ST407 Monte Carlo Methods 2019/20

Assignment 2

Issued: Friday, November 22nd, 2019

Deadline: Midday on Thursday, December 5th, 2019

Note: Completed work should be submitted to the student support office no later than the above deadline, by posting it in the appropriate year box. You *must* attach and sign an "Assignment Cover Sheet", which can be found beside the postboxes outside the student support office. The cover sheet includes a section where you declare whether the work is your own. In Warwick assessment is taken very seriously and the work you submit must be your own, it must be written in your own words and any computer code which is included must be your own original work. This does not preclude you from *discussing* the assignment with others. If it is unclear what the difference is, please consult prior to completing any work a copy of the "Assignment Cover Sheet" to understand fully what you are declaring on submission.

1. Consider the following multi-modal univariate density, where s_i is the i^{th} digit of you your seven digit student ID number (for instance, if your ID number is **u1712345**, then $s_4 = 2$ and $s_7 = 5$):

$$f(x) \propto s_7 \cdot \exp\left\{-\sin\left(\frac{s_1 \cdot x^2}{15 - s_1}\right) - \frac{(x - 3 - s_2 \cdot \pi)^2}{2 \cdot (5 + s_3)^2}\right\} + 2 \cdot (1 + s_7) \cdot \exp\left\{-\frac{x^2}{32}\right\}$$
$$+ (10 - s_7) \cdot \exp\left\{-\cos\left(\frac{s_4 \cdot x^2}{15 + s_4}\right) - \frac{(x + 3 + s_5 \cdot \pi)^2}{2 \cdot (5 + s_6)^2}\right\}.$$

(a) Write an R function to evaluate this density (up to its normalising constant), and produce a plot to illustrate this density.

(b) For each of the following parts: i) implement the methodology indicated; ii) provide appropriately documented R code snippets for the methodology; iii) discuss the selection of any tuning parameters; iv) provide appropriate diagnostics for the methodology; v) contrast your output directly with the objective, giving a qualitative assessment of the implementation.

i. Implement a simulated annealing algorithm to provide an estimate of the global mode of f.

ii. Implement a random walk Metropolis algorithm with f being its invariant distribution.

iii. Implement a slice sampler with f being its invariant distribution.

iv. Implement (an)other Monte Carlo algorithm(s) of your choosing with f as its (/their) invariant distribution. Provide an indication of why your choice may be interesting to investigate (it may of course transpire in practice it is not). If the algorithm wasn't explicitly covered in this course, then provide a short description of the methodology and an appropriate reference.

(c) A helpful Ph.D. student suggests it might be interesting to consider modifying your methodology in b) by replacing the Metropolis-Hasting acceptance ratio ($\alpha_{\text{M-H}}$) in your algorithm(s) with his variant, $\alpha_{\text{G-M}}$. In particular, he suggests instead of accepting a proposal, $\mathbf{X}$ (drawn from an appropriate instrumental distribution), with probability $\alpha_{\text{M-H}}(\mathbf{X}|\mathbf{X}^{(t-1)}) := \min\left\{1, \frac{f(\mathbf{X})\cdot q(\mathbf{X}^{(t-1)}|\mathbf{X})}{f(\mathbf{X}^{(t-1)})\cdot q(\mathbf{X}|\mathbf{X}^{(t-1)})}\right\}$, to instead accept the proposal with probability $\alpha_{\text{G-M}}(\mathbf{X}|\mathbf{X}^{(t-1)}) := \frac{f(\mathbf{X})\cdot q(\mathbf{X}^{(t-1)}|\mathbf{X})}{f(\mathbf{X})\cdot q(\mathbf{X}^{(t-1)}|\mathbf{X}) + f(\mathbf{X}^{(t-1)})\cdot q(\mathbf{X}|\mathbf{X}^{(t-1)})}$.

i. Is this modification valid? (i.e. does this modification to the algorithms in b) result in algorithms producing the correct output?).

ii. Re-consider the methodologies and implementation developed in b) in light of this suggestion. In each case, does the modification help?

(d) If you were to choose a single methodology to sample from your target density, f, which would you choose, and why?

Some guidance:

- If you provide a sensible implementation for each part, then you should achieve a high mark even if your approach is not 'optimal'. Indeed, it is possible that a highly tuned and perfectly executed implementation of a given part *won't* work well for your particular target density – that's the point! - this assessment is an investigation into what is *most* effective, and the limitations of the Monte Carlo algorithms studied during the course.
- You will gain (/ lose) marks on your ability to appropriately convey your answer to each part clearly and concisely. Providing reams of exhaustive analysis indicates a *lack* of understanding!
- I will be judging the quality of your investigations and implementations by using your reported answer in (d), and the results of your answer against my own un-tuned reference implementation. As such, to score highly in this part provide an appropriate quantitative analysis / justification.
- Remember that it is worth just 10% of the overall module mark, and so use your time efficiently.