Applications of Chatterjee's Correlation in MCMC
(UGP, 21-22 Even Semester)

Vivek Kumar Singh, Dootika Vats

Department of Mathematics and Statistics
Indian Institute of Technology, Kanpur

April 17, 2022

Table of Contents

1. Markov chain Monte Carlo
2. Pearson's Correlation Coefficient
3. Chatterjee's Correlation Coefficient
4. Chatterjee's Autocorrelation Function
5. Proof of consistency of estimator
6. Simulations

Markov Chain Monte Carlo

Definition

A Markov chain is a discrete time stochastic process $X_1, X_2, \ldots$ such that the next state of the process depends solely on the present state, i.e.

$$\Pr(X_{n+1}|X_1, \ldots, X_n) = \Pr(X_{n+1}|X_n).$$

- A Markov chain can be specified by two things
  1. The initial distribution, i.e. the marginal of X_1.
  2. The transition probabilites, i.e. the conditional distribution of $X_{n+1}|X_n$

Markov chain Monte Carlo

Definition

A Markov chain is said to be **stationary** if the marginal of X_n is independent of n. This invariant distribution is called the stationary distribution.

Definition

A Markov chain is **ergodic** if the distribution of X_n converges to the invariant distribution.

Definition

A stationary Markov chain is **time-reversible** w.r.t. the stationary distribution if X_n and X_{n+1} are exchangable.

Pearson's Correlation Coefficient

- Pearson's Correlation Coefficient measures the linear correlation between two sets of data.

Definition

Given random variables X, Y, Pearson Correlation ρ is defined as

$$\rho = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)} \cdot \sqrt{\mathrm{Var}(Y)}}.$$

- Pearson Correlation is very powerful in detecting monotone relations and has a well developed asymptotic theory.

Problems with Pearson Correlation

- There are some problems with the Pearson Correlation.

- We would like the correlation to be close to its maximum value iff one variable is a function of the other. In Pearson's case, it is equal to ± 1 iff the variables are linearly dependent.

- We would also like the correlation to be 0 iff the variables are independent. If the variables are independent, Pearson correlation is indeed 0, but the converse is not always true.

- In order to solve these problems, Chatterjee came up with a new measure of dependence in [?]. It overcomes the above mentioned drawbacks, and has a computationally efficient and consistent estimator.

Definition

Given random variables X, Y, where is Y is not a constant, Chatterjee correlation ξ is defined as

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

where μ is the law of Y.

Consistent Estimator of ξ

- Let $\{(X_i, Y_i)\}_{i=1}^{n}$ be i.i.d. pairs following the same distribution as (X, Y). 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

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

Theorem

If Y is not almost surely a constant, then as $n \to \infty$, $\xi_n(X, Y)$ converges almost surely to $\xi(X, Y)$.

Properties of ξ

- $\xi(X, Y) \in [0, 1]$
- $\xi(X, Y) = 0$ if and only if X and Y are independent.
- $\xi(X, Y) = 1$ if and only if atleast one of X and Y is a measurable function of the other.
- ξ is not symmetric in X, Y. This is intentional and useful as we might want to study if Y is a measurable function of X, or X is a measurable function of Y. To get a symmetric coefficient, it suffices to consider $\max(\xi(X, Y), \xi(Y, X))$.
- ξ_n is based on ranks, and for the same reason, it can be computed in $O(n \log n)$.

Chatterjee's Autocorrelation Function

- We present a new autocorrelation function (ACF) using Chatterjee's correlation coefficient.

Definition

Let $X_1, X_2, \ldots$ be a stationary, time homogeneous Markov chain with stationary distribution π. We define the new lag-k ACF as follows

$$\gamma'_{k,n} := \xi(X_n, X_{n+k}).$$

$\gamma'_{k,n}$ is independent of n

Theorem

$\gamma_k = Cov(X_n, X_{n+k})$ is independent of n.

Theorem

$\gamma'_{k,n} = \xi(X_n, X_{n+k})$ is independent of n, where n and k are in $\mathbb{N}$.

- From now on, we'll denote $\gamma'_{k,n}$ by γ'_k

Symmetricity of γ'_k

Theorem

$\xi(X_n, X_{n+k}) = \xi(X_{n+k}, X_n)$ *for time reversal Markov chains for any* $n, k \in \mathbb{N}$.

Convergence of γ'_k

Theorem

$\lim_{n\to\infty} \xi(X_1, X_n) = 0$ *for an Ergodic Markov chain.*

Proof of consistency of estimator

- Chatterjee presented the proof of consistency of the estimator in [**?**], where the samples drawn from (X, Y) are i.i.d.
- In our case of stationary Markov chains, we have correlated but identically distributed draws.
- We aim to prove the consistency of the estimator for our case as well.

Proof of consistency of estimator

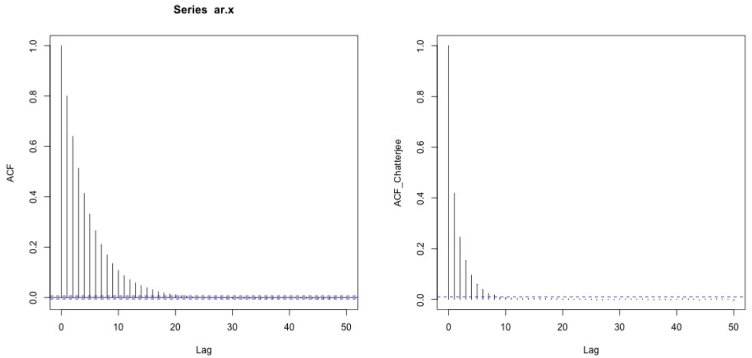
Theorem

Let $X_1, X_2, \dots$ be a stationary time-homogeneous Markov chain with stationary distribution μ. Then $\xi_n(X, Y)$ estimated using the draws from the Markov chain converge to $\xi(X, Y)$ as $n \to \infty$, where X, Y are any two time points in the chain.

- We presented some ideas and a pathway for proving it but could not complete it and is left as future work.

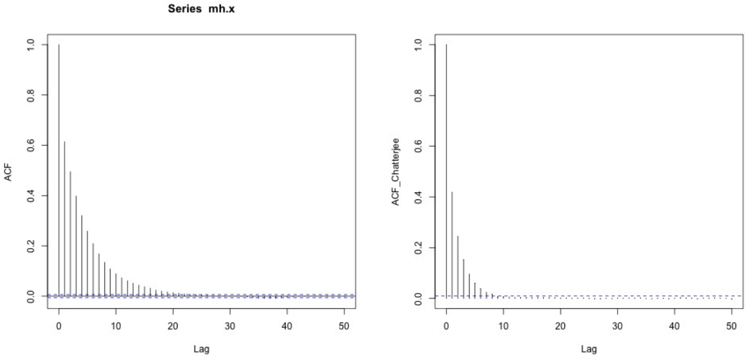
Simulations

- AR(1) process with $\rho = 0.8$.



Simulations

- Metropolis-Hastings algorithm with initial distribution as Exp(0.01) and target distribution $\mathcal{N}(0, 1)$.



Thank you