Statistica Sinica **36** (2026), 1-19
doi:https://doi.org/10.5705/ss.202023.0009

A REGULARIZED LOW TUBAL-RANK MODEL FOR HIGH-DIMENSIONAL TIME SERIES DATA

Samrat Roy* and George Michailidis

Indian Institute of Management Ahmedabad and University of Florida

Abstract: High dimensional time series analysis has diverse applications in macroeconometrics and finance. Recent factor-type models employing tensor-based decompositions prove to be computationally involved due to the non-convex nature of the underlying optimization problem and also they do not capture the underlying temporal dependence of the latent factor structure. This work leverages the concept of tubal rank and develops a matrix-valued time series model, which first captures the temporal dependence in the data, and then the remainder signals across the time points are decomposed into two components: a low tubal rank tensor representing the baseline signals, and a sparse tensor capturing the additional idiosyncrasies in the signal. We address the issue of identifiability of various components in our model and subsequently develop a scalable Alternating Block Minimization algorithm to solve the convex regularized optimization problem for estimating the parameters. We provide finite sample error bounds under high dimensional scaling for the model parameters.

Key words and phrases: High dimensional time series, tensor, tubal rank.

1. Introduction

There has been a lot of interest in modeling high-dimensional time series due not only to traditional application areas in macroeconomics and finance (De Mol, Giannone and Reichlin, 2008; Blanchard and Perotti, 2002; Bernanke, Boivin and Eliasz, 2005), but also emerging ones, including dynamic traffic networks (Chen and Chen, 2022), functional genomics (Michailidis and d'Alché Buc, 2013) and neuroscience (Seth, Barrett and Barnett, 2015). To accommodate high dimensionality, both regularized versions of VAR models (Bańbura, Giannone and Reichlin, 2010; Basu and Michailidis, 2015; Kock and Callot, 2015; Ghosh, Khare and Michailidis, 2019; Wang, Zheng and Li, 2024) and dynamic factor models (Bai and Wang, 2015; Lam and Yao, 2012; Chang, Guo and Yao, 2018) have been developed. Wang, Liu and Chen (2019) and Chen, Yang and Zhang (2022) recently extended the aforementioned factor models to matrix and tensor valued time series by expressing the data into a low dimensional dynamic signal as $A_1 G_t A_2^T$, or $\mathcal{G}_t \times_1 A_1 \times_2 A_2 \times_3 \cdots \times_k A_k$ respectively, with A_i's being the loading matrices and G_t and $\mathcal{G}_t$ are the core factor matrix and tensors.

*Corresponding author. E-mail: samratroy67@yahoo.com