On-line analysis of count-valued time series by dynamic discount factors

Kaoru Irie, The University of Tokyo
Christopher Glynn, University of New Hampshire
Tevfik Aktekin, University of New Hampshire

Abstract

The on-line analysis of streaming data (i.e., monitoring and forecasting of time series) is highly demanded in the practice of modern statistical analysis for the important decision-making problems such as server maintenance, web advertisement and supply control. The time series of integer-valued observations, or counts, make this problem more complicated in developing non-Gaussian state space models that are analytically tractable for the data to be processed sequentially. A solution to this problem has been Poisson-gamma state space models with scaled-beta multiplicative innovations (Smith and Miller, 1986), which exploit the Poisson-gamma conjugacy and realize the sequential update of posterior and predictive distributions in a timely manner without relying on MCMC methods. However, as Chen et al. (2017) pointed out, this simple model often results in overfitting or over-smoothing of the posteriors due to its limited model flexibility controlled by a single parameter called discount factor. In this research, we consider a version of this class of models in which the discount factors are time-varying. The dynamic discount factor is expected to be small only when the model continues to fail to predict future observations accurately, so that the posterior can ``discount'' the past information, adjust the posterior location to the new observation and avoid over-smoothing. Otherwise, the value of discount factor remains large to exploit the accumulated information when the data process is stable, avoiding overfitting to the current data. As a solution to the computational feasibility under the time constraint, a variant of the simple auxiliary particle filtering/particle learning is tailored for this model to implement the sequential analysis; it is observed in application that the computed on-line posteriors do not suffer from the particle degeneracy. The proposed models and computational methodologies are applied to the access log data of FoxNews websites, which are observed every 30 seconds, to illustrate the quick and flexible adjustment of predictions in the presence of ``burst'' of accesses. In addition, the records of Uber demand in New York City, which consist of the pick-up calls aggregated in every 10 minutes, are analyzed to confirm that the predictive distributions are also able to track the moderately slow change of data trend.