Applications of Chatterjee's Correlation in MCMC

Vivek Kumar Singh

BS/MTH/190988

Supervisor: Dr. Dootika Vats

भारतीय प्रौद्योगिकी संस्थान कानपुर

INDIAN INSTITUTE OF TECHNOLOGY KANPUR

April 11, 2022

Contents

1 Introduction **2**

2 Preliminaries **3**

2.1 Introduction to Markov chain Monte Carlo 3

2.2 Basic Theorems from Measure Theory 3

3 Problems with Pearson correlation coefficent **4**

4 Chatterjee's autocorrelation function **5**

5 Sketch of Proof of consistency **7**

6 Some simulation plots and conclusion **8**

Chapter 1

Introduction

Chapter 2

Preliminaries

2.1 Introduction to Markov chain Monte Carlo

Definition 2.1. Continuous time markov chain

Definition 2.2. markov transition kernel

Definition 2.3. time homogeneity

Definition 2.4. stationarity

Definition 2.5. ergodicity

Definition 2.6. time reversibility

Definition 2.7. total variation norm

Definition 2.8. chapman Kolmogorov

2.2 Basic Theorems from Measure Theory

Theorem 2.9. Lebesgue's DCT

Chapter 3

Problems with Pearson correlation coefficent

Definition 3.1. Pearson correlation coefficient measures the linear correlation of two sets of data. Given a pair of random variables (X, Y), pearson correlation ρ is defined as

$$\rho = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}\left[X\right]} \cdot \sqrt{\mathrm{Var}\left[Y\right]}}$$

Pearson's correlation coefficent is very powerful in detecting monotone relations, has a well developed asymptotic theory. The autocorrelation function that we look at in MCMC is also defined using this.

There are two most common problems with this coefficent.

1. Firstly, we would like that the correlation would be close to its maximum value if and only if one variable looks like a noiseless function of the other variable. This is not the case for the Pearson's Coefficient as it is close to ±1 iff one variable is a noiseless *linear* function.

2. Second, we would like the correlation to be close to its minimum value if and only if both the variables are independent of each other. In the case of the Pearson's correlation, it is zero when the variables are independent but the converse is not always true.

Chapter 4

Chatterjee's autocorrelation function

Sourav Chatterjee proposed a new correlation coefficient in his [add reference]. This coefficient is (a) as simple as the classical ones, (b) is a consistent estimator of some measure of dependence which is 0 iff the variables are independent, and 1 iff one is a measurable function of the other, and (c) has a simple asymptotic theory under the hypothesis of independence, like the classical coefficients.

Let (X, Y) be a pair of random variables, where Y is not a constant (for our purposes, both X and Y are continuous). Let $\{(X_i, Y_i)\}_{i=1}^n$ be i.i.d. pairs following the same distribution as (X, Y).

1. The case when $X_i's$ and $Y_i's$ have no ties. Rearrange the data as $(X_{(1)}, Y_{(1)}), \ldots, (X_{(n)}, Y_{(n)})$ such that $X_{(1)} < \cdots < X_{(n)}$. Let r_i be the rank of $Y_{(i)}$, i.e. the number of j such that $Y_{(j)} \leq Y_{(i)}$. Then the correlation coefficient ξ_n is defined to be

$$\xi_n(X, Y) := 1 - \frac{3\sum_{i=1}^{n-1} |r_{i+1} - r_i|}{n^2 - 1}$$

.

2. In the case of ties. If there are ties in $X_i's$, choose an increasing arrangement as follows and break ties uniformly at random. Let r_i defined as

above, and define l_i to be the number of j such that $Y_{(j)} \geq Y_{(i)}$. Define

$$\xi_n(X,Y) := 1 - \frac{n\sum_{i=1}^{n-1}|r_{i+1}-r_i|}{2\sum_{i=1}^{n-1}l_i(n-l_i)}$$

. When there are no ties among the $Y_i's, l_1, \ldots, l_n$ is just a permutation of $1, \ldots, n$ and the denominator is just $n(n^2-1)/3$, which reduces to the definition in the no ties case.

Theorem 4.1. If Y is not almost surely a constant, then as $n \to \infty$, $\xi_n(X,Y)$ converges almost surely to the deterministic limit

$$\xi(X,Y) := \frac{\int Var(\mathbb{E}(1_{\{Y\geq t\}}|X))d\mu(t)}{\int Var(1_{\{Y\geq t\}})d\mu(t)}$$

where μ is the pdf of Y. This limit belongs to the interval $[0,1]$. It is 0 iff X and Y are independent, and it is 1 iff there is a measurable function $f : \mathbb{R} \to \mathbb{R}$ such that $Y = f(X)$ almost surely.

Chapter 5

Sketch of Proof of consistency

Chapter 6

Some simulation plots and conclusion