Monte Carlo Methods for Computation and Simulation (048715)

Problem Set 3: MCMC

Submission: June 24

Markov Chains:

1. Consider an irreducible Markov chain with *N* states.
   a. Show that for *N*=2 the chain is always reversible.
   b. Give an example of a Markov chain that is not reversible, with the minimal possible *N*.

2. **Random walk with reflecting boundaries**: Consider a Markov chain over the state space $\{-m,-m+1,\dots,m\}$, and transition probabilities (p_{ij}) given by
$$p_{i,\min\{i+1,m\}} = 1 - p_{i,\max\{i-1,-m\}} = \alpha \in (0,1)$$
   a. Is this chain irreducible? Periodic?
   b. Show that the chain is reversible, and compute the stationary distribution.
   c. ***Simulation****:* For $m=1$, choose some α and initial conditions, and compute numerically $\pi^{(t)}$. Plot all three components and verify convergence to the stationary distribution.

3. **AR(1):** Consider an order-1 auto-regressive model with parameter a:
$$X_{t+1} = aX_t + w_t\,, \quad t = 0,1,2,\dots$$
Here (w_t) is an *iid* sequence with $w_t \sim N(0,\sigma^2)$, $\sigma > 0$, $|a| < 1$, and $X_t \in \mathbb{R}$.
   a. Write down the transition function $f(y\,|\,x)$, namely the pdf of X_{t+1} conditioned on $X_t = x$.
   b. Show that (X_t) is a reversible Markov chain with stationary distribution $\pi = N(0,\sigma_1^2)$, and compute σ_1^2.
   c. How will your answers change if $|a| > 1$? Explain briefly.

Metropolis-Hastings:

4. **The Metropolis-Hasting-Green (MHG) Algorithm:** The MHG algorithm is a generalization of MH that allows state-dependent mixing of several transition matrices. Let $\{Q_i = (q_i(x,y)),\ i \in I\}$ be a finite collection of transition functions over the same (finite) state space $\mathbb{X}$.
For each state x, let $\beta(x) = (\beta_i(x))_{i\in I}$ be a probability vector.
Each step of the MHG algorithm proceeds as follows:
   1. Starting from $X_t = x$, choose an index i with probability $\beta_i(x)$.
   2. Sample *Y* from $q_i(x,\cdot)$.
   3. Set $X_{t+1} = Y$ with probability $\alpha_i(x)$, and $X_{t+1} = x$ otherwise, where
$$\alpha_i(x,y) = \min\{1, \rho_i(x,y)\}, \qquad \rho_i(x,y) = \frac{f(y)\beta_i(y)q_i(y,x)}{f(x)\beta_i(x)q_i(x,y)}$$

a. Write down an expression for the transition probabilities $p(X_{t+1} = y \mid X_t = x)$.

b. Show that f is a stationary distribution of the Markov chain (X_t).

5. **Sampling Spanning Trees:** Let *G=(V,E)* be an undirected and fully connected graph. A simple MCMC algorithm to sample uniformly from the set of spanning trees of *G* is the following: Start with some spanning tree; add uniformly-at-random some edge from *G* (so that one cycle forms); remove uniformly-at-random some link from this cycle; repeat.

   Suppose now that the graph is positively weighted, i.e., each edge $e \in E$ has some cost $c_e > 0$. The weight of any sub-graph of *G* is the sum of costs of its edges.

   a. Suggest an MCMC algorithm that samples from the set of spanning trees of *G*, with a probability that is proportional to their weights.

   b. Suppose we wish to estimate the average weight of a spanning tree. Suggest two variants of the above MCMC algorithms that provide this estimate.

6. ***Simulation* (spanning trees):** Implement the algorithms of Problem 5 on some small (but non-trivial) weighted graph of your choosing:

   a. Implement the algorithm in 5a. Show graphs of the empirical frequencies and verify convergence to the correct values.

   b. Implement the two algorithms from 5b. Examine and compare their convergence visually.

   c. Use the batch method (with 20 batches) to estimate the standard deviation of these two algorithms. Estimate the number of samples required to obtain a 1% accuracy (with 95% confidence).

The Gibbs Sampler:

8. Show that the Random Sweep Gibbs Sampler induces a reversible Markov chain.

9. **Bivariate Normal Sampling**. Apply the Gibbs sampler to sampling from the bivariate normal distribution, $X = (x_1, x_2)^T \sim N(\mu, \Sigma)$.

   a. Compute the conditional distributions $f_i(x_i \mid x_j)$, and write down the systematic Gibbs sampling algorithm for this problem.

   ***Simulation**:* Apply the Gibbs sampler with parameters $\mu = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$, $\Sigma = \begin{pmatrix} 0.8 & b \\ b & 1 \end{pmatrix}$, and two values of b: $b = 0.2$ and $b = 0.85$. Adjust the number of samples (and other parameters) according to the simulation results to obtain meaningful conclusions.

   b. Plot the empirical covariance between the components of X as a function of time, verify convergence to ρ and compare convergence rates.

   c. A common measure for the mixing properties of a sampled Markov chain is the autocorrelation function, namely $R(k) = \mathrm{cov}(X(t), X(t+k))$ as a function of $k \geq 0$. Estimate and plot the autocorrelation function for the first component x_1 of X. Compare and discuss briefly.