

层次时间序列预测方法与应用综述

向易¹ 丛丽丽² 王玮鹏² 周晓航²

1 武汉商学院信息工程学院 武汉 430113

2 青岛城市学院工商管理学院 山东 青岛 266106

(xiangyi.ie@wbu.edu.cn)

摘要 层次时间序列是解决具有层级约束的多元时间序列,上层节点的数据是其所有子节点数据的累加。层次时间序列预测的主要难点是在准确预测每个序列的同时,还要保证不同级别之间的一致性,即预测在层次结构中满足累加的约束。随着大规模数据的涌现,这一复杂而具有挑战性的问题展现出更大的研究价值和广泛的应用前景。通过对层次时间序列预测相关方法和文献的综述,从分类方法和应用理论两个方面进行总结和归纳,同时探讨了该技术面临的挑战和实际应用中存在的缺口。分析表明,层次时间序列预测方法主要可分为预测模型和修订模型两个阶段,逐步引入机器学习和深度学习,并演化为将预测和修订模型融合的端对端方法。这些方法广泛应用于商业运营和政府治理领域。在未来的研究趋势方面,首先需要关注海量数据对两阶段方法预测准确度的影响;其次是深入研究端对端层次时间序列预测模型,以避免两阶段参数不连动的问题。此外,政府管理和商业运营的研究可以侧重于对具体问题导致不同层级关注度差异进行建模。

关键词: 层次时间序列预测;一致性;机器学习;深度学习

中图分类号 TP181

Comprehensive Review of Hierarchical Time Series Forecasting Methods and Applications

XIANG Yi¹, CONG Lili², WANG Weipeng² and ZHOU Xiaohang²

1 School of Information Engineering, Wuhan Business University, Wuhan 430113, China

2 School of Management, Qingdao City University, Qingdao, Shandong 266106, China

Abstract Hierarchical time series involve multiple time series with hierarchical constraints, where the data of the upper-level nodes are the cumulative sum of all their child node data. The main challenge in hierarchical time series forecasting is to accurately predict each series while ensuring consistency across different levels, i. e., the forecast must satisfy the additive constraints within the hierarchical structure. With the emergence of large-scale data, this complex and challenging problem has demonstrated greater research value and a broad range of application prospects. This study reviews the literature on hierarchical time series forecasting methods, summarizing and generalizing from the aspects of classification methods and theoretical applications, while also discussing the challenges faced by this technology and the gaps in practical applications. The analysis indicates that hierarchical time series forecasting methods can mainly be divided into two stages, which are forecasting models and revision models, gradually introducing machine learning and deep learning methods, and evolving into end-to-end methods that integrate forecasting and revision models. These methods are widely applied in the fields of business operations and government governance. In terms of future research trends, the first area of focus should be the impact of massive data on the accuracy of two-stage method forecasts, followed by in-depth research on end-to-end hierarchical time series forecasting models to avoid the issue of non-coherent parameters between stages. Additionally, research in government management and business operations can focus on modeling the differences in attention levels at various hierarchical levels caused by specific issues.

Keywords Hierarchical time series forecasting, Consistency, Machine learning, Deep learning

1 引言

在多元时间序列问题的研究领域中,一类自然具有层次约束的多元时间序列被称为层次时间序列,其层次结构中上层的时间序列是下层时间序列的聚合物^[1-2]。这些层次结构可能代表不同时间尺度的汇总数据——如年度、季度、月度和周度数据,或不同地理位置的数据——包括国家、省、地级市

和县区等级别。尽管每个层次都是一个时间序列,但仅对各个层级的时间序列独立进行预测并不能确保预测结果遵循层次约束(也称为聚合约束),即上级时间序列的预测值应为其下级时间序列预测值之和^[3]。零售销售和电力预测是此类层次时间序列预测的典型应用场景。零售预测需要分别追踪不同商店、区域级别和国家级别的销售数据^[4-5],而电力预测需要个人、电网和区域级别的电力消费/生产量^[6-7]。尽管底层

基金项目:教育部人文社会科学基金青年项目(24YJC790248)

This work was supported by the Ministry of Education of Humanities and Social Science Project(24YJC790248).

通信作者:周晓航(xiaohang950510@163.com)

时间序列往往稀疏、含噪且缺乏上层汇总级别的明显模式,但在所有层次上进行准确预测均具有价值。底层预测有助于实施自动化决策,而顶层预测则支持战略性决策制定^[8]。例如,零售商依据销售数据进行店铺级别的预测,而制造商则依据汇总的需求序列预测来安排生产计划^[9]。对于负责配送规划的运输经理而言,需求的空间分布、各配送渠道的货物规模(补货订单)、发货时间及货物类型等信息至关重要。库存经理可能关注于物料类型、数量及需求发生时间的预测^[10]。供应链管理可以根据供应链中各环节的重要性,对总需求进行相应维度的分解。因此,层次时间序列预测模型的应用在精细化管理中发挥着关键作用,使得该领域的研究变得尤为重要。

层次时间序列预测的挑战在于不仅要保证各个时间序列预测的准确性,还要满足时间序列中的加和一致性。此任务的难度不只是简单地融合多种信息,关键在于将明确的层次约束整合进模型训练之中。尽管深度学习模型在各种预测领域展现出卓越性能,但是深度学习是由数据驱动模型,在拟合层次时间序列预测任务时不易学习到其中的层次结构信息。这种忽略可能导致虽然各个序列的预测准确度高,但整体预测结果不满足层次约束,从而失去对层次时间序列分析的实际意义。此外,不同层级的预测结果对决策制定者而言均具备价值:顶层预测有助于策略规划,而底层预测则能提升管理效率^[8]。因此,在层次时间序列预测中,除了捕捉并预测时间序列的历史变化外,预测结果还需满足层次序列间固定关系的约束。

传统的层次时间序列预测方法通常采用两阶段模型:首先对部分或全部时间序列进行预测,随后通过修正模型确保预测结果满足层次约束,即预测结果的加和关系。随着深度学习模型在时间序列预测领域显示出优异性能,越来越多的研究人员尝试将深度学习模型整合进层次时间序列的预测和修正阶段。然而,两阶段过程中预测与修正模型参数的独立学习可能导致信息损失。为了克服这一问题,一些研究者提出了基于深度学习的端到端模型,该模型能够实现预测和修正阶段参数信息的联动,已在市场营销和交通预测等领域取得了显著成效。尽管许多学者致力于层次时间序列预测技术及其应用的研究,但目前对这些研究成果的系统梳理尚显不足。

本文综述了层次时间序列预测领域的研究进展,概括了当前层次时间序列预测技术的研究状况;通过商业运营和政府治理的视角,细致归纳了层次时间序列预测方法的应用情境;进一步,深入分析了在层次时间序列预测任务中遭遇的技术挑战及应用上的缺口,并展望了未来的研究方向与发展趋势。

2 层次时间序列预测的问题描述

当多元时间序列之间存在层次关系,且不同层次的时间序列需满足一定的加和约束时,称其为层次时间序列。图1用树结构描述了一个层次时间序列的层级结构的示例。在这个多元层次时间序列中存在8组时间序列数据 $y_{1:T}^1, y_{1:T}^2, y_{1:T}^3, y_{1:T}^4, y_{1:T}^5, y_{1:T}^6, y_{1:T}^7, y_{1:T}^8$,分别用 $(A), (B_1), (B_2), (C_1), (C_2), (C_3), (C_4), (C_5)$ 表示。

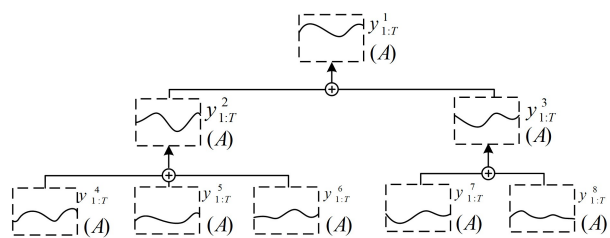


图1 层次时间序列结构示例

Fig. 1 Example of hierarchical time series structure

如图1所示,对于位于非叶节点的每个时间序列,聚合一致性需要满足:

$$y^i = \sum_{j \in C_i} y^j \quad (1)$$

其中, C_i 代表第*i*个时间序列下的所有子节点。式(1)中的所有聚合一致性约束满足:

$$y^A = y^{B_1} + y^{B_2}, y^{B_1} = y^{C_1} + y^{C_2} + y^{C_3}, y^{B_2} = y^{C_4} + y^{C_5} \quad (2)$$

上述的聚合约束可以通过一个统一的公式表示。假设 $Y_t = [y_t^1, \dots, y_t^N] \in \mathfrak{R}^N$,表示*t*时刻所有时间序列的值。可以将 Y_t 分为底部项和聚合项,其中底部项包括叶节点处的时间序列,而聚合项对应于其余节点(即聚合的序列)。使用 $a_t \in \mathfrak{R}^r$ 来表示包含在 Y_t 中的聚合(底部)项,其中 $r(m)$ 是聚合(底部)项的量,有 $Y_t = \begin{bmatrix} a_t \\ b_t \end{bmatrix}$ 和 $N = r + m$ 。然后,按照层次结构,可以根据 b_t 定义一个聚合矩阵 $S \in \{0, 1\}^{r \times m}$,如下所示:

$$S = \begin{bmatrix} s_{\text{sum}} \\ I_m \end{bmatrix}$$

其中, $s_{\text{sum}} \in \{0, 1\}^{r \times m}$ 是一个求和矩阵; I_m 是 $m \times m$ 的单位矩阵,对应于层次结构。最后,聚合约束可以简化为以下向量形式:

$$Y_t = S b_t \Leftrightarrow \begin{bmatrix} a_t \\ b_t \end{bmatrix} = \begin{bmatrix} s_{\text{agg}} \\ I_m \end{bmatrix} b_t$$

以图1中的层次结构为例, $a_t = [y_t^1, y_t^2, y_t^3]^T \in \mathfrak{R}^3$, $b_t = [y_t^4, y_t^5, y_t^6, y_t^7, y_t^8]^T \in \mathfrak{R}^5$ 。聚合矩阵 S 为:

$$S = \begin{bmatrix} s_{\text{agg}} \\ I_m \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ & & & & I_5 \end{bmatrix}$$

此外,通过定义 $A := [I_r | -S_{\text{agg}}] \in \{0, 1\}^{r \times N}$, $Y_t = S b_t$ 可以等价地表示为: $A Y_t = 0$ 。

因此,对层次时间序列预测问题的描述如下:给定 M 步历史时间序列 $Y_t = [y_{1-M:t-1}^1, \dots, y_{1-M:t-1}^N] \in \mathfrak{R}^{N \times M}$ 和层次信息 Γ ,获得最终预测 $\hat{Y}_t \in \mathfrak{R}^N$ 。值得注意的是,预测的时间序列需满足聚和约束。因此,一个预测模型 \mathbb{F} 被用来获得预测 \hat{Y}_t ,如下所示:

$$\hat{Y}_t = \mathbb{F}(Y_t, \Gamma, \theta)$$

其中, \hat{Y}_t 应满足 $A \hat{Y}_t = 0$ 并且接近 Y_t 。 $Y_t(\hat{Y}_t)$ 必须包含所有的时间序列。

3 层次时间序列预测方法

3.1 基于传统层次时间序列预测的两阶段方法

层次时间序列预测方法主要依赖于两阶段的处理策略,

其首先应用传统时间序列预测模型对全体或部分序列进行预测,随后通过修正模型采用聚合或分解等手段以实现整体序列的一致性预测。在此过程中,所使用的传统的时间预测模型包括 AR^[11],ETS^[12]和 ARMA^[12]。预测模型的主要不同,则体现在研究使用了不同的修正模型。典型的修正模型包括 3 种主要方法。

1) 自下而上方法(Bottom-Up Approach):此方法从最基层的时间序列出发进行预测,然后利用各层级之间的关联关系进行聚合,从而实现对整体时间序列的一致性预测。该策略的优势在于能够充分利用底层数据的预测结果,这些结果往往相对准确,因为较低层级的数据通常表现出更高的稳定性和较小的波动。每一层的预测结果均基于底层数据累积而成,这使得对总体预测结果的解释与分析更为直观,但同时也因数据粒度较细而要求更高的数据完整性,预测过程中的噪声可能对建模和预测效果带来较大挑战^[13]。

2) 自上而下方法(Top-Down Approach):该方法着手于顶层时间序列的预测,之后通过历史比例的分解来实现对整体时间序列的一致性预测。从最高层次开始的预测随后向下细化到较低层次,这一策略可以减小计算复杂度,节约时间与资源。然而,历史比例的变化可能随时间推移而变动,从而影响整体预测的准确性^[13]。

3) 中间向外方法(Middle-Out Approach):通过预测中间层级的时间序列,再结合自下而上和自上而下的方法,分别向上聚合及向下分解,以获得整体的一致性预测结果。这种方法集成了自下而上与自上而下两种方法的优点,尽管其实施复杂度相对较高。

历史上,不同学者对这 3 种方法的有效性有着不同的看法和实证研究。例如,Grunfeld 等^[14]早期提出,细分数据易

出错,从而认为自上而下方法可能更准确。Fliedner^[15]也指出,使用聚合级别的数据能够获得更好的聚合预测性能。相对地,Orcutt 等^[16]和 Edwards 等^[17]通过研究认为,在某些条件下,自下而上的方法在聚合过程中信息损失较少,因而可以获得更准确的预测。Schwarzkopf 等^[18]的研究也支持这一观点,指出在存在数据缺失或不可靠情况下,自下而上方法表现更佳。

不管采用以上哪种修订方法,其预测结果往往都会受到聚合或分解阶段中噪声传递的影响。为了克服这一问题,Hyndman 等^[19]在 2011 年提出了一种创新方法,通过单独预测每一个时间序列,然后利用线性回归模型来优化所有时间序列的预测,从而实现对整体时间序列的一致性预测。此方法采用了广义最小二乘估计(Ordinary Least Squares,OLS)技术,并需要估计由不一致性引起的误差的协方差矩阵。通过这种优化方式,层次时间序列预测不再局限于对部分序列的预测,而是能够实现对整体的一致性预测。进一步,Wickramasuriya 等^[20]在 2019 年提出了一种新的预测协调方法(修正方法),称为最小迹(MinT)方法。该方法旨在最小化整个层次结构上一致性预测的均方误差,同时维持无偏性的约束条件。采用 MinT 方法所得到的修正预测在保持一致性和无偏性的同时,还具备了在所有组合预测中最小的方差,这一特点显著提升了预测的准确性和可靠性。此外,Taieb 等在 2019 年提出了一种名为经验风险最小化(Empirical Risk Minimization,ERM)的优化方法。与 MinT 方法相比,ERM 方法放宽了无偏估计的条件,更加贴合实际数据的特征,从而使层次时间序列预测的修正模型更加符合现实世界的复杂性和变动性。这些方法的发展不仅提高了层次时间序列预测的准确度,也为处理实际问题提供了更为灵活和高效的工具。

表 1 传统层次时间序列中的修订模型比较

Fig. 1 Comparison of revision models in traditional hierarchical time series

| 方法 | 部分/全部 | 特点 | 文献 |
|----------------------|--------|------------------------------------|---------------|
| 自上而下 (bottom-up) | 部分时间序列 | 强调对底层细节的精细建模,适用于数据较为细致且有较强异质性的情况 | [14-15,18,22] |
| 自下而上 (top-down) | 部分时间序列 | 强调整体趋势和战略规划,适用于顶层数据的趋势相对较为明显的情况 | [16,18] |
| 中间向外 (Middle-out) | 部分时间序列 | 强调中间层次的协同和整合,平衡了底层的细致建模和顶层的全局规律 | [19] |
| 最小二乘法 (OLS) | 全部时间序列 | 对于线性关系较强的层次时间序列,OLS 可以提供较为简单但有效的修正 | [19] |
| 最小迹 (MinT) | 全部时间序列 | 考虑多个因素和多个指标之间的关系,以更全面地修正预测 | [20] |
| 经验风险最小化 (ERM) | 全部时间序列 | 放松了 MinT 方法的无偏估计条件 | [3] |

3.2 融合机器学习与层次时间序列预测的两阶段方法

鉴于传统时间序列预测技术在处理高维度大数据集以及有效表征复杂函数关系方面的局限性,近年来层次时间序列预测领域开始广泛采用机器学习方法来提高预测的准确性和效率。Abolghasemi 等^[23]在 2019 年首次探索了将机器学习技术应用于层次时间序列预测的潜力。他们通过 XGBoost 模型、支持向量机(SVM)和人工神经网络等方法分别对部分或全部时间序列进行预测,并采用自上而下(Top-down)、自下而上(Bottom-up)、中间向外(Middle-out)以及传统修正方法等不同策略来实现整体的一致性预测。这项研究表明,机器学习方法在原始预测准确度方面的优势显著,能够有效提

升整体预测的一致性。此外,为了解决层次时间序列预测中的一致性问题的研究,研究者开始探索深度学习模型在此领域的应用。Shiratori 等^[24]于 2020 年尝试通过在深度学习模型的目标函数中添加一致性约束的正则项,以引导模型在预测底层时间序列时自然地考虑到一致性问题。最终,这些预测结果通过传统的修订模型进行调整,以获得符合一致性要求的最终预测输出。

在层次时间序列预测领域,如何有效地获得一致性预测结果是一个重要且具有挑战性的问题。近年来的研究趋势表明,越来越多的工作开始集中于利用机器学习方法改进修订模型,以提升预测的一致性和准确性。Burba 等^[25]通过结合自回归模型和 Xgboost 模型生成的原始预测,采取了一种

创新的方法来直接对神经网络结构进行基于层次结构矩阵的裁剪。这种方法使得深度学习模型的预测结果在结构上自然具有一致性。然而,这种方式对模型结构有较强的人为干预。对此,研究者进一步通过引入正则项来使模型自动学习层次约束,从而提高了修订结果的泛化能力。Mishchenko 等^[26]在层次时间序列预测的损失函数中为第一层和第二层的预测结果增加了惩罚项,以促使模型在训练过程中自然而然地保证这两层的一致性。在 Gleason 等^[27]通过将贝叶斯网络的预测结果投影到层次结构空间,并计算全局惩罚,实现了修订模型具有全局一致性约束的预测结果。Han 等^[28]采用深度学习方法获得预测结果,并通过逐层分位数误差的惩罚手段,确保了预测结果具有层次一致性约束。相比之下,大多数研究集中在点估计上,未能有效量化预测风险。为此,Taieb 等^[3]提出了一种创新方法,在预测模型中获取预测值的概率分布参数而非单一均值。通过采样方法生成预测值,并采用自上而下的策略进行聚合,以实现一致性预测,这一方法为量化预测不确定性提供了新的视角。表 2 列出了部分基于机器学习的两阶段层次时间序列预测研究方法。

表 2 基于机器学习的两阶段层次时间序列预测研究

Table 2 Research on two-stage hierarchical time series prediction based on machine learning

| 文献 | 预测模型 | | 修订模型 | |
|------|------|------|------|------|
| | 传统方法 | 机器学习 | 传统方法 | 机器学习 |
| [3] | | ✓ | ✓ | |
| [23] | | ✓ | ✓ | |
| [26] | | ✓ | | ✓ |
| [27] | | ✓ | | ✓ |
| [24] | | ✓ | ✓ | |
| [25] | ✓ | ✓ | | ✓ |
| [26] | | ✓ | | ✓ |

3.3 基于深度学习端对端的层次时间序列预测的两阶段方法

随着层次时间序列预测方法的不断创新和发展,研究的最终目标始终是在保持预测结果一致性的同时,提高预测的准确率。在层次时间序列预测中,两阶段方法由于预测和修订模型之间的参数独立性,预测模型无法充分利用层级结构信息,而修订模型又无法有效捕捉数据变化模式。为了解决这一问题,近年来,研究者开始探索通过端对端的深度学习模型直接在层次时间序列预测中进行应用,旨在同时提升预测的准确性和结果的一致性。端对端的层次时间序列预测模型因能够更好地融合模型参数与层次结构信息,逐渐成为研究的热点。

Rangapuram 等^[29]提出的 E2Ehier 模型便是这一研究方向的代表之一。该模型通过深度学习直接获取每个时间序列的概率参数,并将采样得到的预测值输入到一个特定的投影算子中,以获得符合一致性要求的预测结果,实现了层次时间序列预测的端对端建模。在多个公开数据集的测试中,E2Ehier 模型相较于两阶段方法显示出更优的预测效果。尽管如此,E2Ehier 模型在网络结构中缺少直接的层次约束,这是其需要进一步改进的地方。另一方面,考虑到现实世界中的数据往往不满足高斯分布的假设,Wang 等^[30]提出利用正规化流(Normalizing Flows,NF)的方法来改善基础预测的准确度。正规化流能够通过变换使得复杂分布的数据适应模型处理,从而提高预测的准确性。此外,针对深度学习模型的数

据驱动特性,Xiang 等^[31]提出了将自上而下的方法融入模型的创新策略,通过引入 top-down 功能层并嵌入到模型中,使得在训练参数的同时考虑到层次约束。实验结果表明,通过增强层次结构的约束,能够有效提升预测的效果。

4 层次时间序列预测方法的应用

4.1 商业运营

层次时间序列预测,对于满足复杂组织结构中的需求预测、层级管理和决策协调至关重要。随着技术的发展,研究者们通过各种实证研究,探索了不同方法在实际应用中的效果,以寻找更有效的预测模型。层次预测法被广泛应用于产品需求的预测中,对企业的供应链管理起到了至关重要的作用。Rostami-Tabar 等^[32]在比较自上而下(Top-Down,TD)与自下而上(Bottom-Up,BU)的方法时,发现在非稳态过程中,自下而上方法相对于稳态过程具有更高的优越性。这项研究通过分析欧洲杂货店 1798 个 SKU 在 103 周内的销售预测数据,展示了在动态变化的市场环境中,底层向上方法能更准确地捕捉需求变化。

Mircetic 等^[33]的研究则聚焦于一家主要的欧洲啤酒厂,他们发现修订后的预测方法能够提供比基础预测更优的结果。特别地,他们提出了一种方法,将不同层次的预测方法进行组合,使用自底向上方法来构建其余层次的预测。这种策略的优势在于它消除了选择单一层次方法的需求,并相比于协调方法获得了轻微的优势。Villegas^[34]提出了一种在状态空间公式中包含横截面层次协调的方法,并将其与预测模型结合。他们通过模拟研究和对西班牙一家杂货零售商的实证分析发现,虽然修订后的预测在短期(1-3 天)内表现良好,但在较长期(4-7 天)内的表现并未达到预期。除此之外,Abolghasemi 等^[35]、Abolghasemi 等^[36]以及 Spiliotis 等^[37]的研究贡献也不容忽视。他们分别在澳大利亚食品制造商的数据上应用层次预测,进一步证明了层次时间序列预测方法在不同行业 and 市场中进行需求预测的适用性和有效性。

层次时间序列预测作为一种强大的预测工具,在旅游业、能源领域以及交通领域等多个行业中发挥着至关重要的作用。这种方法能够处理和分析从微观到宏观不同层次的数据,为决策提供支持,优化资源分配,并提高管理效率。在旅游业中,层次时间序列预测被用于估计游客流量,帮助旅游经营者和地方政府做出更加有效的营销决策和资源规划。Karmy 等^[38]的研究表明,自底向上的方法在预测旅游零售行业游客流量方面表现出色。这一发现对于旅游目的地管理和旅游零售业的库存管理具有实际意义,尤其是在旅游旺季期间,准确的游客流量预测可以显著提升顾客满意度和销售业绩。在能源领域,准确预测电力和石油等战略性产品的需求对于保障能源安全和优化能源分配至关重要。层次时间序列预测方法能够综合考虑各种因素对能源需求的影响,提高预测准确性。例如,Mancuso 等^[13]利用基于深度学习的端对端层次时间序列预测模型,成功提升了对电力需求的预测能力,这对于电网运营商在确保电网稳定性和规划未来能源投资方面具有重要价值。在交通领域,层次时间序列预测对于客流量管理、交通规划和拥堵缓解等具有重要作用。通过准确预测不同时间段内各个交通节点的客流量,交通管理部门可以更加有效地进行交通调度和资源分配。此外,层次时间序列预

测还能帮助公共交通运营商优化运力安排,有效助力智慧城市建设。Benarmas 等通过探索道路网络的层次结构,提出了一种用于交通流量分层预测的深度学习方法^[39],首先通过 N-beasts 神经网络获得道路的交通流量总预测,再基于自顶向下的思想,利用卷积神经网络和递归神经网络向下细分,获得所有道路流量预测。

4.2 政府治理

宏观经济学中对层次时间序列预测的运用体现了这一方法在理解和预测总体经济现象方面的重要性。GDP 和 CPI 作为宏观经济的关键指标,其预测准确性直接影响到政策制定者和经济分析师对经济健康状况的判断和预测。GDP 被构建为个体组成部分的总和,支出法将 GDP 视为以消费、投资、进口和出口为特征的支出总和。针对 GDP 预测,将汇集各经济部门与经济活动所产生的基础数据,并按照层次时间序列预测方法进行整合;通过修正模型,使中下层节点的数据与顶层节点(GDP 总量)保持一致与协调。Athanasopoulos 等^[40]、Bisaglia、Di Fonzo 和 Girolimetto^[41]以及 Di Fonzo 和 Girolimetto^[42]分别在点估计和概率估计的两种情境下,都发现了支持的证据。这些研究表明,在经过修订模型后,中下层节点的原始预测对顶层节点的预测准确度有显著提升,高于直接获取顶层节点预测的情况。通过比较不同修正模型的效果发现,MinT 方法在层次时间序列上的整体效果最佳,而当注意力仅限于顶层系列时,加权最小二乘法更为准确。另一个具有层次结构的重要宏观经济变量是消费者价格指数(CPI)。CPI 被构建为一篮子商品的加权价格指数。因此,层次约束矩阵 S 不再仅仅由 1 和 0 组成, S 矩阵中可以包含权重,且可以随时间变化。Weiss^[43]在对英国 CPI 进行预测时发现,短期的一个月预测用 OLS 方法修订好于传统方法;但是对于中长期的 CPI 预测,中间向外则是更好的优化方法。Weiss^[43]还考虑了通货膨胀,发现自下而上的效果更好。

宏观经济学中层次结构的存在以及数据集构建的复杂性,为层次时间序列预测方法的进一步研究和应用提供了新的方向。随着经济数据获取技术的进步和统计方法的发展,对修订模型的探索需要不断适应经济变量特性的变化和预测需求的细化。尤其是在处理具有时间变化权重的层次结构数据时,预测模型需要更加灵活和精确,以反映宏观经济动态和结构的实际变化。

5 层次时间序列预测的挑战和趋势

在当前快速变化和竞争激烈的社会环境中,组织面临着诸多不确定性和挑战。为了在这种环境中生存并取得成功,组织必须具备高度的灵活性和适应性,能够迅速响应环境变化和市场需求的调整。在这个背景下,层次时间序列预测作为精细化管理的支撑工具,为组织提供了一个强大的手段来更好地预测未来趋势和应对潜在风险。

5.1 现存问题和挑战

首先,随着数字化的迅速发展,海量数据的涌现使得时间序列数据的复杂性和噪声程度显著增加,进而给层次时间序列预测模型的预测和修订带来了更大的挑战。众所周知,层次时间序列预测的核心首先在于原始预测结果,更精确的预测能够降低一致性模型拟合的难度。在大数据时代,如何

有效地利用丰富的数据资源,从时间序列的历史数据中识别复杂的变化规律成为了亟待解决的问题。其次,对修订模型的拟合难度也随之增加。大数据时代下,层次结构更为精细,复杂度更高,这给从原始预测中获得一致性预测带来了巨大挑战。此外,最明显的是计算资源的需求。层次时间序列预测的一个显著特征是需要满足层次约束,常见的修订模型都需要使用层次约束矩阵来确保预测结果的一致性。然而,复杂的层次结构使得层次约束矩阵的计算变得更加困难,从而影响了现有层次时间序列预测的准确性。由于矩阵规模较大,计算机需要更多时间来执行矩阵运算,同时更高的存储需求可能导致内存不足或频繁的数据交换,进而降低了模型的运行效率。

在当前数据量日益庞大的情况下,两阶段方法所导致的参数不连动问题依然存在,并且在误差扩散方面可能表现得更加明显。尽管针对修订模型的研究已经显而易见地突破了单层方法的局限性,避免了仅仅使用单个层次的预测来实现整体预测,但两阶段方法由于预测和修订模型之间的参数独立性,预测模型无法充分利用层级结构信息。基于深度学习的端对端的层次时间序列预测模型能够有效地弥补两阶段方法的不足之处。当预测模型和修订模型融合在一起时,模型的参数训练可以同时考虑到层次约束。但基于深度学习的端对端层次时间序列预测模型面临的挑战是如何对多元时间序列之间的层次约束进行建模。鉴于机器学习方法在当前已经可以很好地拟合历史的变化规律,如何在当前模型中有效地建模层次约束显得十分关键,这将极大地影响端对端层次时间序列预测研究的发展方向。

5.2 未来研究趋势

尽管现有方法已经在提高层次时间序列预测效果方面取得显著进步,但在面对更加复杂的管理需求和大数据场景时,进一步提高预测的准确性仍将极大地提升管理效能和降低决策风险,这也为层次时间序列的预测技术的发展指明了方向。

1) 层次时间序列预测的两阶段方法研究将还是针对修订模型进行优化。面对越来越多的大数据加速计算效率和优化存储效果,将是重点关注的方向。2) 在端对端模型中有效地表示或实施层次约束,仍然是一个需要进一步探索的新问题。引入针对层次约束的模块在机器学习/深度学习的预测模型中是另一个趋势,这将有助于提高模型对复杂数据结构的理解和处理能力。3) 1 针对市场运营和政府管理,都将以概率估计为主,概率预测不仅提供了对未来事件的点估计,还提供了与估计相关的不确定性信息。

结束语 本文关于层次时间序列预测方法与应用文献综述,涵盖了层次时间序列的基本理论、预测方法的历史发展以及应用场景这 3 个方面。尽管层次时间序列预测技术已经取得了显著进展,但在提高预测准确度的同时保证一致性仍然是这一任务的最大理论挑战,而形成端对端的预测结果则是实际应用中亟待解决的问题。因此,针对层次时间序列预测任务的研究具有重要的理论价值和现实意义。本文全面系统地回顾层次时间序列预测方法的工作,将极大地推动层次时间序列预测任务的发展。

参 考 文 献

[1] PETROPOULOS F, APILETTI D, ASSIMAKOPOULOS V,

- et al. Forecasting: Theory and practice[J]. *International Journal of Forecasting*, 2022, 38(3): 705-871.
- [2] ZHOU F, PAN C, MA L, et al. SLOTH: Structured learning and task-based optimization for time series forecasting on hierarchies [C]// *Proceedings of the AAAI Conference on Artificial Intelligence*. 2023: 11417-11425.
- [3] TAIEB S B, TAYLOR J W, HYNDMAN R J. Coherent probabilistic forecasts for hierarchical time series[C]// *International Conference on Machine Learning*. 2017: 3348-3357.
- [4] SEEGER M, RANGAPURAM S, WANG Y, et al. Approximate Bayesian inference in linear state space models for intermittent demand forecasting at scale[J]. *arXiv:1709.07638*, 2017.
- [5] SEEGER M, SALINAS D, FLUNKERT V. Bayesian intermittent demand forecasting for large inventories[C]// *Proceedings of the 30th International Conference on Neural Information Processing Systems*. 2016: 4653-4661.
- [6] JEON J, PANAGIOTELIS A, PETROPOULOS F. Probabilistic forecast reconciliation with applications to wind power and electric load[J]. *European Journal of Operational Research*, 2019, 279(2): 364-379.
- [7] TAIEB S B, TAYLOR J W, HYNDMAN R J. Hierarchical probabilistic forecasting of electricity demand with smart meter data[J]. *Journal of the American Statistical Association*, 2021, 116(533): 27-43.
- [8] JANUSCHOWSKI T, KOLASSA S. 3. 2 A classification of business forecasting problems[M]// *Business Forecasting: Practical Problems and Solutions*. Germany: Wiley, 2016: 171.
- [9] CHOPRA S, MEINDL P. Supply chain management. Strategy, planning & operation[M]. Springer, 2007.
- [10] CAPLICE C G, SHEFFI Y. ESD. 260J/1. 260J/15. 770J Logistics Systems, Fall 2003 [M/OL]. Massachusetts Institute of Technology, 2003.
- [11] BOX G E, JENKINS G M, REINSEL G C, LJUNG G M. Time series analysis: Forecasting and control [M]. John Wiley & Sons, 2015.
- [12] MONTGOMERY D C, JENNINGS C L, KULAHCI M. Introduction to time series analysis and forecasting[M]. John Wiley & Sons, 2015.
- [13] MANCUSO P, PICCIALI V, SUDOSO A M. A machine learning approach for forecasting hierarchical time series[J]. *Expert Systems with Applications*, 2021, 182: 115102.
- [14] GRUNFELD Y, GRILICHES Z. Is aggregation necessarily bad? [J]. *The Review of Economics and Statistics*, 1960, 42(1): 1-13.
- [15] FLIEDNER G. An investigation of aggregate variable time series forecast strategies with specific subaggregate time series statistical correlation[J]. *Computers & Operations Research*, 1999, 26(10-11): 1133-1149.
- [16] ORCUTT G H, WATTS H W, EDWARDS J B. Data aggregation and information loss[J]. *The American Economic Review*, 1968, 58(4): 773-787.
- [17] EDWARDS J B, ORCUTT G H. Should aggregation prior to estimation be the rule? [J]. *The Review of Economics and Statistics*, 1969, 51(4): 409-420.
- [18] SCHWARZKOPF A B, TERSINE R J, MORRIS J S. Top-down versus bottom-up forecasting strategies[J]. *The International Journal of Production Research*, 1988, 26(11): 1833-1843.
- [19] HYNDMAN R J, AHMED R A, ATHANASOPOULOS G, et al. Optimal combination forecasts for hierarchical time series [J]. *Computational Statistics & Data Analysis*, 2011, 55(9): 2579-2589.
- [20] WICKRAMASURIYA S L, ATHANASOPOULOS G, HYNDMAN R J. Optimal forecast reconciliation for hierarchical and grouped time series through trace minimization[J]. *Journal of the American Statistical Association*, 2019, 114(526): 804-819.
- [21] TAIEB S B, KOO B. Regularized regression for hierarchical forecasting without unbiasedness conditions[C]// *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 2019: 1337-1347.
- [22] SHLIFER E, WOLFF R W. Aggregation and proration in forecasting[J]. *Management Science*, 1979, 25(6): 594-603.
- [23] ABOLGHASEMI M, HYNDMAN R J, TARR G, et al. Machine learning applications in time series hierarchical forecasting[J]. *arXiv preprint arXiv:1912.00370*, 2019.
- [24] SHIRATORI T, KOBAYASHI K, TAKANO Y. Prediction of hierarchical time series using structured regularization and its application to artificial neural networks [J]. *Plos One*, 2020, 15(11): e0242099.
- [25] BURBA D, CHEN T. A trainable reconciliation method for hierarchical time-series[J]. *arXiv preprint arXiv:2101.01329*, 2021.
- [26] MISHCHENKO K, MONTGOMERY M, VAGGI F. A self-supervised approach to hierarchical forecasting with applications to groupwise synthetic controls[J]. *arXiv:1906.10586*, 2019.
- [27] GLEASON J L. Forecasting hierarchical time series with a regularized embedding space[C]// *MileTS20: 6th KDD Workshop on Mining and Learning from Time Series*. 2020: 883-894.
- [28] HAN X, DASGUPTA S, GHOSH J. Simultaneously reconciled quantile forecasting of hierarchically related time series[C]// *International Conference on Artificial Intelligence and Statistics*. 2021: 190-198.
- [29] RANGAPURAM S S, WERNER L D, BENIDIS K, et al. End-to-end learning of coherent probabilistic forecasts for hierarchical time series[C]// *International Conference on Machine Learning*. PMLR, 2021: 8832-8843.
- [30] WANG S, ZHOU F, SUN Y, et al. End-to-end modeling of hierarchical time series using autoregressive transformer and conditional normalizing flow-based reconciliation[C]// *2022 IEEE International Conference on Data Mining Workshops (ICDMW)*. 2022: 1087-1094.
- [31] XIANG Y, SUN H, TU W. HDResNet: Hierarchical-decomposition residual network for hierarchical time series forecasting [C]// *2023 International Joint Conference on Neural Networks (IJCNN)*. 2023: 1-8.
- [32] ROSTAMI-TABAR B, BABAI M Z, DUCQ Y, et al. Non-stationary demand forecasting by cross-sectional aggregation[J]. *International Journal of Production Economics*, 2015, 170: 297-309.
- [33] MIRCETIC D, ROSTAMI-TABAR B, NIKOLICIC S, et al. Forecasting hierarchical time series in supply chains: An empirical investigation [J]. *International Journal of Production Research*, 2022, 60(8): 2514-2533.

- [34] VILLEGAS M A, PEDREGAL D J. Supply chain decision support systems based on a novel hierarchical forecasting approach [J]. *Decision Support Systems*, 2018, 114: 29-36.
- [35] ABOLGHASEMI M, HYNDMAN R J, SPILIOTIS E, et al. Model selection in reconciling hierarchical time series [J]. *Machine Learning*, 2022, 111: 739-789.
- [36] ABOLGHASEMI M, TARR G, BERGMEIR C. Machine learning applications in hierarchical time series forecasting: Investigating the impact of promotions [J]. *International Journal of Forecasting*, 2024, 40(2): 597-615.
- [37] SPILIOTIS E, ABOLGHASEMI M, HYNDMAN R J, et al. Hierarchical forecast reconciliation with machine learning [J]. *Applied Soft Computing*, 2021, 112: 107756.
- [38] KARMIY J P, MALDONADO S. Hierarchical time series forecasting via support vector regression in the European travel retail industry [J]. *Expert Systems with Applications*, 2019, 137: 59-73.
- [39] BENABDALLAH BENARMAS R, BEGHDAD BEY K. A deep learning hierarchical approach to road traffic forecasting [J]. *Journal of Forecasting*, 2024, 43(5): 1294-1307.
- [40] ATHANASOPOULOS G, GAMAKUMARA P, PANAGIOTELIS A, et al. Hierarchical forecasting [M] // *Macroeconomic Forecasting in the Era of Big Data*. *Advanced Studies in Theoretical and Applied Econometrics*. 2020, 52: 689-719.
- [41] BISAGLIA L, DI FONZO T, GIROLIMETTO D. Fully reconciled GDP forecasts from income and expenditure sides [R]. arXiv: 2004. 03864, 2020.
- [42] REDELICO F O, PROTO A N, AUSLOOS M. Hierarchical structures in the Gross Domestic Product per capita fluctuation in Latin American countries [J]. *Physica A: Statistical Mechanics and its Applications*, 2009, 388(17): 3527-3535.
- [43] WEISS C. *Essays in hierarchical time series forecasting and forecast combination [D]*. University of Cambridge, 2018.



XIANG Yi, born in 1993, Ph. D. His main research interests include deep learning and hierarchical time series forecasting.



ZHOU Xiaohang, born in 1995, Ph. D. Her main research interests include information management, e-commerce and business intelligence.