A spatio-temporal model based on discrete latent variables for the analysis of COVID-19 incidence

Francesco Bartolucci[1]; Alessio Farcomeni[2]

[1] Department of Economics, University of Perugia (IT)
[2] Department of Economics and Finance, University of Rome «Tor Vergata» (IT)

Abstract:
We propose a Bayesian model based on discrete latent variables, which are spatially associated and time specific, for the analysis of incident cases of COVID-19 infections. The model may be seen as an extension of a hidden Markov model with covariates (Bartolucci et al., 2013). We assume that for each area the sequence of latent variables across time follows a Markov chain with initial and transition probabilities that also depend on the latent variables referred to the neighboring areas. Moreover, every area- and time-specific response variable follows a distribution belonging to the exponential family with mean depending on the corresponding covariates and regression parameters depending on the underlying latent state. The model is estimated by a Markov chain Monte Carlo algorithm based on a data augmentation scheme and which alternates two steps: (i) new values of the latent variables are drawn for each area and time occasion given the model parameters and the observed data; (ii) new values of the model parameters are updated given the latent variables and the observed data. The algorithm also includes suitable checks in order to face the label switching problem. As an illustration we analyze incident cases of COVID-19 collected in Italy at regional level for the period from February 24, 2020, to January 17, 2021, corresponding to 48 weeks, where we use a Poisson formulation with the number of swabs as an offset. Our model identifies a common trend based on splines of suitable order and knots and, for every week, it assigns each region to one among five distinct risk groups corresponding to the latent states. This number of latent states is selected on the basis of the data by the WAIC (Watanabe, 2010). A detailed description of the proposed approach and analyses may be found in Bartolucci and Farcomeni (2021).

Keywords:
Data augmentation; Hidden Markov models; Markov chain Monte Carlo; SARS-CoV-2; Swabs

References:

1. Bartolucci, F., Farcomeni, A., and Pennoni, F. (2013), *Latent Markov Models for Longitudinal Data*, Chapman and Hall/CRC press, Boca Raton, FL.
2. Bartolucci, F. and Farcomeni, A. (2021), A spatio-temporal model based on discrete latent variables for the analysis of COVID-19 incidence, *Spatial Statistics*, in press.
3. Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory, *Journal of Machine Learning Research* **11**, 3571-3594.