On model selection from a finite family of possibly misspecified time series models

Hsiang-Ling Hsu[1]*, Ching-Kang Ing[2], Howell Tong[3]

[1] National University of Kaohsiung, Taiwan
[2]National Tsing Hua University, Taiwan
[3]China and London School of Economics, United Kingdom

Abstract

Consider finite parametric time series models. `I have n observations and k models, which model should I choose on the basis of the data alone?' is a frequently asked question in many practical situations. This poses the key problem of selecting a model from a collection of candidate models, none of which is necessarily the true data generating process (DGP). Although existing literature on model selection is vast, there is a serious lacuna in that the above problem does not seem to have received much attention. In fact, existing model selection criteria have avoided addressing the above problem directly, either by assuming that the true DGP is included among the candidate models and aiming at choosing this DGP, or by assuming that the true DGP can be asymptotically approximated by an increasing sequence of candidate models and aiming at choosing the candidate having the best predictive capability in some asymptotic sense. In this article, we propose a misspecification-resistant information criterion (MRIC) to address the key problem directly. We first prove the asymptotic efficiency of MRIC whether the true DGP is among the candidates or not, within the fixed-dimensional framework. We then extend this result to the high-dimensional case in which the number of candidate variables is much larger than the sample size. In particular, we show that MRIC can be used in conjunction with a high-dimensional model selection method to select the (asymptotically) best predictive model across several high-dimensional misspecified time series models.