Assignment 05

CMPE 58N, Monte Carlo Methods

Department of Computer Engineering, Boğaziçi University, Istanbul, Turkey

Instructor: A. T. Cemgil

Due: 12 June 2009, Fri, 15:00.

1. (10) (**Cross Entropy Method**) This assignment is based on de Boer, Kroese, Mannor and Rubinstein tutorial, page 13. Implement the cross entropy method for rare event simulation and create a table similar to table 1 for a network with the same topology but $u = (0.25, 0.4, 0.1, 0.5, 0.3)$.

2. (10) (**Permanent of a binary matrix**) This assignment is based on Liu 4.2, page 90.

   We define a so-called restriction matrix A where $A_{i,j} \in \{0,1\}$. You can think of A as an adjacency matrix of a directed (not necessarily acyclic) graph.

   Now, imagine that on each node i, there is a particle labeled i. The A matrix denotes the set of edges that particle i can travel. A particle moves synchronously to a new node j where $A_{i,j} = 1$ (or stays where there are if $A_{i,i} = 1$). We call the final configuration a valid configuration if there is only a single particle on each node. The permanent of a matrix A the total number of valid configurations.

   Clearly, each configuration is a permutation of particles $1 \ldots N$. We denote a permutation by σ and the set of all permutations by Π. For example, if $(1,2,3) \to (1,3,2)$, then $\sigma(1) = 1$, $\sigma(2) = 3$ and $\sigma(3) = 2$. A configuration is valid only if

$$\prod_{i=1}^{N} A_{i,\sigma(i)} = 1$$

   The permanent is

$$\mathrm{perm}(A) = \sum_{\sigma \in \Pi} \prod_{i=1}^{N} A_{i,\sigma(i)}$$

   (a) Write a program to exactly compute the permanent of a matrix. This is practical only for N about 12. Test your program with the test matrices to be posted on the course web page.

   (b) Implement the importance sampling strategy described by Liu and compare with exact results.

   (c) (optional) Implement the cross entropy method for the same problem

3. (10) (**Sequential Monte Carlo**) Consider a nonlinear dynamical system (Noisy Sinusoidal with frequency modulation) covered in the class. The model is

$$\begin{aligned} \Delta_k &\sim \mathcal{N}(\Delta_k; \Delta_{k-1}, Q) \\ \phi_k &= \phi_{k-1} + \Delta_k \\ y_k &\sim \mathcal{N}(y_k; \sin(\phi_k), R) \end{aligned}$$

(a) Write a program to generate realisations from this model. Find reasonable values for Q, R and Δ_0 via experimentation so that you generate sequences qualitatively similar as the one given on the slides.

(b) Implement a sequential Monte Carlo algorithm to estimate the mean and variance of $\Delta_k | y_{1:k}$.

(c) For a given observation sequence $y_{1:K}$, compute the marginal likelihood, i.e., $p(y_{1:K}|Q, R, \Delta_0)$.

4. (10) (**Model Selection 1**) Consider the following *coal mining* dataset. The table contains the number of deadly accidents in coal mines occurring between 1851 and 1962 in the UK.

   [4 5 4 1 0 4 3 4 0 6 3 3 4 0 2 6 3 3 5 4 5 3 1 4 4 1 5 5 3 4 2 5 2 2 3 4 2 1 3 2 2 1 1 1 1 3 0 0 1 0 1 1 0 0 3 1 0 3 2 2 0 1 1 1 0 1 0 1 0 0 0 2 1 0 0 0 1 1 0 2 3 3 1 1 2 1 1 1 1 2 4 2 0 0 0 1 4 0 0 0 1 0 0 0 0 0 1 0 0 1 0 1]

   (a) Write down an appropriate model for this problem (consider Poisson intensity multiple change point model covered in class)

   (b) Using an MCMC method or exact computation, find the positions when the intensity has changed. Can you detect the points when new health and safety regulations are introduced in the late 18th century?

5. (10) (**Model Selection 2**)

Consider the following generalised linear model.

$$\begin{aligned} r_{i,j} &\sim \mathcal{B}(r_{i,j}; \pi) \\ s_{i,j}|r_{i,j} &\sim \mathcal{N}(s_{i,j}; 0, \Sigma(r_{i,j})) \\ \mathbf{x}_j|s_{1:W,j} &\sim \mathcal{N}(\mathbf{x}; Cs_{1:W,j}, V) \\ C &\equiv [\ C_1 \ \ \dots \ \ C_i \ \ \dots \ \ C_W\ ] \end{aligned}$$

Here, $r_{i,j} \in \{0,1\}$ for $i = 1 \dots W$ and $j = 1 \dots N$, $\Sigma(0) = 0$, $\Sigma(1) = 1000$, $\pi = 0.1$ and $V = 0.01 I_W$. Each data vector $\mathbf{x}_j$ is $L \times 1$. We let $X = [\mathbf{x}_1, \dots, \mathbf{x}_N]$, $S = \{s_{i,j}\}$, $R = \{r_{i,j}\}$.

(a) This model is closely related to a matrix factorisation approach where $X \approx CS$. Interpret X, C and S and explain the link.

(b) Write a program to sample from this generative model.

(c) Assume that the matrix C is known. Develop a MCMC algorithm to sample from $p(S, R|X, C)$. Test your algorithm with the following data

$$\begin{aligned} C &= \begin{pmatrix} 1 & 0 & 1 & 2 & 1 \\ 0 & 1 & 1 & 1 & 2 \end{pmatrix} \\ X &= \begin{pmatrix} 5 & 1 & 4 & 2 \\ 0 & 1 & 1 & 2 \end{pmatrix} \end{aligned}$$

and compare your solutions (e.g., $\langle S \rangle_{p(S|X,C)}$ or a configuration that is sampled by your chain) with the least squares solution $S = C^{\dagger} X$. In matlab, you can compute the pseudo-inverse via `pinv` function. Interpret the results geometrically on a $2-D$ plot.

(d) Develop a MCMC algorithm to sample from $p(S, R, C|X)$.

(e) (Tempering) Define a sequence of problems with $\Sigma_\tau(0) = \tau$ such that $\tau \to 0$. Implement simulated tempering by slowly changing τ. Does this improve your sampler?