Frontier of Statistical Physics and Information Processing 2013

Learning Coefficient for Bayesian Learning and Markov Chain Monte Carlo Method

Kenji Nagata[1]
1. Graduate School of Frontier Sciences, The University of Tokyo

Stochastic learning models with hierarchical structures such as neural networks, mixture of gaussians and hidden Markov models are difficult to clarify the generalization performance because the asymptotic normality does not hold. In recent studies, the algebraic geometrical method has been established, and clarify the learning coefficient, which characterize the generalization performance, for Bayesian estimation for some singular learning machines. However, applying the algebraic geometrical method is not generally easy, and there are many singular learning machines not to clarify the learning coefficient.

On the other hand, the Markov chain Monte Carlo (MCMC) method is well used for the Bayesian estimation. The MCMC method is the algorithm to sample from arbitrary probability distribution, and is suitable for the Bayesian estimation which requires the expectation over the Bayesian posterior distribution.

When we implement the MCMC method, the step size for the Metropolis algorithm and the temperature setting for the exchange Monte Carlo method have important effects on the efficiency of the algorithm. For the criteria of these settings, we analyze the average of the acceptance rate for the Metropolis algorithm, and the average of the exchange rate for the exchange Monte Carlo method. These analyses clarify that those average values have relationship with the learning coefficient for Bayesian estimation, and that the learning coefficient can be calculated by the behavior of the MCMC method.

In this presentation, we introduce the analytic results for the MCMC method, and summarize the relationship between the behavior of the MCMC method and the learning coefficient for Bayesian estimation.

[1] Nagata, K. and Watanabe, S.: Asymptotic behavior of exchange ratio in exchange Monte Carlo method. Neural Networks, vol.21, pp.980-988, 2008.
[2] Nagata, K. Watanabe, S. and Okada, M.: Analysis of acceptance rate for Metropolis algorithm and relationship between Bayesian learning, International Society on Bayesian Analysis (ISBA) 2012 Conference, 2012.