Properties of the Stochastic Approximation EM Algorithm with Mini-batch Sampling

Estelle Kuhn, Catherine Matias, Tabea Rebafka

Properties of the Stochastic Approximation EM Algorithm with Mini-batch Sampling

Estelle Kuhn[1], Catherine Matias[2] and Tabea Rebafka[2]

[1] MaIAGE INRA, Université Paris-Saclay, Jouy-en-Josas, France.
[2] Sorbonne Université, Université Paris Diderot, Centre National de la Recherche Scientifique, Laboratoire de Probabilités, Statistique et Modélisation, Paris, France.

Abstract

For models where the classical EM algorithm cannot be applied directly, stochastic variants such as Monte Carlo EM, Stochastic Approximation EM (SAEM) and Monte Carlo Markov Chain SAEM (MCMC-SAEM) exist. However, their computing time is very long when the sample size and hence the number of latent variables is large. As a solution mini-batch sampling has been proposed recently, which consists in using only a part of the observations and simulating only a portion of the latent variables at each iteration. Intuitively, when the so-called mini-batch size, that is the size of the data subset selected at every iteration, is small, the computing time is shortened, while the computed estimator may be less accurate.

In this talk, we propose a mini-batch version of the MCMC-SAEM algorithm, which is appropriate when the latent data cannot be simulated exactly from the conditional distribution, as for instance in nonlinear models or non-Gaussian models. As the underlying stochastic approximation procedure only requires the simulation of a single instance of the latent variable at every iteration, MCMC-SAEM is much more computing efficient than MCMC-EM. Nevertheless, when the dimension of the latent variables is huge, the sampling step can still be time-consuming and thus our mini-batch version is computationally more efficient than the original algorithm.

When the model belongs to the exponential family, we prove almost-sure convergence of the sequence of estimates generated by the mini-batch MCMC-SAEM algorithm as the number of iterations increases. Moreover, we provide results in the same regime that quantify the impact of the mini-batch size on the limit distribution of the estimator compared to the classical batch MCMC-SAEM algorithm. Simulation experiments and real data examples show that an appropriate choice of the mini-batch size results in an important speed-up of the convergence in nonlinear mixed effects models, frailty models and the stochastic block model.