Statistica Sinica **36** (2026), 1-24
doi:https://doi.org/10.5705/ss.202023.0017

MULTISCALE AUTOREGRESSION ON ADAPTIVELY DETECTED TIMESCALES

Rafal Baranowski, Yining Chen* and Piotr Fryzlewicz

London School of Economics and Political Science

Abstract: We propose a multiscale approach to time series autoregression, in which linear regressors for the process in question include features of its own path that live on multiple timescales. We take these multiscale features to be the recent averages of the process over multiple timescales, whose number or spans are not known to the analyst and are estimated from the data via a change-point detection technique. The resulting construction, termed Adaptive Multiscale AutoRegression (AMAR) enables adaptive regularisation of linear autoregressions of large orders. The AMAR model is designed to offer simplicity and interpretability on the one hand, and modelling flexibility on the other. Our theory permits the longest timescale to increase with the sample size. A simulation study is presented to show the usefulness of our approach. Some possible extensions are also discussed, including the Adaptive Multiscale Vector AutoRegressive model (AMVAR) for multivariate time series, which demonstrates promising performance in the data example on UK and US unemployment rates. The R package amar (Baranowski, Chen and Fryzlewicz, 2022) provides an efficient implementation of the AMAR framework.

Key words and phrases: Multiscale modelling, piecewise-constant approximation, regularised autoregression, time series.

1. Introduction

1.1. Motivation and main idea

Autoregression in time series modelling is arguably the most frequently used device to characterise temporal dependence in data. The classical linear autoregressive model of order p, known as AR(p), for univariate time series X_t assumes that X_t is a linear but otherwise unconstrained function of its own past values $X_{t-1}, \ldots, X_{t-p}$, plus white-noise-like innovation ε_t, that is $X_t = \beta_1 X_{t-1} + \cdots + \beta_p X_{t-p} + \varepsilon_t$ for $t = 1, \ldots, T$. However, in situations where the application of this model yields a large or even moderate p, either in absolute terms or relative to T (perhaps in an attempt to reflect long-range dependence in X_t), it may be tempting to consider instead an alternative approach, in which X_t is regressed explicitly on some other features of its own past, rather than directly on the individual variables $X_{t-1}, \ldots, X_{t-p}$.

*Corresponding author. E-mail: Y.Chen101@lse.ac.uk