都是成正比的。不过正如之前的文献所述,这个特点在许多宏观研究中被证实是实际存在的,即许多宏观经济变量往往会一起波动,并且波动方向一致。具体使用了这种共同随机波动的方法也可以参考 Mumtaz (2016) 的文章等<sup>[33][34]</sup>。

#### 3.3.3 序列相关方差

通常在研究中会假设方差序列非相关,但 (3.5) 式中模型的结构可以处理序列相关的方差,一个具体的例子如下。

假设  $\mathbf{u}_i$  服从一个 MA(2) 过程:

$$\mathbf{u}_i = \varepsilon_i + \psi_1 \varepsilon_{i-1} + \psi_2 \varepsilon_{i-2}$$

其中  $\varepsilon_i \sim N(\mathbf{0}, \mathbf{\Sigma})$ ,  $\psi_1$  和  $\psi_2$  满足可逆条件。这个过程可以嵌套在以下矩阵框架中,此时

$$\Omega = \begin{pmatrix} \omega_0 & \omega_1 & \omega_2 & 0 & \Lambda & 0 \\ \omega_1 & \omega_0 & \omega_1 & 0 & 0 & M \\ \omega_2 & \omega_1 & \omega_0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \omega_2 \\ M & 0 & 0 & 0 & 0 & \omega_1 \\ 0 & \Lambda & 0 & \omega_2 & \omega_1 & \omega_0 \end{pmatrix}$$

其中  $\omega_0 = \psi_1^2 + \psi_2^2 + 1$ ,  $\omega_1 = \psi_1(1 + \psi_2)$ ,  $\omega_2 = \psi_2$ 。

但是 (3.5) 式的结构有一个缺点, 便是需要要求所有元素的 MA 系数必须相同, 当 MA 过程变得更加复杂, 如  $\psi_1$  和  $\psi_2$  变成了  $n \times n$  维的系数矩阵时, 方差便无法被写成 Kronecker 结构, 那么原 VAR 的框架便无法支撑了。不过这种受限的序列相关性形式仍能够被用于捕获数据的持久影响性。

现在我们假设  $\mathbf{u}_t$  服从一个 MA(1) 过程, 这样  $\Omega$  的结构即为一个带状矩阵而不是对角矩阵:

$$u_{it} = \gamma_{it} + \psi \gamma_{i,t-1}$$

其中,  $|\psi| < 1$  并且  $\gamma_{it}$  服从正态分布  $N(0,1)$ 。设定该过程的初始化为  $u_{i1} \sim N(0, 1 + \psi^2)$ , 则 MA(1) 过程的矩阵可以写成

$$\mathbf{u}_i = \mathbf{H}_\psi \boldsymbol{\gamma}_i$$

其中  $\mathbf{u}_i = (u_{i1}, \dots, u_{iT})'$ ,  $\boldsymbol{\gamma}_i = (\gamma_{i1}, \dots, \gamma_{iT})'$ , 且

$$\mathbf{H}_\psi = \begin{pmatrix} 1 & 0 & \Lambda & 0 \\ \psi & 1 & \Lambda & 0 \\ M & 0 & 0 & M \\ 0 & \Lambda & \psi & 1 \end{pmatrix}$$

因此  $\boldsymbol{\gamma}_i$  服从  $N(\mathbf{0}, \boldsymbol{\Sigma}_\psi)$  分布,  $\boldsymbol{\Sigma}_\psi = \text{diag}(1 + \psi^2, 1, \dots, 1)$ , 所以  $\mathbf{u}_i$  的协方差矩阵是  $\mathbf{H}_\psi \boldsymbol{\Sigma}_\psi \mathbf{H}_\psi'$ 。则协方差结构的表达式为  $\boldsymbol{\Sigma} \otimes \Omega$ , 其中  $\Omega = \mathbf{H}_\psi \boldsymbol{\Sigma}_\psi \mathbf{H}_\psi'$ , 等式右边的  $\boldsymbol{\Sigma}_\psi$  和  $\mathbf{H}_\psi$  均为带状矩阵。需要说明的是, 对于一般的 MA(q) 过程, 只需要重新正确定义  $\boldsymbol{\Sigma}_\psi$  和  $\mathbf{H}_\psi$ , 然后重复相同的过程即可。

## 4. 实证分析

在前面的部分中我们说明了如何通过先验假设解决大数据下应用高维 BVAR 对中国宏观经济进行预测时存在的过度参数化问题, 并且介绍了几种协方差的时变结构。在本章中我们将通过具有不同先验假设以及不同时变协方差结构的模型对中国宏观经济变量进行点预测和密度预测, 根据预测比较结果来判断这些模型的预测能力, 并且探讨能够提高中国宏观经济预测精确度的因素。

### 4.1 样本外预测结果

在本节中, 我们将通过递归的样本外预测来对中国宏观经济进行实时研究, 应用第 2 章说明的 20 个变量以及第 3 章中讨论的相关先验假设与误差协方差结构, 对不同的多维 BVAR 模型进行说明, 从而评估不同模型在点预测和密度预测方面的能力, 同时探讨预测中国宏观经济时能够提高精确度的因素。

#### 4.1.1 模型及先验假设

预测时, 我们主要关注以下四个变量: IP、CPI、国家财政收入以及一年期国债收益率。我们对比了各模型在预测当期数据、提早 1 个月和提早 6 个月数据时的结果。由于报告的滞后性, 在时间为  $t$  时可获得的实时数据仅包含到  $t-1$  月度的观测值。

在模型选择上, 我们通过为基础模型中逐步增加协方差结构来比较预测结果, 这样能够

更直观地看出不同协方差结构对预测结果的改进作用。我们还将高维 BVAR 与第 3 章中讨论的不同的先验假设相结合,以便对先验分布不同的各个模型的预测效果进行比较,从而探讨何种先验假设更适合在预测中国宏观经济时应用。此外,为了观察更大量的数据集对预测结果是否有改进作用,我们另外选取了两个小型 BVAR,即设定为明尼苏达先验的 BVAR-small 以及设定为 SSVS 先验的 BVAR-SSVS-small,其预测时只使用了四个核心变量即 IP、CPI、国家财政收入以及一年期国债收益率。具体选择的模型完整描述见下表 4.1。

表 4.1 应用于预测的模型说明

模型类型	说明
BVAR-small	小型 BVAR, 含 4 个变量, 为明尼苏达先验
BVAR-SSVS-small	小型 BVAR, 含 4 个变量, 为 SSVS 先验
BVAR-Minn	BVAR 模型, 含 20 个变量, 为明尼苏达先验
BVAR-NCP	BVAR 模型, 含 20 个变量, 为自然共轭先验
BVAR-CSV	BVAR 模型, 含 20 个变量, 协方差为共同随机波动结构
BVAR-t-CSV	BVAR 模型, 含 20 个变量, 协方差为 t 分布与共同随机波动结构
BVAR-t-CSV-MA	BVAR 模型, 含 20 个变量, 协方差为 t 分布、共同随机波动与 MA 结构

当比较不同模型的预测能力时,除了准确的点预测外,还应考虑预测分布的不确定性即密度预测。为了便于比较,并为每个模型中同样的参数选择一致的赋值,同时为误差协方差结构不同的模型选择相同的先验条件。

除了说明中为明尼苏达先验的模型外,其他模型中  $(\mathbf{B}, \Sigma)$  的联合密度函数服从

$$\Sigma \sim IW(S_0, \nu_0), (\text{vec}(\mathbf{B}) | \Sigma) \sim N(\text{vec}(\mathbf{B}_0), \Sigma \otimes \mathbf{V}_B)。$$

为了方便对比,当模型中具有共同的参数时,我们会给它们赋予同样的数值。在先验为明尼苏达先验的模型中,令  $\mathbf{B}_0 = \mathbf{0}$ , 并且令  $\mathbf{V}_B$  中包含的三个超参数的值分别为  $c_1 = 0.2^2$ 、 $c_2 = 0.1^2$ 、 $c_3 = 10^2$ , 这些超参数的取值已被应用于多种文献。在先验为自然共轭先验的模型中,我们仍令  $\mathbf{B}_0 = \mathbf{0}$ , 并且令协方差矩阵  $\mathbf{V}_B$  中包含的两个超参数的值分别为  $c_1 = 0.2^2$ 、 $c_2 = 10^2$ 。同时,令  $\Sigma$  的概率密度分布参数  $\nu_0 = n + 3$ 、 $\mathbf{S}_0 = \text{diag}(s_1^2, \dots, s_n^2)$ , 这里的  $s_i^2$  代表误差方差对角矩阵中第  $i$  行的标准 OLS 估计。

关于模型中误差协方差的参数设置如下:对于误差协方差为非高斯 t 分布结构的模型,我们假设其中自由度参数  $\nu$  服从一个统一先验即  $\nu \sim U(2, 50)$ 。对于结构为共同随机波动的模型,我们假设参数  $\rho \sim N(\rho_0, V_\rho)$  (其中  $|\rho| < 1$ ),  $\sigma_h^2$  具有独立的反伽马先验分布即  $\sigma_h^2 \sim IG(\nu_h, S_h)$ 。为了使  $\rho$  的分布中心在 0.9 处且  $\sigma_h^2$  的先验均值为  $0.1^2$ , 我们将初始化参数分别设为  $\nu_{h_0} = 5$ ,  $S_{h_0} = 0.04$ ,  $\rho_0 = 0.9$ ,  $V_\rho = 0.2^2$ 。对于误差协方差为序列相关 MA

(1) 的模型,我们假设其系数  $\psi$  服从截断正态分布,即  $\psi \sim N(\psi_0, V_\psi)$  (其中  $|\psi| < 1$ ), 则 MA (1) 是一个可逆的过程。令  $\psi_0 = 0$ ,  $V_\psi = 1$ , 则此时先验的中心在 0 附近,并且分布区间为  $(-1, 1)$ 。考虑到先验方差较大,这个设定相对不包含信息。

#### 4.1.2 预测评价指标

最终在正式比较所提出的高维 BVAR 模型的预测性能时,主要通过模型对目标变量的点预测和密度预测结果来判断。这里我们给出两个常用的评价指标。

假设数据的时间跨度为从 1 到  $t$ , 表示为  $\mathbf{Y}_{1:t}$ , 并且使用 MCMC 采样器获得这些给定数据的后验分布。然后令  $E(y_{i,t+m} | \mathbf{Y}_{1:t})$  表示变量  $i$  的后验预测分布期望值,令  $p(y_{i,t+m} | \mathbf{Y}_{1:t})$  表示观察值在后验预测密度函数中的取值。接下来,我们向前移动一个周期,使用数据  $\mathbf{Y}_{1:t+1}$  重复整个过程,依此类推。最终,我们从时间  $t = t_0$  到  $t = T - m$  重复了这个过程。

##### (1) 均方根预测误差 (RMSFE)

这个矩阵评估了模型的点预测结果。令  $y_{i,t+m}^o$  表示  $y_{i,t+m}$  在  $t+m$  时期的真实值,则矩阵

公式为

$$RMSFE = \sqrt{\frac{\sum_{t=t_0}^{T-m} (y_{i,t+m}^o - E(y_{i,t+m} | \mathbf{Y}_{1,t}))^2}{T-m-t_0+1}}$$

显然, RMSFE 的值越小表明该模型的点预测结果准确程度越高。

#### (2) 对数预期似然平均 (ALPL)

为了评估模型的密度预测结果, 其中一个方法是计算预期似然概率, 即  $y_{i,t+m}$  在真实值  $y_{i,t+m}^o$  处的预期密度分布。对数结果与分配给实际发生事件的概率 (密度) 单调相关, 并且这种方法不用考虑所采用的模型是否为真实的<sup>[35]</sup>。令  $p(y_{i,t+m} = y_{i,t+m}^o | \mathbf{Y}_{1,t})$  表示变量  $i$  的预期似然概率, 如果实际结果贴近预测值, 则预期似然的值将很大, 反之则很小。定义 ALPL 为

$$ALPL = \frac{1}{T-m-t_0+1} \sum_{t=t_0}^{T-m} \log p(y_{i,t+m} = y_{i,t+m}^o | \mathbf{Y}_{1,t})$$

则此时 ALPL 的取值越大, 代表模型的密度预测结果越好。

### 4.2.3 模型预测结果比较

表 4.2 至表 4.5 给出了几种模型的预测结果比较, 包括各模型对当期数据、提前 1 个月和提前 6 个月数据的预测结果。我们将含有 20 个变量、先验假设为明尼苏达先验的 BVAR-Minn 模型当做基准 BVAR, 并以此为基准同其他模型进行比较。需要指出, 对于均方根预测误差 (RMSFE) 指标, 下面所有表格中给出的数值都是给定模型的 RMSFE 值与 BVAR-Minn 的 RMSFE 值之比, 因此小于 1 的值意味着该模型的点预测结果优于 BVAR-Minn; 对于对数预期似然平均 (ALPL) 指标, 下面所有表格中给出的是该模型的 ALPL 值与 BVAR-Minn 的 ALPL 值之差, 因此若差值为正数, 则表示该模型的密度预测结果优于基准 BVAR-Minn 模型。

表 4.2 预测工业生产指数 (IP) 时相对于基准 BVAR 模型 (BVAR-Minn) 的结果

模型类型	RMSFE 比值			ALPL 差值		
	当期	1 个月	6 个月	当期	1 个月	6 个月
BVAR-small	1.00	1.00	1.00	0.00	0.10	0.00
BVAR-SSVS-small	1.00	1.00	1.00	0.00	0.10	0.00
BVAR-NCP	1.00	1.00	1.00	-0.10	0.00	0.00
BVAR-CSV	0.88	1.00	1.00	0.30	0.20	0.30
BVAR-t-CSV	0.88	1.00	1.00	0.40	0.20	0.30
BVAR-t-CSV-MA	1.00	1.00	1.00	0.30	0.20	0.30

表 4.3 预测消费者物价指数 (CPI) 时相对于基准 BVAR 模型 (BVAR-Minn) 的结果

模型类型	RMSFE 比值			ALPL 差值		
	当期	1 个月	6 个月	当期	1 个月	6 个月
BVAR-small	0.67	1.00	0.67	0.10	0.10	0.10
BVAR-SSVS-small	0.67	1.00	1.00	0.10	0.10	0.10
BVAR-NCP	1.00	1.00	1.00	0.00	0.00	0.10
BVAR-CSV	1.00	1.00	1.00	0.00	0.10	0.10
BVAR-t-CSV	1.00	1.00	1.00	0.20	0.20	0.10
BVAR-t-CSV-MA	1.00	1.00	1.00	0.10	0.20	0.10

表 4.4 预测国家财政收入时相对于基准 BVAR 模型 (BVAR-Minn) 的结果

模型类型	RMSFE 比值			ALPL 差值		
	当期	1 个月	6 个月	当期	1 个月	6 个月
BVAR-small	0.98	0.95	1.00	-0.10	-0.10	-0.10
BVAR-SSVS-small	1.04	1.00	1.00	0.10	0.10	0.20
BVAR-NCP	1.04	1.08	1.00	0.00	0.00	0.10
BVAR-CSV	0.74	0.94	1.00	1.00	0.60	0.70
BVAR-t-CSV	0.74	0.94	1.00	1.00	0.50	0.70
BVAR-t-CSV-MA	0.80	0.94	1.00	0.90	0.50	0.60

表 4.5 预测一年期国债收益率时相对于基准 BVAR 模型 (BVAR-Minn) 的结果

模型类型	RMSFE 比值			ALPL 差值		
	当期	1 个月	6 个月	当期	1 个月	6 个月
BVAR-small	0.06	0.06	0.06	2.90	2.80	2.80
BVAR-SSVS-small	0.06	0.06	0.06	2.90	2.80	2.80
BVAR-NCP	1.12	1.10	1.04	0.00	0.00	0.00
BVAR-CSV	0.96	1.00	0.98	0.20	0.20	0.20
BVAR-t-CSV	0.96	1.00	0.98	0.20	0.20	0.20
BVAR-t-CSV-MA	1.04	1.00	0.98	0.20	0.20	0.20

总体来看,当先验分布不同时,对于小型模型,具有明尼苏达先验的 BVAR-small 和具有 SSVS 先验的 BVAR-SSVS-small 两个模型在点预测上的表现大致相同,BVAR-small 在预测国家财政收入时略好;对于大型模型,具有自然共轭先验的模型预测表现,无论是点预测还是密度预测都并未优于 BVAR-Minn。

在同种先验下,在模型中添加时变协方差结构能改进模型的预测性能。对于四个主要经济变量的预测,具有时变协方差结构的模型与基准模型 BVAR-Minn 的 RMSFE 比值绝大多数都小于或等于 1,说明具有时变协方差的模型能更好地对宏观经济变量进行点预测;在密度预测方面,时变协方差对模型预测性能的总体改进要好于对点预测的改进,大部分具有时变协方差结构的模型与 BVAR-Minn 的 ALPL 差值都是大于 0 的数,这表示预测值更贴近实际结果。

为了判断大量数据能否改善预测结果,我们使用了仅具有 4 个关键变量的小型模型 BVAR-small 同其他具有 20 个变量且同样为明尼苏达先验的大型模型进行比较。结果表明除了预测一年期国债收益率时小型模型表现较好外,在预测其他三个变量时大量宏观数据确实能够提高模型预测的精确度,这突出了额外的宏观经济和金融数据在提高模型预测方面的价值。在预测其他三个变量时,具有时变协方差结构的高维 BVAR 模型与基准模型的 ALPL 差值不论在预测当期值还是提前 1 或 6 个月的数据时几乎都大于小型模型与基准模型的差值。这说明大数据在密度预测方面对模型精确度有较大提高。更多的宏观经济金融数据能够在预测时提供更多的信息,使得预测值更多地分布在实际结果周围。

接下来我们针对四个表中的预测结果分别展开详细解释:在对工业生产指数(IP)进行预测时,仅有 4 个变量的 BVAR-small、BVAR-SSVS-small 与 BVAR-NCP 模型同 BVAR-Minn 的 RMSFE 比值均为 1,这三种模型与 BVAR-Minn 的 ALPL 差值也基本都为 0,说明在对 IP 的预测上,这四种模型的表现基本一致。而对于 IP 当期值的预测,BVAR-CSV 和 BVAR-t-CSV 模型与 BVAR-Minn 的 RMSFE 之比为 0.875 均小于 1,并且其 ALPL 差值也为正数,说明这两者在点密度和密度预测方面的表现明显优于另外五种模型,即具有更灵活的协方差结构的 BVAR 更具优势。对于 BVAR-t-CSV-MA 模型,其在点密度预测上与基准模型表现一致,在密

度预测上表现略好，但不如 BVAR-t-CSV。

在对消费者物价指数（CPI）进行预测时，没有任何模型能在点预测和密度预测方面始终优于基准模型。在点预测方面，除了 BVAR-small 与 BVAR-SSVS-small 模型外，其他四种模型在预测当期、提前 1 个月和 6 个月的数据时表现均与 BVAR-Minn 一致。在密度预测方面，BVAR-small 和 BVAR-SSVS-small 与基准模型的 ALPL 差值略大于 0，说明其在密度预测上的表现上略好于 BVAR-Minn。但此时，具有 20 个变量且协方差结构更灵活的高维 BVAR 体现出了更好的密度预测，尤其是 BVAR-t-CSV 模型。这与先前的文献研究结果一致，Bernanke 和 Boivin 等人的文章也证明了在模型中添加变量有助于解决价格难题<sup>[36]</sup>。

在对国家财政收入进行预测时，除了具有自然共轭先验的 BVAR-NCP 模型，另外五种模型在点或密度预测方面的结果均较基准模型有所改进。其中，在点密度预测上，改变模型的先验似乎没什么改进效果，改变协方差结构则效果更加明显。BVAR-CSV、BVAR-t-CSV 和 BVAR-t-CSV-MA 这三种对误差协方差结构进行灵活改造过的模型在对当期值以及提早 1 个月的值进行预测时表现都较为突出，但在对提前半年的数据进行预测时则与 BVAR-Minn 的结果一致。这说明这三种模型在对国家财政收入进行点预测时更倾向于短期内的预测，随着预测时间变长其精准度会下降。在密度预测方面，具有 SSVS 先验的小型 BVAR-SSVS-small 模型与 BVAR-Minn 的 ALPL 差值略大于 0，但此时改变协方差结构更能提高预测精确度。BVAR-CSV、BVAR-t-CSV 与 BVAR-t-CSV-MA 依旧体现出倾向短期预测的趋势。通过比较来看，在预测国家财政收入时，BVAR-CSV 和 BVAR-t-CSV 表现较为优秀。

在对一年期国债收益率的预测上，具有自然共轭先验的 BVAR-NCP 模型结果相对于基准模型依旧没有体现较好的改善。在点密度预测上，依旧是 BVAR-CSV 和 BVAR-t-CSV 两种模型体现出较大程度的改进，而 BVAR-t-CSV-MA 的点预测表现相比于没有 MA 结构时则有所下降。在密度预测上，这三种模型得到的结果都好于基准模型。令人意外的是，仅具有 4 个变量的小型 BVAR-small 模型与小型 BVAR-SSVS-small 模型同基准模型的 RMSFE 之比均远小于 1，且其与基准模型的 ALPL 之差高于其他模型一个数量级。这说明在预测一年期国债收益率时，BVAR-small 和 BVAR-SSVS-small 的点预测和密度预测要远好于 BVAR-Minn 以及另外四种模型。这可能是由于影响国债收益率的主要因素是经济发展状况、利率的变动预期、通货膨胀等宏观经济因素，这几个主要影响因素已经包含在小型模型内，对于其他经济变量而言短期内不会有较大波动，这表明短期国债收益率对其他宏观经济因素造成的影响敏感性不高。

我们的研究表明，通过对所有结果进行比较发现，当预测中国宏观经济指标时，不含有序列相关、但具有共同随机波动误差协方差结构的高维 BVAR 模型，即 BVAR-CSV 与 BVAR-t-CSV 模型，其预测结果相比其他模型表现较好，尤其是在密度预测方面。其中，BVAR-t-CSV 模型的表现更好，其往往能提供所有层面的最佳点预测和密度预测。这是因为，理论上 BVAR-CSV 应对较大冲击的主要渠道基本上只有增加共同随机波动结构，这可以增加模型内参数的波动性。而相比之下，BVAR-t-CSV 模型中除了共同随机波动结构外还具有非高斯结构，该模型能使参数概率分布的两侧尾部更厚从而更全面地捕获数据特性，这突出了允许重尾创新的经验相关性。同时我们也发现，仅具有自然共轭先验但没有灵活的误差协方差分布的 BVAR-NCP 模型的表现较差，其预测结果通常不如误差协方差中具有共同随机波动分布和非高斯 t 分布的模型。最后，我们指出在误差协方差中添加 MA 结构对中国宏观经济变量的预测精确度提高程度有限。这种受限的序列相关结构主要用于捕捉数据的持久性影响，由于我们的模型中变量已然含有 4 个滞后项，此时 MA 结构则变得不那么重要。

## 5. 结论

宏观经济预测已成为当下中国经济研究的重要组成部分，通过预测提前掌握国民经济活动的运行状态和发展趋势，从而为政府提供政策制定支持、提高市场运行效率，对于健全我国宏观经济调控体系尤为重要。随着大数据时代到来，宏观经济变量的数据量及复杂程度日益提高，单一的传统预测模型已经无法满足要求，提高模型的计算速度以及预测精确度成为

亟待解决的问题。

为了能够更准确地研究中国宏观经济的发展情况,本文采用了一类新的高维贝叶斯向量自回归(BVAR)模型,该模型能够应用大量数据进行预测,从而有效避免由于样本数据较少产生的局限性。贝叶斯推断方法认为模型会随时间发生随机改变,并且将概率附加到每个模型参数上。在估计模型参数时可以根据不同数据特征对其施加不同的先验信息,从而更准确地刻画经济变量的动态变化,提高模型预测的精准性。同时,文中采用的高维BVAR模型的另一个优点是其误差协方差中含有Kronecker结构,该结构允许非高斯、共同随机波动以及序列相关的创新,并且各模型进行上万次后验绘制所需的时间均在几分钟之内,因此可以用于大大加快模型的计算速度。

为了进一步提高模型预测精确度,首先我们需要解决大数据下VAR模型常见的过度参数化问题。我们选择设定不同的先验,如明尼苏达先验、自然共轭先验和SSVS先验来对模型系数进行收缩估计,以减轻因宏观经济数据过多而导致的过度参数化和过度拟合的影响。提高预测精确度的另一点则是使模型中误差协方差具有时变性质。我们在误差协方差中添加了如非高斯、共同随机波动以及序列相关的时变结构,并在实证中对拥有不同协方差结构的模型进行了样本内拟合、边际似然比较以及样本外预测比较。

在实证分析中,本文选取了我国自2005年1月至2020年11月期间20个宏观经济指标的月度数据,变量包括IP、CPI、国家财政收入、债券收益率等其他宏观经济和金融指标。为了判断能提高中国宏观经济预测精确度的具体因素,我们将拥有明尼苏达先验的BVAR-Minn模型作为基准,通过仅有4个变量的BVAR-small模型和BVAR-SSVS-small模型、拥有20个变量的假设为自然共轭先验的BVAR-NCP模型、具有共同随机波动误差的BVAR-CSV模型、误差协方差为非高斯与共同随机波动分布的BVAR-t-CSV模型和协方差同时具有非高斯、共同随机波动与MA(1)三种结构的BVAR-t-CSV-MA这六个模型与基准模型递归样本外预测结果的比较,展示了各模型在点预测和密度预测方面的结果。

结果表明:对中国宏观经济预测而言,相对于基准模型BVAR-Minn,改变模型先验对预测精确度的提升不大,中国宏观经济指标更适合用拥有不同误差协方差结构灵活的BVAR来进行预测,这能较好地改进点预测和密度预测的精确度。所有模型中BVAR-t-CSV模型的预测表现最佳,即高维贝叶斯模型的误差协方差中具有非高斯结构和共同随机波动结构是主要能够提高预测精确度的因素。在实际经济活动中,各个宏观变量都会随时间发生整体的变动,而该模型的协方差中具有共同时变因子,这能够更好地捕捉这种变动,并且通过增加参数的波动性和令分布尾部更厚来更好地应对宏观波动冲击。同时,大量数据能改善模型的预测结果,尤其是在提高密度预测的精确度方面,更多的数据在预测时提供更多更准确的信息,从而使预测值更贴近实际值。因为不受限制的高维BVAR模型能够考虑行业和变量间联合信息,因此更能应对较强冲击,这一点在预测IP时被格外凸显出来。说明在针对某些变量预测时,应额外考虑多种宏观经济和金融数据的影响,以便提高预测结果的精确度。

在经济全球化与大数据背景下,各个经济体间相互联系日益密切,值得研究的宏观经济变量逐渐增加,各指标包含的数据量也呈现爆炸性增长趋势。在这种复杂条件下,如何更好地把握我国宏观经济的运行规律和趋势成为重要议题。本文这项研究拓展了我国现有的宏观预测模型,为未来研究中国宏观经济走向提供了新思路 and 较为全面的参考,在实际问题的应用上具有广泛前景。

#### 参考文献

- [1] Bańbura M, Giannone D, Reichlin L. Large Bayesian vector auto regressions[J]. *Journal of Applied Econometrics*, 2010, 25(1):71-92.

- [2] George E I, Sun D, Ni S. Bayesian stochastic search for VAR model restrictions[J]. Journal of Econometrics, 2008, 142(1):553-580.
- [3] Carriero A, Clark T E, Marcellino M. Large Bayesian vector autoregressions with stochastic volatility and non-conjugate priors[J]. Journal of Econometrics, 2019.
- [4] Giannone D, Reichlin L, Small D H. Nowcasting GDP and inflation: The real-time informational content of macroeconomic data releases[J]. Working Paper Series, 2006.
- [5] Bańbura M, Giannone D, Modugno M, et al. Chapter 4 – Now-Casting and the Real-Time Data Flow[M]. Elsevier B.V., 2013.
- [6] Giannone D, Reichlin L. Does information help recovering structural shocks from past observations?[J]. ULB -- Universite Libre de Bruxelles, 2006, 4(2-3):455-465.
- [7] Christiano L J, Eichenbaum M S, Evans C L. Monetary policy shocks: What have we learned and to what end?[J]. Working Paper Series, Macroeconomic Issues, 1997.
- [8] Doan T, Litterman R B, Sims C A. Forecasting and conditional projection using Realistic prior distributions[J]. Staff Report, 1984, 3(1):1-100.
- [9] Koop G, Korobilis D. Bayesian multivariate time series methods for empirical macroeconomics[J]. MPRA Paper, 2009, 3(4):267-358.
- [10] Stock J H, Watson M W. Forecasting using principal components from a large number of predictors[J]. Journal of the American Statistical Association, 2002, 97.
- [11] Stock J H, Watson M W. Implications of dynamic factor models for VAR analysis[J]. Social Science Electronic Publishing.
- [12] Mol C D, Giannone D, Reichlin L. Forecasting using a large number of predictors: Is Bayesian shrinkage a valid alternative to principal components?[J]. Journal of Econometrics, 2008, 146(2):318-328.
- [13] Carriero A, Clark T E, Marcellino M. Common Drifting Volatility in Large Bayesian VARs[J]. Journal of Business and Economic Statistics, 2016.
- [14] Berg T O, Henzel S R. Point and density forecasts for the Euro area using Bayesian VARs[J]. International Journal of Forecasting, 2015, 31(4):1067-1095.
- [15] Ramos F F R. Forecasts of market shares from VAR and BVAR models: a comparison of their accuracy[J]. International Journal of Forecasting, 2003, 19(1):95-110.
- [16] Bouri E, Gupta R, Hosseini S, et al. Does global fear predict fear in BRICS stock markets? Evidence from a Bayesian graphical structural VAR model[J]. Emerging Markets Review, 2018, 34.
- [17] Frühwirth-Schnatter S, Wagner H. Stochastic model specification search for Gaussian and partial non-Gaussian state space models[J]. Journal of Econometrics, 2009, 154(1):85-100.
- [18] Korobilis D. Hierarchical shrinkage priors for dynamic regressions with many predictors[J]. International Journal of Forecasting, 2013, 29(1):43-59.
- [19] Blake A P, Mumtaz H. Applied Bayesian econometrics for central bankers[J]. Technical Books, 2012.
- [20] Karlsson S. Chapter 15. Forecasting with Bayesian vector autoregression[J]. Handbook of Economic Forecasting, 2013.

- [21] Primiceri G E. Time varying structural vector autoregressions and monetary policy[J]. Review of Economic Studies, 2005(3):821-852.
- [22] Clark T E, Ravazzolo F. Macroeconomic forecasting performance under alternative specifications of time-varying volatility[J]. Journal of Applied Econometrics, 2015, 30.
- [23] Raghavan M, Athanasopoulos G, Silvapulle P. Canadian monetary policy analysis using a structural VARMA model[J]. Canadian Journal of Economics/Revue canadienne d'économique, 2016, 49(1).
- [24] 李仲莲, 施明义. 我国货币流通量的宏观预测[J]. 金融研究, 1990(06):18-23.
- [25] 赵昕东, 陈飞, 高铁梅. 我国货币政策工具变量效应的实证分析[J]. 数量经济技术经济研究, 2002(07):25-30.
- [26] 张劲帆, 刚健华, 钱宗鑫, 等. 基于混频向量自回归模型的宏观经济预测[J]. 金融研究, 2018, No. 457(07):38-52.
- [27] 张崇, 吕本富, 彭赓, 等. 网络搜索数据与CPI的相关性研究[J]. 管理科学学报, 2012, 15(7):50-59.
- [28] 张伟, 田金方, 曹灿. 基于混频大数据的宏观经济总量实时预测研究[J]. 宏观经济研究, 2020(2).
- [29] 范辉, 尹余芳, 董依兰. 用于宏观经济预测的VAR、BVAR和VARMA模型的对比[J]. 行政事业资产与财务, 2011, 000(004):8-8.
- [30] Litterman R B. Forecasting with Bayesian vector autoregressions - five years of experience[J]. Journal of Business and Economic Statistics, 1986, 4(1):25-38.
- [31] Chiu C W, Mumtaz H, Pintér G. Forecasting with VAR models: Fat tails and stochastic volatility[J]. International Journal of Forecasting, 2015, 33(4):1124-1143.
- [32] Geweke J. Bayesian treatment of the independent student-t linear model[J]. Journal of Applied Econometrics, 1993, 8(S1):S19-S40.
- [33] Mumtaz H. The Evolving transmission of uncertainty shocks in the United Kingdom[J]. Econometrics, 2016, 4(1):16.
- [34] Norman D, Walker T. Co-movement of Australian state business cycles[J]. Australian Economic Papers, 2007.
- [35] Geweke J, Amisano G. Prediction with misspecified models[J]. The American economic review, 2012, 102(3):482-486.
- [36] Bernanke B S, Boivin J. Monetary policy in a data-rich environment[J]. Journal of Monetary Economics, 2003, 50.

## A Comparative Study on the Forecasting Ability of High-Dimensional Bayesian Vector Autoregressive Models on China's Macroeconomic

Li Yue

(Centre for Economics, Finance and Management Studies, Hunan University, Changsha/Hunan, 410000)

**Abstract:** With the gradual establishment and improvement of China's macroeconomic control system, macroeconomic forecasting has become an important reference for the country when formulating economic policies and regulating the market. It can predict future economic trends by predicting the variable behavior of the economy and the financial system. The emergence of big data in recent years has allowed researchers to apply more information and to treat economic activities as a process that changes over time to make real-time predictions. This improves the prediction results but also brings challenges to the prediction work. In order to improve the accuracy of our country's macroeconomic forecasts and explore the main factors that can improve the forecast accuracy under such increasingly complex conditions, based on theoretical and empirical analysis, this paper uses a large Bayesian vector autoregression (BVAR) model with a specific Kronecker structure in the error covariance to perform fitting and forecasting research on 20 major macroeconomic and financial variables in China. On the one hand, the BVAR model used in this article can split the covariance matrix through the Kronecker structure, thereby speeding up the calculation of a large number of parameters to be estimated in the high-dimensional model under big data; on the other hand, it can be assumed that the variance of the regression equation is time-varying, and it allows the existence of non-Gaussian, common stochastic volatility and sequence-related innovative structures alone or at the same time in the error covariance. The fitting and prediction results of China's macroeconomic data under big data show that models with three innovative covariance structures can improve the accuracy of China's macroeconomic forecasts to a certain extent. In other words, the covariance has a time-varying factor that can better capture the nature of China's macro data. The  $t$  distribution and the common stochastic volatility distribution in the model error covariance structure are the main factors to improve the accuracy of China's macroeconomic forecasts, which can provide greater improvements in both point and density prediction. At the same time, compared with small models, large BVARs with more variable information can better consider the joint information between variables, so as to improve the performance of China's macroeconomic forecasts. At present, there are few literature on forecasting China's macroeconomics through large BVARs with different time-varying error covariance structures. Therefore, this article fills in the blank in this respect, and is helpful to improve the accuracy of macro forecast in China under the environment of big data. Also, it provides some new ideas for the choice of models when studying our country's macroeconomic trends in the future.

**Keywords:** China's macroeconomic forecast; Large BVARs; Heavy-tailed distribution; Common stochastic volatility; Sequence correlation