



Data-driven model selection for same-realization predictions in autoregressive processes

Kare Kamila¹

Received: 11 April 2021 / Revised: 13 September 2022 / Accepted: 11 October 2022 /
Published online: 27 November 2022
© The Institute of Statistical Mathematics, Tokyo 2022

Abstract

This paper is about the one-step ahead prediction of the future of observations drawn from an infinite-order autoregressive $AR(\infty)$ process. It aims to design penalties (fully data driven) ensuring that the selected model verifies the efficiency property but in the non-asymptotic framework. We show that the excess risk of the selected estimator enjoys the best bias-variance trade-off over the considered collection. To achieve these results, we needed to overcome the dependence difficulties by following a classical approach which consists in restricting to a set where the empirical covariance matrix is equivalent to the theoretical one. We show that this event happens with probability larger than $1 - c_0/n^2$ with $c_0 > 0$. The proposed data-driven criteria are based on the minimization of the penalized criterion akin to the Mallows's C_p .

Keywords Model selection · Oracle inequality · Efficiency · Autoregressive process · Data driven

Kare Kamila has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 754362.

✉ Kare Kamila
kamilakare@gmail.com

¹ SAMM, Université Paris 1, Panthéon-Sorbonne, 90, rue de Tolbiac 75634 Paris, France