

# 一种用于高阶数据预测的张量自回归方法研究\*

李昂, 刘金杰

(重庆师范大学 重庆国家应用数学中心, 重庆 401331)

**摘要:**通过张量方法为高阶高维时间序列构建一种新的回归模型。基于张量奇异值分解,在张量T-积下利用管秩为张量时间序列预测问题提出了一种新型的低秩自回归模型,并采用交替极小化算法进行模型参数估计。数值实验结果显示,该模型在预测精度和计算时间方面均表现出一定优势。该结果验证了这个低秩自回归模型的可行性与有效性。

**关键词:**张量奇异值分解;自回归模型;张量T-积;管秩;交替极小化算法

**中图分类号:**O242.2;O175.2

**文献标志码:**A

**文章编号:**1672-6693(2025)05-0020-09

## 1 研究背景

回归分析是数据分析领域中的核心工具,主要用于分析2个或多个变量之间的相关性、方向、强度等关系,并通过建立数学模型来预测研究者感兴趣的变量。传统回归分析方法已经形成了多个研究分支,包括广义线性回归<sup>[1]</sup>、偏最小二乘法<sup>[2]</sup>、高斯过程法<sup>[3]</sup>、随机森林法<sup>[4]</sup>、支持向量回归法<sup>[5]</sup>以及神经网络<sup>[6-7]</sup>等。这些方法通常采用向量或矩阵形式对回归系数建模,以刻画输入变量与目标变量之间的关系。然而,随着互联网和数据技术的发展,现代数据以更复杂的形式呈现,超越了传统向量和矩阵的表达方式。面对这种多维数据结构,传统回归分析方法在提取有效信息方面逐渐显得力不从心。张量(高维数组)的引入,为处理高阶高维数据提供了自然的扩展路径。近些年来,基于张量结构的回归模型在理论与应用方面获得了很大的进展,已成为大规模数据分析的重要工具,并在图像分析<sup>[8]</sup>、时空数据分析<sup>[9]</sup>、机器学习<sup>[10-12]</sup>等领域得到了广泛使用。

相较于传统的向量或矩阵数据结构,张量能够更有效地保留数据的潜在关系,从而更准确地表达多维信息,在数据存储和处理中展现出独特的优势。现实中,许多领域(如工程学、经济学等)的观测数据往往以高阶高维结构呈现,尤其是与时间点密切相关的复杂数据,通常就可以自然地表示为张量形式。例如,按时间段记录的网络交通流量数据中不同终端对之间的交互关系,即可构成1个三阶张量时间序列。然而,传统回归模型往往仅适用于低维数据,难以充分利用高阶数据中蕴含的结构信息,因此限制了高阶数据在复杂时间序列建模中的表现。为解决此问题,Rogers等人<sup>[13]</sup>针对张量时间序列首次提出了高阶预测模型:多重线性动态系统(multilinear dynamical systems, MLDS),该模型通过追踪潜在的张量状态和观测到的张量序列来实现预测,为这一领域提供一种可行的解决方案。

近些年来,张量分解作为一种从张量数据中提取关键信息的强大工具受到了研究者的广泛关注。它包括CP(candecomp/parafac)分解<sup>[14]</sup>、Tucker分解<sup>[15]</sup>、张量火车(tensor train, TT)分解<sup>[16]</sup>等诸多方法。特别地,在2011年,Kilmer等人<sup>[17]</sup>通过将张量转化为相应的管矩阵,并在管矩阵上进行分解,提出了张量奇异值分解(tensor singular value decomposition, T-SVD)。该方法不仅可以有效地降低计算量,还可以充分保留管纤维之间的信息数据。张量奇异值分解方法一经提出,便迅速在众多领域得到了广泛应用。随着张量分解与张量运算技术的不断发展,研究者们提出了多种基于不同分解方法的自回归模型,进一步推动高阶时间序列预测的研究

\* 收稿日期:2025-03-04 修回日期:2025-04-23 网络出版时间:2025-05-16T06:59

资助项目:重庆市英才计划“包干制”项目(No. cstc2022ycjh-bgzxm0040);重庆市教育委员会科学技术研究计划重点项目(No. KJZD-K202200506);重庆师范大学博士启动基金、人才引进项目(No. 22XLB005);重庆市科学技术局科学研究一般项目(No. ncamc2022-msxm02)

第一作者简介:李昂,男,研究方向为应用数学,E-mail:liang\_strive2022@163.com;通信作者:刘金杰,女,副教授,博士,E-mail:jinnie.liu@cqnu.edu.cn

网络出版地址:https://link.cnki.net/urlid/50.1165.N.20250507.1635.006

进展。2019 年, Gao 等人<sup>[18]</sup>基于张量核范数(tensor nuclear norm, TNN), 借助张量奇异值分解, 建立张量线性回归模型进行数据分析。2021 年, Chen 等人<sup>[19]</sup>提出保留矩阵时间序列结构的自回归模型(autoressive models for matrix-valued time series, MAR)。相较于一般的向量自回归模型(vector autoregressive modeling, VAR)<sup>[20]</sup>, MAR 模型可以在保留数据结构的同时拥有 VAR 模型的优点。在此基础上, Li 等人<sup>[21]</sup>通过张量 Tucker-分解在张量模态积的定义下结合 MAR 模型, 给出了一种新的张量自回归模型(tensor autoregressive models, TenAR), 并在不同方法下的自回归参数估计下, 该模型能直接得到和原数据相同结构的预测值。同年, Hill 等人<sup>[22]</sup>在张量积的基础上提出具有鲁棒性的张量自回归模型, 并证明该模型在计算参数估计上具有很大的优势。2022 年, Wang 等人<sup>[23]</sup>在减秩思想的基础上, 利用张量 Tucker-分解实现同时沿 3 个方向限制参数空间, 从而达到自回归参数的大规模降维。同年, Cates 等人<sup>[24]</sup>则通过结合自回归模型以及张量分解提出了 L-变换下的张量自回归模型(L-TAR), 并使用最小二乘法完成模型的参数估计。2024 年, Wang 等人<sup>[25]</sup>通过截断的张量奇异值分解建立了低秩张量自回归预测器。

值得注意的是, 通过张量奇异值分解定义的张量管秩总是少于通过 Tucker 分解定义的 Tucker 秩, 这使得基于张量奇异值分解的模型在计算效率上通常优于基于 Tucker 分解的模型。受此启发, 本文在张量 T-积的定义框架下, 基于张量管秩, 将传统的自回归模型与张量奇异值分解相结合, 在保证不破坏原有张量数据结构的情况下对它进行回归建模, 进而提出一种新型的低秩张量自回归模型。本文还设计了基于交替极小化方法的求解算法, 并通过在合成数据和真实时间序列数据集上的实验验证所提方法的有效性和优越性。

## 2 符号说明与预备知识

在本文中, 小写字母  $a, b, c, \dots$  表示标量, 加粗小写字母  $\mathbf{a}, \mathbf{b}, \mathbf{c}, \dots$  以及加粗大写字母  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \dots$  分别表示向量和矩阵, 加粗大写花体字母  $\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$  表示张量。实数域记为  $\mathbf{R}$ , 复数域记为  $\mathbf{C}$ , 单位矩阵为  $\mathbf{I}$ 。对于任意的正整数  $n$ , 定义  $[n] := \{1, 2, \dots, n\}$ 。

对于一个三阶张量  $\mathcal{A} \in \mathbf{C}^{m \times n \times p}$ ,  $m, p$  均为正整数,  $(i, j, k)$  位置的元素表示为  $\mathcal{A}_{ijk}$ 。对张量  $\mathcal{A}$  进行正面切片的第  $k$  片表示为  $\mathcal{A}(:, :, k)$ 。令管标量  $\mathbf{a}_{ij} \in \mathbf{C}^{1 \times 1 \times p}$  表示张量  $\mathcal{A}$  的第  $(i, j)$  个管纤维, 即  $\mathbf{a}_{ij} = \mathcal{A}(i, j, :)$ , 其中  $i \in [m], j \in [n]$ 。此外, 令  $\mathbf{a}_{ij}(k)$  表示管纤维  $\mathbf{a}_{ij}$  的第  $k$  个元素, 其中  $k \in [p]$ 。相应地, 根据管纤维的相关表示方法, 并引入记号  $\mathbf{C}_p^{m \times n}$  代替  $\mathbf{C}^{m \times n \times p}$ , 可以将三阶张量  $\mathcal{A} \in \mathbf{C}^{m \times n \times p}$  视为一个由长度为  $p$  的管标量构成的  $m \times n$  管矩阵, 记作  $\mathcal{A} = (\mathbf{a}_{ij}) \in \mathbf{C}_p^{m \times n}$ 。在此设定下, 张量  $\mathcal{A} \in \mathbf{C}_p^{m \times n}$  的第  $k$  个正面切片可表示为  $\mathbf{A}_{(k)} = (\mathbf{a}_{ij}(k)) \in \mathbf{C}^{m \times n}$ ,  $\mathbf{A}_{(k)}$  恰好是一个  $m \times n$  矩阵。任意 2 个管矩阵  $\mathcal{A} = (\mathbf{a}_{ij}), \mathcal{B} = (\mathbf{b}_{ij}), \mathcal{A}, \mathcal{B} \in \mathbf{C}_p^{m \times n}$ , 定义内积为  $\langle \mathcal{A}, \mathcal{B} \rangle = \sum_{i=1}^m \sum_{j=1}^n \langle \mathbf{a}_{ij}, \mathbf{b}_{ij} \rangle$ 。 $\mathcal{A}$  的 Forbenious 范数定义为  $\|\mathcal{A}\|_F = \sqrt{\langle \mathcal{A}, \mathcal{A} \rangle}$ 。对于 2 个给定的矩阵  $\mathbf{A}, \mathbf{B} \in \mathbf{C}^{m \times n}$ , 内积表示为  $\langle \mathbf{A}, \mathbf{B} \rangle = \text{tr}(\mathbf{A}^H \mathbf{B})$ ,  $\text{tr}$  为矩阵的迹,  $\mathbf{A}^H$  为  $\mathbf{A}$  的共轭转置。对于矩阵  $\mathbf{A} \in \mathbf{C}^{m \times n}$ , 将它的核范数表示为  $\|\mathbf{A}\|_* = \sum_{i=1}^{\min\{m, n\}} \sigma_i(\mathbf{A})$ ,  $\sigma_{\max}(\mathbf{A})$  和  $\sigma_{\min}(\mathbf{A})$  分别表示矩阵  $\mathbf{A}$  的最大奇异值和最小奇异值。

对任意给定的三阶张量  $\mathcal{C} = (\mathbf{c}_{il}) \in \mathbf{C}_p^{m \times n}$  以及矩阵  $\mathbf{D} = (d_{ij}) \in \mathbf{C}^{m \times s}$ , 定义张量和矩阵运算  $\mathcal{G} = \mathcal{C} \odot \mathbf{D} = (\sum_{l=1}^n \mathbf{c}_{il} d_{lj})$ 。可以注意到  $\mathcal{G}$  的第  $k$  个正面切片  $\mathcal{G}_{(k)}$  由  $\mathbf{C}_{(k)} \mathbf{D}$  得到, 这意味着:

$$\|\mathcal{G}\|_F = \sqrt{\sum_{k=1}^p \|\mathcal{G}_{(k)}\|_F^2} \leq \sqrt{\sum_{k=1}^p \|\mathbf{C}_{(k)}\|_F^2 \|\mathbf{D}\|_2^2} = \|\mathcal{C}\|_F \|\mathbf{D}\|_2.$$

对于任意矩阵  $\mathbf{C} \in \mathbf{C}^{m \times n}, \mathbf{D} \in \mathbf{C}^{m \times s}$  均有  $\|\mathbf{C} \mathbf{D}\|_F \leq \|\mathbf{C}\|_2 \|\mathbf{D}\|_F$ 。类似地, 对于任意张量  $\mathcal{C} = (\mathbf{c}_{ij}) \in \mathbf{C}_p^{m \times n}$  以及矩阵  $\mathbf{D} = (d_{il}) \in \mathbf{C}^{s \times m}$ , 定义  $\mathcal{G} = \mathbf{D} \odot \mathcal{C} = (\sum_{l=1}^n d_{il} \mathbf{c}_{lj})$ 。同样有  $\|\mathcal{G}\|_F \leq \|\mathcal{C}\|_F \|\mathbf{D}\|_2$ 。

$\mathbf{Q} \in \mathbf{C}^{p \times p}$  为一个正交阵。对于任意给定的向量  $\mathbf{a} \in \mathbf{C}^p$ , 定义映射  $\varphi_{\mathbf{Q}}: \mathbf{C}^p \rightarrow \mathbf{C}^p$ , 在该映射下的像  $\bar{\mathbf{a}} := \varphi_{\mathbf{Q}}(\mathbf{a}) = \mathbf{Q} \mathbf{a}$ 。定义该映射的逆映射为  $\varphi_{\mathbf{Q}}^{-1}: \mathbf{C}^p \rightarrow \mathbf{C}^p$ , 逆映射下的像为  $\varphi_{\mathbf{Q}}^{-1}(\mathbf{a}) = \mathbf{Q}^{-1} \mathbf{a}$ 。任意给定  $\mathbf{a}, \mathbf{b} \in \mathbf{C}^p$ , 定义在矩阵  $\mathbf{Q}$  下的张量积运算为  $\mathbf{a} \odot_{\mathbf{Q}} \mathbf{b} = \varphi_{\mathbf{Q}}^{-1}(\varphi_{\mathbf{Q}}(\mathbf{a}) \circ \varphi_{\mathbf{Q}}(\mathbf{b}))$ 。其中  $\circ$  表示向量的 Hadamard 积。为了方便叙述, 下文的表述中将省略  $\mathbf{Q}$ , 用  $\mathbf{a} \odot \mathbf{b}$  表示 2 个向量在  $\mathbf{Q}$  下的张量 T-积。

**定义 1**<sup>[26]</sup> ( $\mathbf{Q}$  变换下的张量 T-积) 给定正交阵  $\mathbf{Q} \in \mathbf{C}^{p \times p}$ , 以及张量  $\mathcal{A} = (\mathbf{a}_{il}) \in \mathbf{C}_p^{m \times s}, \mathcal{B} = (\mathbf{b}_{ij}) \in \mathbf{C}_p^{s \times n}$ , 张

量  $\mathcal{A}$  与张量  $\mathcal{B}$  的积为张量  $\mathcal{C} \equiv (c_{ij}) \in \mathbf{C}_p^{m \times n}$ , 记为  $\mathcal{A} \otimes \mathcal{B}$ 。其中:

$$c_{ij} = \sum_{l=1}^s a_{il} \odot b_{lj}, i \in [m], j \in [n].$$

这种变换后的张量 T-积运算是通过将张量  $\mathcal{A}$  与张量  $\mathcal{B}$  等效地转化为管矩阵进行运算, 在此定义下可以保留矩阵乘法, 但需要将标量乘法替换为处理后的管运算。如果矩阵  $Q$  为快速傅里叶变换矩阵, 则  $\mathcal{A} \otimes \mathcal{B}$  退化为 Kilmer 等人<sup>[17]</sup> 提出的张量 T-积。文献[24]提出的余弦变换张量积也是该张量 T-积定义下的特殊例子。

若  $Q \in \mathbf{C}^{p \times p}$  为正交阵, 对  $\mathcal{A} = (a_{ij}), \mathcal{B} = (b_{ij}), \mathcal{A}, \mathcal{B} \in \mathbf{C}_p^{m \times n}$ , 定义运算  $\overline{\mathcal{A}} := \varphi_Q(\mathcal{A}) = \varphi_Q(a_{ij}), \overline{\mathcal{B}} := \varphi_Q(\mathcal{B}) = \varphi_Q(b_{ij})$ 。显然可得  $\langle \mathcal{A}, \mathcal{B} \rangle = \langle \overline{\mathcal{A}}, \overline{\mathcal{B}} \rangle$  以及

$$\mathcal{C} = \mathcal{A} \otimes \mathcal{B} \Leftrightarrow \overline{c}_{ij}(k) = \sum_{l=1}^s \overline{a}_{il}(l) \overline{b}_{lj}(l), k \in [p].$$

上式也等同于:

$$\overline{\mathcal{C}}_{(k)} = \overline{\mathcal{A}}_{(k)} \overline{\mathcal{B}}_{(k)}, k \in [p].$$

若张量  $\mathcal{A} \in \mathbf{C}_p^{m \times m}$  是  $f$ -对角张量且所有对角管均为  $e_Q, e_Q = \varphi_Q^{-1}(e), e = (1, 1, \dots, 1)^T \in \mathbf{C}_p$ , 则称  $\mathcal{A}$  为矩阵  $Q$  下的单位张量, 记为  $\mathcal{I}_m$ 。显然  $\mathcal{A}$  为矩阵  $Q$  下的单位张量当且仅当  $\overline{\mathcal{A}}_{(k)}$  对任意  $k \in [p]$  均为  $m \times m$  的单位矩阵。若存在张量  $\mathcal{B} \in \mathbf{C}_p^{m \times m}$  使得  $\mathcal{A} \otimes \mathcal{B} = \mathcal{B} \otimes \mathcal{A} = \mathcal{I}_m$ , 则称张量  $\mathcal{A} \in \mathbf{C}_p^{m \times m}$  为可逆张量, 记  $\mathcal{A}^{-1} = \mathcal{B}$ 。特别地, 若  $\mathcal{A}^{-1} = \mathcal{A}^H$ , 则称  $\mathcal{A}$  为  $\otimes$ -酉张量; 若对于张量  $\mathcal{A} \in \mathbf{C}_p^{m \times q} (q \leq m)$  也满足  $\mathcal{A}^H \mathcal{A} = \mathcal{I}_q$ , 则称  $\mathcal{A}$  为部分  $\otimes$ -酉张量。对于张量  $\mathcal{A} \in \mathbf{C}_p^{m \times n}, \mathcal{B} \in \mathbf{C}_p^{m \times s}$  且  $\mathcal{B}$  为部分  $\otimes$ -酉张量; 若存在  $\mathcal{C} \in \mathbf{C}_p^{s \times n}$ , 使得  $\mathcal{A} = \mathcal{B} \otimes \mathcal{C}$ , 则定义  $\mathcal{B}$  中的管列向量  $\{\mathcal{B}_{\cdot 1}, \mathcal{B}_{\cdot 2}, \dots, \mathcal{B}_{\cdot s}\}$  为张量  $\mathcal{A}$  的管列向量在运算  $\otimes$  下的一组正交基, 也可等价表示为  $\mathcal{A} = \mathcal{B} \otimes \mathcal{B}^H \otimes \mathcal{A}$ 。

通过上述变换后的张量 T-积, 可以得到三阶张量在变换矩阵  $Q$  下的张量奇异值分解 (tensor singular value decomposition, T-SVD)。对于给定的三阶张量  $\mathcal{A} \in \mathbf{C}_p^{m \times n}$ , 存在 2 个变换  $Q$  下的正交张量  $\mathcal{U} \in \mathbf{C}_p^{m \times m}$  和  $\mathcal{V} \in \mathbf{C}_p^{n \times n}$ , 以及一个  $f$ -对角张量  $\mathcal{S} \in \mathbf{C}_p^{m \times n}$ , 使得  $\mathcal{A} = \mathcal{U} \otimes \mathcal{S} \otimes \mathcal{V}^{-1}$ , 这被称为张量  $\mathcal{A}$  在变换  $Q$  下的张量奇异值分解。

**定义 2**<sup>[26]</sup> ( $Q$  变换下的张量管秩) 给定正交阵  $Q \in \mathbf{C}^{p \times p}$ , 张量  $\mathcal{A} = (a_{il}) \in \mathbf{C}_p^{m \times s}$  在变换  $Q$  下 ( $\overline{\mathcal{A}} = \varphi_Q(\mathcal{A})$ ) 的第  $k$  个正面切片的秩记为  $r_k$ , 即  $r_k = \text{rank}(\overline{\mathcal{A}}_{(k)}), k \in [p]$ 。令  $\mathcal{A}$  在变换  $Q$  下的管秩记为  $\text{rank}_t(\mathcal{A})$ , 则有  $\text{rank}_t(\mathcal{A}) = \#\{i \mid \mathcal{S}(i, i, :) \neq 0\} = \max\{r_k \mid k \in [p]\}$ , 其中  $\#$  表示集合的基数 (即元素个数),  $\mathcal{S}$  来源于  $\mathcal{A}$  在变换  $Q$  下的张量奇异值分解  $\mathcal{A} = \mathcal{U} \otimes \mathcal{S} \otimes \mathcal{V}^{-1}$ 。

### 3 新模型的提出与求解

本节首先提出了一个用于三阶张量时间序列的张量自回归模型, 再借助稀疏张量优化技术对该自回归模型中涉及的系数张量求解提出了一种可行的计算方法。

#### 3.1 模型简介

传统的自回归模型首先用于以向量为数据结构的时间序列预测, 其中每天的观测数据均为向量形式, 自回归参数为矩阵结构,  $S$  阶的传统自回归模型为:

$$\mathbf{x}_t = \sum_{s=1}^S \mathbf{A}_s \mathbf{x}_{t-s} + \mathbf{r}_t, t = S+1, \dots, T.$$

其中:  $\mathbf{x}_t$  为每天观测的向量数据,  $\mathbf{A}_s$  为矩阵自回归参数。该模型通过传统的矩阵乘法构造出与之对应的自回归模型。然而, 对于以张量等更高阶形式出现的数据, 此模型则无法利用矩阵乘法将自回归参数与张量数据进行关联, 仍需以破坏张量数据结构的形式将之转化为向量后再代入模型进行回归预测。因此, 这就需要引入张量积运算替代传统的矩阵乘法, 再结合自回归模型得到本文提出的张量自回归模型。

具体来说, 本文模型中第  $t$  天的张量观测值均由先前的观测值经过  $\otimes$ -乘法的运算后所得。假如考虑训练时间长度为  $T$  的时间序列, 每一个时间点  $t$  处的观测值表示为一个三阶张量  $\mathcal{X}_t \in \mathbf{C}_p^{m \times n}$ , 则有下面滞后项长度为  $S$  的张量自回归模型:

$$\mathcal{X}_t = \sum_{s=1}^S \mathcal{A}_s \otimes \mathcal{X}_{t-s} + \mathcal{R}_t, t = S+1, \dots, T, \quad (1)$$

其中:  $\mathcal{A}_s \in \mathbf{C}_p^{m \times m}$  为自回归系数张量,  $\mathcal{R}_t \in \mathbf{C}_p^{m \times n}$  为白噪声张量,  $1 \leq s \leq T$ 。

和传统的自回归模型相比,本文提出的模型具有以下 2 个优势:1) 可以有效地保留原张量的数据结构,更好地反应数据之间的潜在关系;2) 能有效减少模型中的参数数量。本模型可以在  $\otimes$ -乘法下将张量的管行之间的信息进行关联起来,从而更有效地利用数据。比如:当  $S=1$  时,不难发现张量  $\mathcal{A}_1$  反映的是管行之间的关系。以交通数据集为例,近些年来,以三阶张量为数据结构出现的交通数据集越来越多,其中数据的第 1 个维度表示对应的观测站点,例如需要同时对 300 个站点进行观测,则张量的第 1 个维度大小为 300;张量的第 2 个维度则对应每天的观测频率,例如每天以 5 min 为间隔作 1 次记录,全天 24 h 不间断记录,则对应数据的第 2 个维度为 288;张量的第 3 个维度表示观测的特征,例如车辆速度、车辆数量等。通过这样的三阶张量对数据进行表示,在构建的张量自回归模型中,就可以通过  $\otimes$ -乘法将同一个时间节点的所有观测值进行信息整合,统一充分利用。

### 3.2 模型求解

为求解模型(1)中的自回归系数  $\mathcal{A}_S$ ,本文建立了如下最优化问题:

$$\min_{\mathcal{A} \in \mathbf{C}^{m \times (Sm)}} \alpha \text{rank}_t(\mathcal{A}) + \frac{1}{2} \sum_{t=S+1}^T \|\mathcal{X}_t - \mathcal{A} \otimes \mathcal{Y}_t\|_F^2 + \frac{\beta}{2} \|\mathcal{A}\|_F^2, \quad (2)$$

其中:  $\mathcal{A} = [\mathcal{A}_1, \mathcal{A}_2, \dots, \mathcal{A}_S]$ ,  $\mathcal{Y}_t = [\mathcal{X}_{t-1}, \mathcal{X}_{t-2}, \dots, \mathcal{X}_{t-S}]^T$ ,  $t = S+1, \dots, T$ 。

由于问题(2)中存在管秩  $\text{rank}_t$ ,所以该问题是 NP 难且非凸问题,无法找到一个直接有效的办法进行求解。但是,可以通过核范数  $\|\cdot\|_*$  逼近管秩  $\text{rank}_t$  范数的方法将问题(2)松弛为:

$$\begin{aligned} \min_{\bar{\mathbf{A}}_{(1)}, \dots, \bar{\mathbf{A}}_{(p)} \in \mathbf{C}^{m \times (Sm)}} \alpha \sum_{k=1}^p \|\bar{\mathbf{A}}_{(k)}\|_* + f(\bar{\mathbf{A}}_{(1)}, \dots, \bar{\mathbf{A}}_{(p)}), \\ f(\bar{\mathbf{A}}_{(1)}, \dots, \bar{\mathbf{A}}_{(p)}) = \frac{\beta}{2} \sum_{k=1}^p \|\bar{\mathbf{A}}_{(k)}\|_F^2 + \frac{1}{2} \sum_{t=S+1}^T \sum_{k=1}^p \|\bar{\mathbf{X}}_{t,(k)} - \bar{\mathbf{A}}_{(k)} \bar{\mathbf{Y}}_{t,(k)}\|_F^2. \end{aligned} \quad (3)$$

很容易注意到对问题(3)进行求解可等价转化为对  $p$  个最优化问题求解:

$$\min_{\bar{\mathbf{A}}_{(k)} \in \mathbf{C}^{m \times (Sm)}} \alpha \|\bar{\mathbf{A}}_{(k)}\|_* + \frac{\beta}{2} \|\bar{\mathbf{A}}_{(k)}\|_F^2 + \frac{1}{2} \sum_{t=S+1}^T \|\bar{\mathbf{X}}_{t,(k)} - \bar{\mathbf{A}}_{(k)} \bar{\mathbf{Y}}_{t,(k)}\|_F^2,$$

也等价表示为:

$$\begin{aligned} \min_{\bar{\mathbf{A}}_{(k)} \in \mathbf{C}^{m \times (Sm)}} \alpha \|\bar{\mathbf{A}}_{(k)}\|_* + \frac{\beta}{2} \|\bar{\mathbf{B}}_{(k)}\|_F^2 + \frac{1}{2} \sum_{t=S+1}^T \|\bar{\mathbf{X}}_{t,(k)} - \bar{\mathbf{B}}_{(k)} \bar{\mathbf{Y}}_{t,(k)}\|_F^2, \\ \bar{\mathbf{A}}_{(k)} = \bar{\mathbf{B}}_{(k)}. \end{aligned} \quad (4)$$

令参数  $\gamma \geq 0$ ,问题(4)的增广拉格朗日函数为:

$$\begin{aligned} \mathcal{L}(\bar{\mathbf{A}}_{(k)}, \bar{\mathbf{B}}_{(k)}, \bar{\mathbf{U}}_{(k)}) = \alpha \|\bar{\mathbf{A}}_{(k)}\|_* + \frac{\beta}{2} \|\bar{\mathbf{B}}_{(k)}\|_F^2 + \frac{1}{2} \sum_{t=S+1}^T \|\bar{\mathbf{X}}_{t,(k)} - \bar{\mathbf{B}}_{(k)} \bar{\mathbf{Y}}_{t,(k)}\|_F^2 + \\ \langle \bar{\mathbf{U}}_{(k)}, \bar{\mathbf{A}}_{(k)} - \bar{\mathbf{B}}_{(k)} \rangle + \frac{\gamma}{2} \|\bar{\mathbf{A}}_{(k)} - \bar{\mathbf{B}}_{(k)}\|_F^2, \end{aligned} \quad (5)$$

用交替极小化方法对式(5)进行求解,通过对当前点  $(\bar{\mathbf{A}}_{(k)}^q, \bar{\mathbf{B}}_{(k)}^q, \bar{\mathbf{U}}_{(k)}^q)$  迭代得到下一步  $(\bar{\mathbf{A}}_{(k)}^{q+1}, \bar{\mathbf{B}}_{(k)}^{q+1}, \bar{\mathbf{U}}_{(k)}^{q+1})$ 。

下面为每个参数求解过程。

**子问题  $\bar{\mathbf{A}}_{(k)}$  ( $k=1, 2, \dots, p$ )** 考虑求解最优化问题:

$$\min_{\bar{\mathbf{A}}_{(k)}} \alpha \|\bar{\mathbf{A}}_{(k)}\|_* + \frac{\gamma}{2} \left\| \bar{\mathbf{A}}_{(k)} - \bar{\mathbf{B}}_{(k)}^q + \frac{1}{\gamma} \bar{\mathbf{U}}_{(k)}^q \right\|_F^2.$$

该问题的最优解为:

$$\bar{\mathbf{A}}_{(k)}^{q+1} = \mathcal{D}_\tau \left( \bar{\mathbf{B}}_{(k)}^q - \frac{1}{\gamma} \bar{\mathbf{U}}_{(k)}^q, \frac{\alpha}{\gamma} \right). \quad (6)$$

**注**  $\mathcal{D}_\tau(\cdot, \cdot)$  是著名的奇异值收缩算子,  $\mathcal{D}_\tau(\mathbf{M}, \gamma) = \mathbf{U} \mathcal{S}_\gamma(\boldsymbol{\Sigma}) \mathbf{V}^T$ 。其中:  $\mathbf{M} = \mathbf{U} \boldsymbol{\Sigma} \mathbf{V}^T$  是矩阵  $\mathbf{M} \in \mathbf{C}^{m \times n}$  在降阶形式下的奇异值分解,且秩为  $r$ 。矩阵  $\mathbf{U}$  和  $\mathbf{V}$  分别是  $m \times r$  和  $n \times r$  的列正交矩阵,  $\boldsymbol{\Sigma} = \text{diag}(\{\delta_i\}_{1 \leq i \leq r})$  是对角矩阵。收缩算子定义为  $\mathcal{S}_\gamma(\boldsymbol{\Sigma}) = \text{diag}(\max\{\delta_i - \gamma, 0\})$ , 其中  $\gamma > 0$ 。

**子问题  $\bar{\mathbf{B}}_{(k)}$  ( $k=1, 2, \dots, p$ )** 考虑求解最优化问题:

$$\min_{\bar{\mathbf{B}}_{(k)}} \frac{\beta}{2} \|\bar{\mathbf{B}}_{(k)}\|_F^2 + \frac{1}{2} \sum_{t=S+1}^T \|\bar{\mathbf{X}}_{t,(k)} - \bar{\mathbf{B}}_{(k)} \bar{\mathbf{Y}}_{t,(k)}\|_F^2 + \frac{\gamma}{2} \left\| \bar{\mathbf{A}}_{(k)}^{q+1} - \bar{\mathbf{B}}_{(k)} + \frac{1}{\gamma} \bar{\mathbf{U}}_{(k)}^q \right\|_F^2.$$

该问题的最优解为:

$$\bar{\mathbf{B}}_{(k)}^{q+1} = (\gamma \bar{\mathbf{A}}_{(k)}^{q+1} + \bar{\mathbf{U}}_{(k)}^q + \sum_{t=S+1}^T \bar{\mathbf{X}}_{t,(k)} \bar{\mathbf{Y}}_{t,(k)}^T) ((\gamma + \beta) \mathbf{I} + \sum_{t=S+1}^T \bar{\mathbf{Y}}_{t,(k)} \bar{\mathbf{Y}}_{t,(k)}^T)^{-1}. \quad (7)$$

对子问题  $\bar{\mathbf{U}}_{(k)}$  ( $k=1, 2, \dots, p$ ) 进行如下迭代:

$$\bar{\mathbf{U}}_{(k)}^{q+1} = \bar{\mathbf{U}}_{(k)}^q + \gamma (\bar{\mathbf{A}}_{(k)}^{q+1} - \bar{\mathbf{B}}_{(k)}^{q+1}). \quad (8)$$

综上所述,下面给出本文模型的算法流程。

**算法** 基于 T-SVD 的张量自回归模型算法

步骤 1, 选取  $\alpha, \beta, \gamma, \epsilon > 0$  以及初始点  $(\bar{\mathbf{A}}^0, \bar{\mathbf{B}}^0, \bar{\mathbf{U}}^0)$ ;  $k=1, 2, \dots, p$ ;  $q=0, 1, 2, \dots$ 。

步骤 2, 通过式(6)计算  $\bar{\mathbf{A}}_{(k)}^{q+1}$ 。

步骤 3, 通过式(7)计算  $\bar{\mathbf{B}}_{(k)}^{q+1}$ 。

步骤 4, 通过式(8)计算  $\bar{\mathbf{U}}_{(k)}^{q+1}$ 。

步骤 5, 若  $\max\{\|(\bar{\mathbf{A}}_{(k)}^{q+1}, \bar{\mathbf{B}}_{(k)}^{q+1}) - (\bar{\mathbf{A}}_{(k)}^q, \bar{\mathbf{B}}_{(k)}^q)\|_F, \|\bar{\mathbf{A}}_{(k)}^{q+1} - \bar{\mathbf{B}}_{(k)}^{q+1}\|_F\} \leq \epsilon$  不成立, 则返回步骤 2。

终止迭代。

通过光滑函数的下降引理<sup>[27]</sup>易知, 使用 ADMM 算法求解本模型可以得到收敛性结果。

### 3.3 模型计算复杂度分析

下面分别列出了上述算法单次迭代中每个步骤的计算成本, 以揭示了该算法的效率和性能表现。

在通过式(6)对子问题  $\bar{\mathbf{A}}_{(k)}$  进行迭代求解的过程中, 计算复杂度为  $O(pSm^3)$ 。同理, 子问题  $\bar{\mathbf{B}}_{(k)}$  迭代求解的计算复杂度为  $O(p(T-S)S^2m^2n + S^3m^3)$ , 通过式(8)对子问题  $\bar{\mathbf{U}}_{(k)}$  迭代求解所需的计算复杂度为  $O(pSm^2)$ , 以上所有步骤的总计算复杂度为  $O(pSm^3 + S^3m^3 + p(T-S)S^2m^2n)$ 。

表 1 本模型每次迭代的计算复杂度

Tab. 1 The computational complexity of each iteration of the model

步骤	计算复杂度
更新 $\bar{\mathbf{A}}_{(k)}$	$O(pSm^3)$
更新 $\bar{\mathbf{B}}_{(k)}$	$O(p(T-S)S^2m^2n + S^3m^3)$
更新 $\bar{\mathbf{U}}_{(k)}$	$O(pSm^2)$
总计	$O(pSm^3 + S^3m^3 + p(T-S)S^2m^2n)$

## 4 数值实验

下面介绍本文模型在时间序列预测的背景下对高阶时间序列进行建模预测的数值实验部分。相关实验在 2 个合成数据集以及 4 个真实数据集下进行。所有数值实验均在配备第 11 代 Intel(R) Core (TM) i5-1155G7@2.50 GHz 和 16 GB 内存的笔记本电脑上进行, 对应 Matlab 版本为 2016a。

### 4.1 数据集以及参数设置

本文选取了 2 个合成数据集(Syndata1 和 Syndata2)以及 3 个真实数据集(PEMSD4、PEMSD8 和 SST)进行相关数值实验。模型涉及的参数设置为  $\alpha=0.1, \beta=0.1, \gamma=1, \epsilon=0.001$ , 同时正交变换矩阵  $\mathbf{Q}$  为离散傅里叶变换矩阵。

1) Syndata1: 首先通过随机生成一个三阶张量  $\mathcal{X}_1 \in \mathbf{R}^{100 \times 100 \times 10}$ , 再对  $\mathcal{X}_1$  进行 T-SVD 分解:

$$\mathcal{X}_1 = \mathcal{U} \otimes \mathcal{S} \otimes \mathcal{V}^{-1}. \quad (9)$$

其中,  $\mathcal{U} \in \mathbf{R}^{100 \times 100 \times 10}$ ,  $\mathcal{V} \in \mathbf{R}^{100 \times 100 \times 10}$  均为正交张量,  $\mathcal{S} \in \mathbf{R}^{100 \times 100 \times 10}$ ,  $\mathcal{S}$  为  $f$ -对角张量。由于现实数据集往往具有一定的趋势性, 可通过指数函数生成三阶趋势张量  $\mathcal{O} \in \mathbf{R}^{100 \times 100 \times 10}$ , 其中  $\mathcal{O}$  的每个  $\mathcal{O}(:, :, k)$ ,  $k=1, 2, \dots, p$  均由 linspace 函数生成  $m * n$  个元素并将其排列成矩阵, 最后再对式(9)进行扰动得到:

$$\mathcal{X}_i = \mathcal{U} \otimes \left( \mathcal{S} + \log\left(\frac{i}{T}\right) * \text{rand} * \mathcal{I} \right) \otimes \mathcal{V}^T + \mathcal{O} + \sin(i) * \mathcal{E}, \quad (10)$$

其中,  $i=1, 2, \dots, T$ ,  $\mathcal{I}$  为单位  $f$ -对角张量,  $\mathcal{E}$  为随机张量。由式(10)得到相关合成数据集(Syndata1)。

2) Syndata2:  $i, j \in [100], k \in [10], \mathcal{X}(i, j, k, t) = \sqrt{j} + \text{rand} * \left(\frac{t}{31}\right)^3, t \in [31]$ 。

3) 海面温度(SST)数据集<sup>[24]</sup>:该数据集覆盖范围为北纬 5°到南纬 5°、西经 180°到西经 110°,按照经纬度划分为一个规模为  $5 \times 6$  的网格矩阵,以 1 h 1 次的间隔周期记录,收集了从 1994 年 4 月 26 日 19:00 到 7 月 19 日 3:00 的全部海面温度信息数据,累计记录 2 000 h,数据规模为(5,6,1,2 000)。

4) Tesla 数据集<sup>[24]</sup>:统计了 14 家公司每天股票的最高价、最低价、成交价格、开盘价以及收盘价 5 个股数指标,并且以 1 周(7 d)为 1 次间隔周期记录数据,累计记录 180 周,数据规模为(14,5,7,180)。

5) GÉANT 数据集<sup>[28]</sup>:由 23 个发射器以及 23 个接收器以 15 min 为 1 次间隔周期传递网络流量数据,全天 24 h 记录,共记录 119 d,数据规模为(23,23,96,119)。

6) PEMS8 数据集<sup>[29]</sup>:在 2016 年 7 月 1 日至 8 月 31 日期间,从圣贝纳迪诺地区的 170 个探测器收集的交通信息。数据格式为(17 856,170,3),每个探测器以 5 min 为 1 个间隔周期记录数据,全天 24 h 不间断记录。每个检测器每天总共包含 288 个数据样本,将数据集改写后可以视为一个四阶张量,对应大小为(288,170,3,59)。

在对比实验中,选择同样基于张量 T-SVD 构建的低秩张量线性回归模型 TLRFR 作为对比算法,并进行一系列性能评测。为确保数值实验的有效性,本文采用归一化平均绝对误差作为统一的预测精度指标,记为  $\sigma_{\text{NMAE}}$ ,即:

$$\sigma_{\text{NMAE}} = \frac{\sum_i \sum_{(j,k,l)} |(\mathcal{X}_i^{\text{true}} - \mathcal{X}_i^{\text{pre}})_{jkl}|}{\sum_i \sum_{(j,k,l)} |\mathcal{X}_i^{\text{true}}|_{jkl}},$$

其中:  $\mathcal{X}_i^{\text{true}}$ ,  $\mathcal{X}_i^{\text{pre}}$  分别表示真实数据和预测数据,该指标使本文能够与不同模型在有效预测时间序列数据方面的准确性进行评估。显然,  $\sigma_{\text{NMAE}}$  值越低,说明算法对于数据的恢复效果越好;反之,说明效果越差。

## 4.2 模型评估

所有数值实验中设置训练集天数为  $T$ ,其中包含滞后项天数为  $S$ ,即通过前  $T$  天进行训练后,利用训练集进行自回归参数估计时,需要选取连续的滞后天数  $S$  来作为  $S+1$  天的线性组合。例如训练集中第  $S+1$  天的数据可由前  $S$  天的观测数据线性表示为  $\mathcal{X}_{S+1} = \mathcal{A}_1 \otimes \mathcal{X}_1 + \dots + \mathcal{A}_S \otimes \mathcal{X}_S + \mathcal{R}_{S+1}$ 。依次类推,通过训练集可得  $T-S$  个线性组合,将所有的线性组合代入模型公式,由此估计出回归参数,再通过  $S$  天的数据值对第  $T+1$  天进行预测,通过预测值的效果对本模型进行评估。本节主要从 2 个方面对本模型进行评估:一是精度,对比分析 2 个合成数据集和 4 个真实数据集上的精度预测表现;二是效率,衡量在不同数据集上的预测耗时。

### 4.2.1 合成数据集

本节在合成数据集 Syndata1 和 Syndata2 上进行了模型预测实验,部分预测结果的精度见表 2。实验过程中,针对不同训练集大小  $T$  和滞后项长度  $S$  进行了测试。结果表明,无论训练集规模或滞后项长度如何变化,本文提出的模型均能保持较优的预测性能,体现出良好的可行性和有效性。

表 2 数据集 Syndata1、Syndata2 上不同模型的  $\sigma_{\text{NMAE}}$  值  
Tab.2 The prediction NMAE value for Syndata1, Syndata2

数据集	$T$	$S$	本模型的 $\sigma_{\text{NMAE}}$ 值	TLRFR 模型的 $\sigma_{\text{NMAE}}$ 值	数据集	$T$	$S$	本模型的 $\sigma_{\text{NMAE}}$ 值	TLRFR 模型的 $\sigma_{\text{NMAE}}$ 值
Syndata1	2	2	0.054	0.423	Syndata2	19	2	0.07	0.81
		4	0.005	0.329			4	0.01	0.81
		6	0.004	0.235			6	0.01	0.81
	22	2	0.048	0.200		6	2	0.01	0.81
		4	0.005	0.188			4	0.01	0.81
		6	0.003	0.162			6	0.01	0.81

### 4.2.2 真实数据集

本文在 4 个真实数据集(SST、GÉANT、Tesla 和 Pems08)下分别选取了不同的训练集大小  $T$  以及滞后项  $S$

进行模型预测精度对比实验,部分实验结果见表 3。

表 3 真实数据集上不同模型的  $\sigma_{NMAE}$  值

Tab. 3 The prediction NMAE value for real datasets of different models

数据集	T	S	本模型的 $\sigma_{NMAE}$ 值	TLRFR 模型的 $\sigma_{NMAE}$ 值	数据集	T	S	本模型的 $\sigma_{NMAE}$ 值	TLRFR 模型的 $\sigma_{NMAE}$ 值
SST	5	2	0.003	0.017	Tesla	16	2	0.041	0.067
		4	0.003	0.022			4	0.050	0.077
	10	2	0.002	0.017		24	2	0.035	0.061
		4	0.002	0.019			4	0.032	0.062
GÉANT	12	2	0.291	0.367	Pems08	18	2	0.168	1.289
		4	0.264	0.329			4	0.228	1.152
	19	2	0.254	0.392		20	2	0.178	0.948
		4	0.297	0.344			4	0.210	0.738

通过不同数据集下本文模型与其他模型的  $\sigma_{NMAE}$  值对比可知,在相同的情况下本文所提出的模型预测精度总是要优于 TLRFR 模型。下面展示了在 GÉANT 数据集下的部分预测效果可视化图(图 1)。

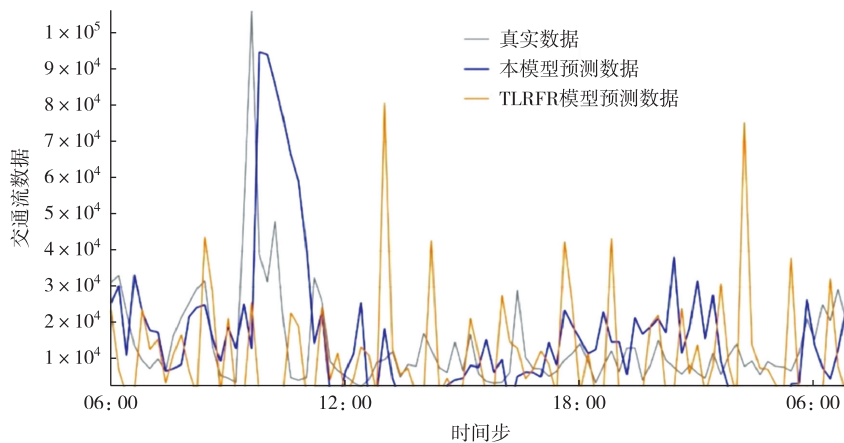


图 1 真实数据集 GÉANT 部分预测结果可视化

Fig. 1 Visualization of partial prediction results of on GÉANT

通过该图可直观对比,本文模型的总体预测结果优于 TLRFR 模型。在选取不同训练集规模和滞后项后,对本文提出的模型在预测耗时方面进行了评估。实验结果表明,在大多数情况下,本模型的计算时间低于 TLRFR 模型,部分结果列于表 4。

表 4 不同真实数据集上不同模型的预测耗时

Tab. 4 Prediction time-consuming cost table for different real datasets of different models

数据集	T	S	本模型耗时/s	TLRFR 模型耗时/s	数据集	T	S	本模型耗时/s	TLRFR 模型耗时/s
SST	5	2	0.010	0.004	Tesla	7	2	0.063	1.420
		4	0.010	0.002			4	0.051	1.399
	10	2	0.010	0.003		10	2	0.064	4.099
		4	0.008	0.004			4	0.056	4.131
GÉANT	12	2	0.291	0.263	Pems08	18	2	1.122	1.161
		4	0.312	0.236			4	1.374	1.452
	19	2	0.301	0.484		20	2	1.658	1.171
		4	0.326	0.478			4	1.722	1.167

通过数值实验,从预测准确度和计算耗时 2 个方面对模型进行对比分析。实验结果表明,在相同训练集规模下,无论是在合成数据上还是在真实数据上,本文模型的预测准确度均优于 TLRFR 模型,且整体计算耗时更少。综合来看,本模型在预测性能和计算效率方面均具备优势,体现出良好的实用性和有效性。

## 5 总结

本文基于张量奇异值分解和传统自回归模型,针对时间序列领域中高阶数据预测的挑战,依托张量积相关理论及运算,提出了一种新的低秩张量自回归模型,并借助最优化理论,通过交替极小化方法计算自回归系数,以获得最优参数。数值实验结果表明,该模型在高维数据张量的预测上具有良好的有效性和实用性。后续将结合张量积定义与自回归积分滑动平均模型,进一步构建基于张量数据结构的新型自回归积分滑动平均模型,以拓展张量在时间序列预测中的应用。

### 参考文献:

- [1] NELDER J A, WEDDERBUM R W. Generalized linear models[J]. *Journal of the Royal Statistical Society Series A: Statistics in Society*, 1972, 135(3): 370-384.
- [2] STONE M, BROOKS R J. Continuum regression: cross-validated sequentially constructed prediction embracing ordinary least squares, partial least squares and principal components regression [J]. *Journal of the Royal Statistical Society Series B (Methodological)*, 1990, 52(2): 237-269.
- [3] RASMUSSEN C E. Gaussian processes in machine learning[M]//Advanced lectures on machine learning. Berlin: Springer, 2004: 63-71.
- [4] LIAW A, WIENER M C. Classification and regression by randomForest[J]. *R news*, 2002, 2(3): 18-22.
- [5] SMOLA A J, SCHÖLKOPF B. A tutorial on support vector regression[J]. *Statistics and Computing*, 2004, 14(3): 199-222.
- [6] HAVKIN S. Neural Network: A comprehensive foundation[M]. New Jersey: Prentice Hall, 1998.
- [7] DREYFUS G, SPRINGERLINK. Neural networks: methodology and applications[M]. Berlin: Springer, 2005.
- [8] YAN K, LIU G N, XIE R D, et al. Novel subject-dependent human-posture recognition approach using tensor regression[J]. *IEEE Sensors Journal*, 2025, 25(1): 1041-1053.
- [9] LIAO T C, ZHANG L, YANG J H, et al. CoATR: a convolutional autoregressive tensor-ring decomposition method for sparse spatio-temporal traffic data[J]. *Neurocomputing*, 2025, 617: 129006.
- [10] ALGARTE R D. High-order tensor regression in sparse convolutional neural networks[EB/OL]. (2025-01-02)[2025-03-04]. <https://arxiv.org/abs/2501.01239>.
- [11] SEMOCHKINA D, JACKSON S E, WOODS D C. Tensor-variate Gaussian process regression for efficient emulation of complex systems: comparing regressor and covariance structures in outer product and parallel partial emulators[EB/OL]. (2025-02-14)[2025-03-04]. <https://arxiv.org/abs/2502.10319>.
- [12] YANG Z, LI W R, HOU J X, et al. Multi-modal spiking tensor regression network for audio-visual zero-shot learning[J]. *Neurocomputing*, 2025, 629: 129636.
- [13] ROGERS M, LI L, RUSSELL S J. Multilinear dynamical systems for tensor time series [C]//Proceedings of the 27<sup>th</sup> International Conference on Neural Information Processing Systems, New York: Curran Associates, Inc., 2013: 2634-2642.
- [14] CARROLL J D, CHANG J J. Analysis of individual differences in multidimensional scaling via an  $N$ -way generalization of "Eckart-young" decomposition[J]. *Psychometrika*, 1970, 35(3): 283-319.
- [15] TUCKER L R. Implications of factor analysis of three-way matrices for measurement of change[J]. *Problems in Measuring Change*, 1963, 15(3): 122-137.
- [16] OSELEDETS I V. Tensor-train decomposition[J]. *SIAM Journal on Scientific Computing*, 2011, 33(5): 2295-2317.
- [17] KILMER M E, MARTIN C D. Factorization strategies for third-order tensors[J]. *Linear Algebra and Its Applications*, 2011, 435(3): 641-658.
- [18] GAO Q X, CHENG J F, XIE D Y, et al. Tensor linear regression and its application to color face recognition[C]//2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW). Piscataway: IEEE, 2019: 523-531.
- [19] CHEN R, XIAO H, YANG D. Autoregressive models for matrix-valued time series[J]. *Journal of Econometrics*, 2021, 222(1): 539-560.
- [20] DAVIS R A, ZANG P F, ZHENG T. Sparse vector autoregressive modeling[J]. *Journal of Computational and Graphical*

- Statistics, 2016, 25(4):1077-1096.
- [21] LI Z B, XIAO H. Multi-linear tensor autoregressive models[EB/OL]. (2021-10-03)[2025-03-04]. <https://arXiv.org/abs/2110.00928>.
- [22] HILL C, LI J, SCHNEIDER M J, et al. The tensor auto-regressive model[J]. Journal of Forecasting, 2021, 40(4):636-652.
- [23] WANG D, ZHENG Y, LIAN H, et al. High-dimensional vector autoregressive time series modeling via tensor decomposition [J]. Journal of the American Statistical Association, 2022, 117(539):1338-1356.
- [24] CATES J, HOOVER R C, CAUDLE K, et al. Forecasting multilinear data via transform-based tensor autoregression[EB/OL]. (2022-05-24)[2025-03-04]. <https://arXiv.org/abs/2205.12201>.
- [25] WANG H, ZHANG L, LI S E. Low-rank tensor autoregressive predictor for third-order time-series forecasting[EB/OL]. (2024-03-05)[2025-03-04]. <https://doi.org/10.48550/arXiv.2403.02835>.
- [26] KERNFELD E, KILMER M, AERON S. Tensor-tensor products with invertible linear transforms[J]. Linear Algebra and Its Applications, 2015, 485:545-570.
- [27] BERTSEKAS D P, TSITSIKLIS J N. Parallel and distributed computation: numerical methods [M]. Belmont: Athena Scientific, 1997.
- [28] UHLIG S, QUOITIN B, LEPROPRE J, et al. Providing public intradomain traffic matrices to the research community[J]. ACM SIGCOMM Computer Communication Review, 2006, 36(1):83-86.
- [29] CHEN Y, SEGOVIA-DOMINGUEZ I, COSKUNUZER B, et al. TAMP-S2GCNets: coupling time-aware multipersistence knowledge representation with spatio-supra graph convolutional networks for time-series forecasting [C]//International Conference on Learning Representations, 2022.

## Operations Research and Cybernetics

### A Tensor Autoregressive Method for High-Order Data Prediction

LI Ang, LIU Jinjie

(National Center for Applied Mathematics in Chongqing, Chongqing Normal University, Chongqing 401331, China)

**Abstract:** Aiming to construct a new regression model for high-order and high-dimensional time series using tensor methods. In recent years, tensor decomposition has become an important tool for data processing and analysis, gaining significant attention. Based on tensor singular value decomposition (T-SVD), a novel low-rank autoregressive model for tensor time series prediction under tensor T-product is proposed. The model parameters are estimated through an alternating minimization algorithm. Numerical experiments demonstrate the model's superior prediction accuracy. The established model have its effectiveness and practical applicability.

**Keywords:** tensor singular value decomposition; autoregressive model; Tensor Product; tubal rank; alternating minimization algorithm

(责任编辑 黄 颖)