Assigned: 08 April 2020

Project #8 – Markov Chain Monte Carlo

EE 511: Spring 2020

Due: Wednesday, 29 April 2020 at 14:00. No late submissions.

1. Use a total sample budget of $n = 1000$ samples to obtain Monte Carlo estimates and variances for the following integrals in two dimensions (x_1 and x_2). Then implement stratification and importance sampling in your Monte Carlo estimation simulation using the same sample budget. Compare the three different Monte Carlo integral estimates, the quality of the estimates, and their sample variances.

   a. $\exp\left(\sum_{i=1}^{2} 5 \cdot |x_i - 5|\right)$ for $0 \le x_i \le 1$.

   b. $\cos\left(\pi + \sum_{i=1}^{2} 5x_i\right)$ for $-1 \le x_i \le 1$.

2. Let X_i, $i = 1, 2, 3$ be independent exponentials with mean 1. Use Gibbs sampling to estimate the following.

   a. $\mathrm{E}[X_1 + 2X_2 + 3X_3 | X_1 + 2X_2 + 3X_3 > 15]$.

   b. $\mathrm{E}[X_1 + 2X_2 + 3X_3 | X_1 + 2X_2 + 3X_3 < 1]$.

3. The Schwefel function is a standard optimization benchmark because it has many local minimas and a single global minimum. The Schwefel function is given by

$$f(x) = 418.9829d - \sum_{i=1}^{d} x_i \sin\left(\sqrt{|x_i|}\right)$$

   over the hypercube $-500 \le x_i \le 500$ for $i = 1, \ldots, d$ where d is the dimension. For this problem consider the surface in 2-dimensions ($d = 2$). Draw a contour plot of the 2-dimensional Schwefel surface. Find the global minimum of the surface using simulated annealing. Begin each simulation at the origin $(0,0)$. Compare the behavior of your simulation when using exponential, polynomial, and logarithmic cooling schedules. Run your algorithm for with different iteration counts (e.g. $N = \{50, 200, 1000, 10000\}$) and generate a histogram showing the computed global minimum for each case. For the simulation that achieves the best estimate of the global minimum overlay the 2-D sample path on a contour plot of the surface and comment on the behavior of the search.

4. You and a friend decide to take a summer road trip through every state capitol in the contiguous United States (the 48 states excluding Hawaii and Alaska) now that you have finished another tough year at USC. Refer to the addendum data files that contain names and (x, y)-coordinates (these are based on cylindrical projection onto the plane) of all 50 US state capitals. Use a simulated annealing simulation to determine a minimal path between these 48 state capitals (removing Juneau, Alaska and Honolulu, Hawaii). Use the Euclidean distance to estimate the distance between two cities given (x, y)-coordinates:

$$d(c_1, c_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$$

   Begin your tour in Sacramento, California. Initialize your simulation with an arbitrary trip through the remaining state capitols. Choose new candidate circuits by randomly choosing two cities and then reversing the path between them. Estimate the distance of the shortest path that visits each of the 48 state capital cities. Plot the best path on the x-y axis and comment on the graphical path. Generate a plot of the total tour distance as a function of the simulation time (step number). Comment on the number of iterations required to estimate this path and on the convergence rate of the estimate toward your minimum.