STA3431H (Monte Carlo Methods), Winter 2010

Homework #2

Due: In class by 2:10 p.m. **sharp** on Monday March 22.

NOTES:

- **Late homeworks, even by one minute, will be penalised!**
- Homework assignments are to be solved by each student individually; you may discuss assignments in general terms with other students, but you must solve it on your own, including doing all of your own computing and writing.
- Your homework should be neatly and clearly presented and explained. Computer programs should include complete source code and program output, with comments etc. so they are easy to follow.
- Reminders: project due March 29 at 2:10pm; test April 5 at 2:10 in SS2105.

The assignment:

1. Consider an independence sampler algorithm on $\mathcal{X} = (1, \infty)$, where $\pi(x) = 3\, x^{-4}$ and $q(x) = r\, x^{-r-1}$ for some choice of $r > 0$, with identity functional $h(x) = x$.

(a) For what value of r will the algorithm provide i.i.d. samples?

(b) For what values of r will the sampler be geometrically ergodic?

(c) For $r = 1/20$, find a number n such that $D(x, n) < 0.01$ for all $x \in \mathcal{X}$.

(d) Run the algorithm (on a computer) for the cases $r = 1/20$ and $r = 12$, each with $M = 10^5$ and $B = 10^4$, to estimate $\mathbf{E}_\pi(h)$. [Hint: if $U \sim$ Uniform$[0, 1]$, then what is the density of $U^{-1/r}$?]

(e) For both cases, use repeated runs (from an overdispersed starting distribution) to approximate the standard error of the estimate.

(f) For both cases, compute the standard error factor "varfact", and hence obtain different approximations of the standard error of the estimate.

(g) Compare the results of parts (e) and (f).

(h) Use the results of parts (e) and (f) to compare the two cases, to determine (with explanation) which case leads to a "better" sampling algorithm.

2. Let $\mathcal{X} = \mathbf{R}$, and let $\pi(x) = c\,g(x)$, where $g(x) = e^{-|x|/10}(1 + \cos(x)\sin(x^3))$, and let $h(x) = x + x^2$. With appropriate choice of M and B and σ and starting distribution $\mathcal{L}(X_0)$, estimate $\mathbf{E}_\pi(h)$ in each of two different ways:

(a) With a usual random-walk Metropolis algorithm for π, with usual proposals $Y_n \sim N(X_{n-1}, \sigma^2)$.

(b) With a Langevin (Metropolis-Hastings) algorithm with proposals $Y_n \sim N(X_{n-1} + \frac{1}{2}\,\sigma^2\, g'(X_{n-1}) \,/\, g(X_{n-1}),\, \sigma^2)$. [Note: Here $g'(x)$ is the usual derivative of g, and should be computed analytically by you in advance and entered into your program.]

(c) Compare the two algorithms and discuss which one is "better".

3. Consider the standard variance components model described in lecture, with $K = 6$ and $J_i \equiv 5$, and $\{Y_{ij}\}$ again the famous "dyestuff" data, but now with prior values $a_1 = a_2 = b_1 = b_2 = 1$, $a_3 = 0$, and $b_3 = 4$. Estimate (as best as you can, together with standard errors) the posterior mean of W/V, in each of three ways:

(a) With a random-walk Metropolis algorithm.

(b) With a Metropolis-within-Gibbs algorithm.

(c) With a Gibbs sampler. [Note: first <u>derive</u> from scratch the conditional distributions claimed in class.]

(d) Finally, discuss the relative merits of all three algorithms for this example.

4. Consider the homerun baseball data in the file "Rhomerun", giving the number of homeruns H_i and number of attempts (at-bats) A_i for players $1 \le i \le 12$. Consider the A_i to be fixed, known constants, and the H_i to be observed data. Assume that $H_i \sim \text{Binomial}(A_i, \theta_i)$ (cond. ind.), where $\theta_i \sim \text{Beta}(1001,\ 1 + 1000\,S)$ (cond. ind.) are unknown. Finally, put a prior $S \sim \text{Poisson}(3)$ on S (note: S is integer-valued).

(a) Specify (up to a normalising constant) the joint posterior distribution of $S, \theta_1, \ldots, \theta_{12}$.

(b) Run at least one MCMC algorithm of your choice for this posterior distribution, to estimate (as best as you can, together with standard errors), the posterior means of each of the 3 variables S, θ_1, θ_2.

(c) For $i = 1, 2$, compare the estimated posterior mean of θ_i to the value H_i/A_i. Are they different, and if so, how and why?