World Academy of Science, Engineering and Technology
International Journal of Economics and Management Engineering
Vol:12, No:05, 2018

Optimal Data Selection in Non-Ergodic Systems: A Tradeoff between Estimator Convergence and Representativeness Errors

Authors : Jakob Krause

Abstract : Past Financial Crisis has shown that contemporary risk management models provide an unjustified sense of security and fail miserably in situations in which they are needed the most. In this paper, we start from the assumption that risk is a notion that changes over time and therefore past data points only have limited explanatory power for the current situation. Our objective is to derive the optimal amount of representative information by optimizing between the two adverse forces of estimator convergence, incentivizing us to use as much data as possible, and the aforementioned non-representativeness doing the opposite. In this endeavor, the cornerstone assumption of having access to identically distributed random variables is weakened and substituted by the assumption that the law of the data generating process changes over time. Hence, in this paper, we give a quantitative theory on how to perform statistical analysis in non-ergodic systems. As an application, we discuss the impact of a paragraph in the last iteration of proposals by the Basel Committee on Banking Regulation. We start from the premise that the severity of assumptions should correspond to the robustness of the system they describe. Hence, in the formal description of physical systems, the level of assumptions can be much higher. It follows that every concept that is carried over from the natural sciences to economics must be checked for its plausibility in the new surroundings. Most of the probability theory has been developed for the analysis of physical systems and is based on the independent and identically distributed (i.i.d.) assumption. In Economics both parts of the i.i.d. assumption are inappropriate. However, only dependence has, so far, been weakened to a sufficient degree. In this paper, an appropriate class of non-stationary processes is used, and their law is tied to a formal object measuring representativeness. Subsequently, that data set is identified that on average minimizes the estimation error stemming from both, insufficient and non-representative, data. Applications are far reaching in a variety of fields. In the paper itself, we apply the results in order to analyze a paragraph in the Basel 3 framework on banking regulation with severe implications on financial stability. Beyond the realm of finance, other potential applications include the reproducibility crisis in the social sciences (but not in the natural sciences) and modeling limited understanding and learning behavior in economics.

Keywords : banking regulation, non-ergodicity, risk management, semimartingale modeling
Conference Title : ICFEMST 2018 : International Conference on Finance, Empirical Methods and Statistical Techniques
Conference Location : London, United Kingdom
Conference Dates : May 14-15, 2018