probability theory \rightarrow Markov chains

Sampling from Distributions with Constraints

Many hard problems in combinatorics and other fields come from asking questions with nontrivial constraints. As we have seen, often times the most practical solution is just to simulate such problems with MCMC methods. As an example of a problem where a constraint can make brute force sampling highly inefficient, consider the following set-up.¹ We can define a stochastic process Y_t by

$$Y_{t+1} = \rho Y_t + W_t$$

with $W_t \sim \mathcal{N}(0, 1)$ iid random variables, and ρ some real constant. Suppose we wanted to sample from the distribution of (Y_1, Y_1, \ldots, Y_L) subject to the constraint that

$$Y_1 > Y_2 > \dots > Y_L.$$

(a) Code up a brute force algorithm to sample from this distribution by doing the obvious thing, simply generating Y_t sequences according to the rule and discarding them until we find one subject to the constraint. Choosing L = 12, how many tries will your program take, on average, to find a correct sequence? Determine the answer numerically, of course.

This is a very inefficient algorithm. In this part, you will show that we can in fact use the Gibbs sampler to be much more efficient about it.

- (b) Show that $p(Y_t|\cdots, Y_{t-1}, Y_{t+1}, \ldots) = p(Y_t|Y_{t-1}, Y_{t+1}).$
- (c) Show that $p(Y_t|Y_{t-1}, Y_{t+1})$ is a Gaussian PDF with

$$E(Y_t|Y_{t-1}, Y_{t+1}) = \frac{\rho}{1+\rho^2} \left(Y_{t-1} + Y_{t+1} \right), \quad Var(Y_t|Y_{t-1}, Y_{t+1}) = 1 - \frac{\rho}{1+\rho^2}.$$

(d) How would you use the Gibbs sampler to sample from this distribution? Code up the Gibbs sampler, and comment on its performance relative to the previous method.

¹Thanks to Emmanuel Candès for giving this problem to me in a class at Stanford.