Petter Mostad
Applied Mathematics and Statistics
Chalmers and GU

MSA101 / MVE187 Computational methods for Bayesian statistics

Exam 2 January 2018, 14:00 - 18:00
Allowed aids: None.
The appendix of this exam contains informaion about some probability distributions.
Total number of points: 30. To pass, at least 12 points are needed

1. (2 points) Assume a random variable with values on the interval $[0,3]$ has a density $\pi(x)$ that is bounded by B, i.e., $\pi(x) \leq B$ for all $x \in [0,3]$.
   (a) Describe in detail the steps of a general method with which one may generate a sample from this distribution.
   (b) Describe exactly how the efficiency of the method depends on the value of B.
2. (3 points) If the positive random variable X has a Lomax distribution with parameters α and β, then a random value for X may be simulated by first simulating $Y \sim \text{Gamma}(\alpha,\beta)$, and then simulating $X \mid Y \sim \text{Exponential}(Y)$. Compute and simplify the probability density function for X.
3. (3 points) Let $\pi(x \mid \theta)$ be a probability density for real x and θ, let $f_1(\theta),\dots,f_k(\theta)$ be k different prior densities, and let $g_1(\theta \mid x),\dots,g_k(\theta \mid x)$ be the corresponding posterior densities. Assuming θ has the prior
$$\pi(\theta) = c_1 f_1(\theta) + c_2 f_2(\theta) + \cdots + c_k f_k(\theta)$$
   for fixed positive constants $c_1,\dots,c_k$ summing to 1, derive an expression for the postererior density $\pi(\theta \mid x)$ in terms of $\pi(x \mid \theta)$, $f_1(\theta),\dots,f_k(\theta)$, $g_1(\theta \mid x),\dots,g_k(\theta \mid x)$, and $c_1,\dots,c_k$.
4. (2 points) Explain the precise difference between a Bayesian network and causal network.
5. (6 points) Assume we want to simulate from a distribution with a density proportional to a function $f(x)$, where x is some vector of real numbers, using the Metropolis Hastings algorithm. Denote the proposal function by $q(y \mid x)$.
   (a) Write down the steps of the algorithm in terms of the functions above.
   (b) Explain what the detailed balance condition is, and what role it plays in the proof of convergence of the Metropolis Hastings algorithm.
   (c) In order to make inference based from a sample derived with Metropolis Hastings, one needs to assess how close it is to being a sample from the target density. One possibility is to use multiple independent chains, each produced with the algorith above. Explain the idea behind this approach, and outline how one may derive information about convergence from such a set of chains.

6. (8 points) Consider the following model:

$$\begin{aligned}
\mu &\sim \text{Normal}\left(\mu_0, \tau_0^{-1}\right)\\
\tau_1 &\sim \text{Gamma}\left(\alpha, \beta\right)\\
x \mid \mu, \tau_1 &\sim \text{Normal}\left(\mu, \tau_1^{-1}\right)
\end{aligned}$$

where $\mu_0, \tau_0, \alpha, \beta$ are assumed known, with $\alpha > 1$.

(a) Describe in detail how to use Gibbs sampling to obtain an approximate sample from the jont posterior for μ and τ_1 given x. Include the formulas to simulate from.

(b) Assume you want to find the τ_1 maximizing the marginal posterior for τ_1 given x. Describe in detail an iterative algorithm computing this. Include the formulas used in each step.

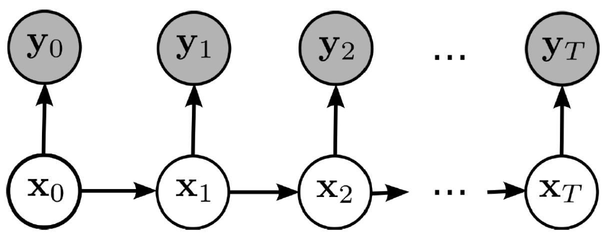

Figure 1: The hidden Markov model used in question 6.

7. (4 points) Consider the model illustrated in Figure 1, with each X_i having possible states 0 and 1, with $\pi(X_{i+1} \mid X_i)$ the same for all i, and with

$$\begin{aligned}
Y_i \mid X_i = 0 &\sim \text{Gamma}(\alpha_0, 1)\\
Y_i \mid X_i = 1 &\sim \text{Gamma}(\alpha_1, 1)
\end{aligned}$$

for fixed and different α_0 and α_1. Assume $Y_0, \dots, Y_T$ are observed, and you want to find the sequence $X_0, \dots, X_T$ maximizing the posterior with these data. Describe an algorithm computing this.

8. (2 points) What is Sampling Importance Resampling? Give a short explanation.

Appendix: Some probability distributions

The Bernoulli distribution

If $x \in \{0, 1\}$ has a Bernoulli(p) distribution, with $0 \le p \le 1$, then the probability mass function is

$$\pi(x) = p^x(1-p)^{1-x}.$$

The Beta distribution

If $x \geq 0$ has a Beta(α, β) distribution with $\alpha > 0$ and $\beta > 0$ then the density is

$$\pi(x \mid \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1}.$$

The Beta-Binomial distribution

If $x \in \{0, 1, 2, \ldots, n\}$ has a Beta-Binomial(n, α, β) distribution, with n a positive integer and parameters $\alpha > 0$ and $\beta > 0$, then the probability mass function is

$$\pi(x \mid n, \alpha, \beta) = \binom{n}{x} \frac{\Gamma(x+\alpha)\Gamma(n-x+\beta)\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)\Gamma(n+\alpha+\beta)}.$$

The Binomial distribution

If $x \in \{0, 1, 2, \ldots, n\}$ has a Binomial(n, p) distribution, with n a positive integer and $0 \leq p \leq 1$, then the probability mass function is

$$\pi(x \mid n, p) = \binom{n}{x} p^x (1-p)^{n-x}.$$

The Cauchy distribution

If $x \geq 0$ has a Cauchy(μ, γ) distribution, with $\gamma > 0$, then the probability density is

$$\pi(x \mid \mu, \gamma) = \frac{1}{\pi\gamma\left(1 + \left(\frac{x-\mu}{\gamma}\right)^2\right)}.$$

The Exponential distribution

If $x \geq 0$ has an Exponential(λ) distribution with $\lambda > 0$ as parameter, then the density is

$$\pi(x \mid \lambda) = \lambda \exp(-\lambda x)$$

and the cumulative distribution function is

$$F(x) = 1 - \exp(-\lambda x).$$

The Gamma distribution

If $x > 0$ has a Gamma(α, β) distribution, with $\alpha > 0$ and $\beta > 0$, then the density is

$$\pi(x \mid \alpha\beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} \exp(-\beta x).$$

The expectation and variance are α/β and α/β^2, respectively, while the mode is $(\alpha - 1)/\beta$ (when $\alpha \geq 1$).

The Geometric distribution

If the non-negative integer x has a Geometric distribution with parameter $p \in [0, 1]$, its probability mass function is given by

$$\pi(x \mid p) = (1 - p)^x p.$$

The Logistic distribution

If x has a Logistic(μ, s) distribution, with $s > 0$, then the density is

$$\pi(x \mid \mu, s) = \frac{\exp\left(-\frac{x-\mu}{s}\right)}{s\left(1 + \exp\left(-\frac{x-\mu}{s}\right)\right)^2}$$

and the cumulative density function is given by

$$F(x) = \frac{1}{1 + \exp\left(-\frac{x-\mu}{s}\right)}.$$

The Normal distribution

If the real x has a Normal distribution with parameters μ and σ^2, its density is given by

$$\pi(x \mid \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{\sigma^2}(x - \mu)^2\right).$$

The Uniform distribution

If $x \in [a, b]$ has a Uniform(a, b) distribution with $b > a$, then the density is given by

$$\pi(x \mid a, b) = \frac{1}{b - a}.$$