Assignment 02

CMPE 58N, Monte Carlo Methods

Department of Computer Engineering, Boğaziçi University, Istanbul, Turkey

Instructor: A. T. Cemgil

Due: 16 Nov 2009, Mon, 9:00.

1. **(Rejection Sampling)** Consider a symmetric triangle distribution p on $[-1, 1]$. Using a uniform distribution as the proposal q, compute the efficiency of rejection sampling, i.e., how many samples from p can be generated per sample from q?

2. **(Importance sampling)** Consider the target distribution

$$p(x) = \mathcal{N}(x; m, S)$$

when using the proposal

$$q(x) = \mathcal{N}(x; \mu, \Sigma)$$

   (a) Derive the weight function for the target distribution

   (b) Derive the analytic expression for the variance of importance weights

3. (**Metropolis and Gibbs**) x_1 and x_2 are two discrete random variables taking values in $\{-1, 1\}$. Suppose we have the joint distribution $p(x_1, x_2) \propto \exp(\theta x_1 x_2)$.

   Suppose we implement a Metropolis algorithm to sample from this target distribution with the following proposal technique: Given the current configuration $x^{(n)} = (x_1^{(n)}, x_2^{(n)})$, for each n, we choose an index $i^{(n)} \in \{1, 2\}$ randomly with probability 0.5 and flip the sign of $x_{i^{(n)}}$.

   (a) *(2 pts)* Write down the state transition diagram of the proposal distribution and indicate the state transition probabilities,

   (b) *(2 pts)* Find an expression for the acceptance probability as a function of θ,

   (c) *(2 pts)* Write the pseudocode for the Metropolis sampler,

   (d) *(2 pts)* Write down the state transition diagram of the transition Kernel T_M of this Metropolis algorithm and indicate the transition probabilities,

   (e) *(2 pts)* Verify if detailed balance condition is satisfied by this particular Metropolis algorithm for all values of θ.

   (f) *(2 pts)* Write down the pseudocode for a deterministic scan Gibbs sampler.

   (g) *(2 pts)* Write down an expression for the Gibbs transition Kernel T_G in terms of θ.

(h) *(2 pts)* Verify detailed balance is satisfied the Gibbs transition Kernel T_G for all values of θ.

4. **(Linear Matrix Inequalities)** This example generalises sampling from circular regions. A Linear Matrix inequality is an expression of the following form

$$0 \preceq F_0 + \sum_{i=1}^{M} x_i F_i$$

Here, F_i for $i = 0 \ldots M$ are symmetric matrices in $\mathbb{R}^{N \times N}$ and $x_1 \ldots x_M$ are real scalars and $x = (x_1 \ldots x_M)^\top$. $0 \preceq F$ means that F is *positive semidefinite*:

$$0 \leq z^\top F z$$

for all $z \in \mathbb{R}^N$. A symmetric matrix is positive semidefinite if all eigenvalues are nonnegative. To see this, consider the eigendecomposition (that always exists when F is symmetric)

$$F = Q^\top \Lambda Q$$

where $Q^\top Q = I$ and $\Lambda = \mathbf{diag}(\lambda_1, \ldots, \lambda_N)$ and define $z^\top F z = (Qz)^\top \Lambda (Qz)$.

(a) Consider a LMI, where $M = 2$

$$0 \preceq F_0 + x_1 F_1 + x_2 F_2$$

Develop a MH method for sampling uniformly from the feasible region of the LMI given a isotropic Gaussian proposal.

(b) Implement and test your methods for $N = 7$ and $M = 2$

5. **(Rapid mixing Markov Chain)** Consider a distribution on a discrete state space

$$\mathcal{X} = \{1, \ldots 20\}$$

and consider a distribution with $x \in \mathcal{X}$ and

$$p(x) \quad \propto \quad x^3$$

We will design a Metropolis-Hastings algorithm with the following proposal

$$\begin{aligned} u &\sim \mathcal{U}(0,1) \\ x'|x &= \begin{cases} (x \mod 20) + 1 & u < \epsilon \\ (x - 1 \mod 20) & u \geq \epsilon \end{cases} \end{aligned}$$

where ϵ is a given parameter of the algorithm. Note that $-1 \mod 20 = 19$.

(a) Write down the proposal in transition matrix form. What does this proposal do for $\epsilon = 0.5$?

(b) Write down the acceptance probability in matrix form.

(c) Write down the MH transition kernel in matrix form and evaluate it numerically for a given ϵ.

(d) By numerically evaluating the second largest eigenvalue, find the optimum ϵ in terms of reducing mixing time. (Show your source code)

(e) Comment about the solution. Why is the optimum different from 0.5 ?

(f) Does the stationary distribution depend on ϵ? Why or why not?

6. **(Estimation of Gaussians)** Consider the following model

$$\begin{aligned} \beta &\sim \mathcal{G}(\beta; \nu, 1) \\ \mu &\sim \mathcal{N}(\mu; 0, 1000) \\ x_i &\sim \mathcal{N}(x_i; \mu, \beta^{-1}) \end{aligned}$$

for $i = 1 \dots N$. $X = \{x_1, \dots, x_N\}$. Furthermore, use

$$\begin{aligned} \mathcal{G}(\beta; \nu, s) &\equiv \exp((\nu - 1)\log\beta - s\beta - \log\Gamma(\nu) + \nu\log s) \\ \mathcal{N}(x; m, s) &= \exp(-\frac{1}{2s}(x-m)^2 - \frac{1}{2}\log(2\pi s)) \end{aligned}$$

(a) (4) Derive the full conditionals and give pseudocode of a Gibbs sampler to sample from

$$p(\mu, \beta | X, \nu)$$

(b) (4) Derive the expression for the acceptance probability of a Metropolis-Hastings algorithm to sample from

$$p(\nu | X, \mu, \beta)$$

Use the following proposal $q = \mathcal{G}(\nu; a, 1)$

(c) (4) Combine (a) and (b) and derive a MH algorithm to sample from

$$p(\nu, \mu, \beta | X)$$