A convergence diagnostic for assessment and numerical accuracy of multivariate outputs of Markov chain Monte Carlo (MCMC) methods.

Esmail Amiri.

Department of Mathematics & Statistics, Imam Khomeini International University,

P.O. Box 288

Ghazvin, IRAN.

e_ amiri @ yahoo.com

1. Introduction

Markov chain Monte Carlo(MCMC) is a well known computation method in Bayesian inference. The method constitutes a very attractive tool for the solution of many awkward Bayesian multiple integration problems. The essential idea of MCMC is to simulate a Markov chain whose equilibrium distribution is the desired posterior distribution, at time t the realization of the Markov chain(MC) can be considered as a sample from the target distribution. Time t is called stopping time of MCMC. There is a great debate concerning the stopping time. In practice this time can not be easily identified. One approach to identify t is named convergence diagnostic(CD). There are many CD methods in the MCMC literature which can be categorized into two main groups. One group only considers convergence of the MC for a few scalar functionals of data(convergence to the marginal distributions of the target), while other group presents those CD methods which consider convergence of MC for many functionals of the data(convergence to the joint distribution of the target). When the inference is based upon only a few scalar functionals of the data the convergence of the MC to the joint target distribution can not be guaranteed. This problem is pointed in Gelfand(1992). We take the last group approach and apply the multivariate two sample Kolmogorov-Smirnov (K-S) test to assess the convergence of MC. Of course the simulated data is correlated. In order to reduce the correlation, we take the batching procedure which is advocated in Ripley(1987).

In section 2 batching procedure and preliminary and final statistical tests are introduced. Section 3 is devoted to the steps of proposed CD method.

2. Batching procedure and statistical tests.

Suppose X_t is the output of a MC at time t, where $X_t \in E \subset \mathbf{R^n}$ and $\pi(X)$ is the target distribution. With MCMC simulation we expect $X_t \xrightarrow[t\to\infty]{d} X \sim \pi(X)$. Consecutive X_t are correlated, therefore ordinary multivariate statistical analysis are not applicable to inference from the output of MC. We take the advise of Geyer(1992) and ignore 1% of the simulated data up to time t, then we follow the batching procedure in Ripley(1987), and divide the data into k successive batches of length m with batch means $B_1, B_2, ..., B_k$. The correlation between batches will decline subject to choosing m and k large enough. We recommend conducting a test (because of the space limitation we have not been able to present proofs and details) to insure shrinkage of serial dependence in(B_i). If the result of this test is positive we can move to the final step and perform multivariate two sample K-S test, otherwise the sample size is not enough.

If the preliminary test is successful, then $B_i,\ i = 1, ..., k$ are considered as samples coming from a distribution which we do not make any assumption concerning this. If we divide $B_i,\ i = 1, ..., k$ into two halves, each half should have the same distribution. This can be tested by conducting a multivariate two sample nonparametric K-S test introduced in Bickel(1969). If we define $\|.\|$ as supremum norm, then the test statistics is $\{k_1 k_2 (k_1 + k_2)^{-1} \|F_{k_1} - G_{k_2}\|$ where F_{k_1} and G_{k_2} are empirical cumulative distributions of the first half and second half of B_i's respictively, and $k_1 + k_2 = k$.

If K-S test shows no difference between the distribution of two halves, chain is in equiblirium, otherwise the sample size is not enough.

3. Steps of CD

Step1: Ignore 1% of simulated data.

Step2: Divide the data into k batches of length m.

Step3: Calculate the batch means.

Step4: Conduct the preliminary test

Step5: If step 4 is successful perform K-S test.

REFERENCES

Bickel, P. J.(1969). A distribution free version of the Smirnov two sample test in the p-variate case. *The annals of Mathematical Statistics.*, **40(1)**, 1-23.

Gelfand, A. E. (1992). Comment on papers by Geyer and Gelman and Rubin. *Statistical Science*,**7(4)**,486-487.

Geyer, C.J. (1992). Practical Markov Chain Monte Carlo *Statistical Science*,**7(4)**,473-511.

Ripley, B. D. (1987). *Stochastic Simulation, chap. 6. Wiley. New York.*

RESUME

Une diagnostique de convergence pour evaluer et détérminer la précision numérique des résultats multivariés des méthodes Chaines fde Markov Monte Carlo.

Nous déterminons la convergence de la Chaine Markov Monte Carlo à la distribution visée dans un contexte multivarié en appliquant le test Kolmogorov-Smirnov multivarié.