Statistics 202C. Spring 2012. Homework 3.

Due: Wednesday 21/May. 2012.

Question 1. Metropolis-Hastings

Describe the Metropolis-Hastings sampling algorithm. What are the proposal and acceptance probabilities? What is detailed balance? Show that Metropolis-Hastings obeys detailed balance. Why does this imply that samples from Metropolis-Hastings will eventually converge to samples from the target distribution?

Implement Metropolis-Hastings on an Ising spin model where each x_i takes values ± 1. (You can choose the proposal probability).

$$\pi(x) = \frac{1}{Z} e^{\mu(x_1 x_2 + x_2 x_3 + \ldots + x_{d-1} x_d)}.$$

Set $d = 10$, $\mu = 1$ and then $\mu = 2$.

Let the *magnetization* be $M = (1/d)(x_1 + \ldots + x_d)$. Plot the autocorrelation of the magnetization as a function of the lag. Discuss how this differs for $\mu = 1$ and $\mu = 2$.

Question 2. Multiple-Try Metropolis-Hastings

Describe multiple-try Metropolis-Hastings (MTMH). What are the advantages of using it compared to standard Metropolis-Hastings?

Let $\pi(x, y) = \sum_{i=1}^{3} \rho_i \frac{1}{2\pi\sigma_i^2} e^{-\{(x-x_i)^2 + (y-y_i)^2\}/(2\sigma_i^2)}$, where $(\rho_1, \rho_2, \rho_3) = (0.4, 0.3, 0.3)$, $(\sigma_1, \sigma_2, \sigma_3) = (1.0, 2.0, 2.0)$, $(x_1, x_2, x_3) = (1.0, 2.0, 1.0)$, and $(y_1, y_2, y_3) = (0.0, 0.0, 1.0)$.

Sketch the probability distribution $\pi(x, y)$ as a function of (x, y).

Use MTMH to sample from π (with $\lambda(x, y) = 1$). Use the proposal probability $T((x, y), (x', y'))$ to be the uniform distribution in the disc $\sqrt{(x' - x)^2 + (y' - y)^2} < 0.5$. Perform this with the number of trial proposals $k = 5$ and $k = 10$. Plot the mean of the samples for both cases.

Question 3. Gibbs Sampling

Describe the Gibbs sampler.

Calculate the conditional distributions for the Ising Spin model:

$$\pi(x) = \frac{1}{Z} e^{\mu(x_1 x_2 + x_2 x_3 + \ldots + x_{d-1} x_d)}.$$

Set $d = 10$, $\mu = 1$ and then $\mu = 2$.

Calculate the magnetization $M = (1/d) \sum_{i=1}^{d} x_i$. Plot the autocorrelation function for both cases ($\mu = 1$ and $\mu = 2$).

How do these results compare to the Metropolis-Hastings algorithm in question 1?

Question 4. Data Augmentation

Describe the Data Augmentation algorithm in 1-D when the data $\{x^i\}$ is generated by a mixture of two Gaussian distributions. Let $\{V^i\}$ be the indicator variables specifying which Gaussian generated the data. Treat the $\{V^i\}$ as missing data. The full distributions are $P(V^i) = e^{V^i \log \alpha + (1-V^i) \log(1-\alpha)}$, $P(\mu_1, \mu_2) = \frac{1}{2\pi\sigma_m^2} e^{-(\mu_1-\alpha_1)^2/(2\sigma_m^2)} e^{-(\mu_2-\alpha_2)^2/(2\sigma_m^2)}$, $P(x^i | V^i, \mu_1, \mu_2) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(x^i - V^i \mu_1 - (1-V^i)\mu_2)^2/(2\sigma^2)}$.

Let $\alpha = 0.5$, $\alpha_1 = 2.0$, $\alpha_2 = 6.0$, $\sigma_m = 10.0$, $\sigma = 2.0$.

Let the data be $0.8, 1.1, 3.4, 3.5, 8.9, 9.8, 7.3, 2.3, 10.8, 4.0$. Implement the Data Augmentation algorithm to obtain 10 i.i.d. samples of the $\{V^i\}, \mu_1, \mu_2$. Hence estimate μ_1, μ_2.