Candidates are admitted to the examination room ten minutes before the start of the examination. On admission to the examination room, you are permitted to acquaint yourself with the instructions below and to read the question paper.

Do not write anything until the invigilator informs you that you may start the examination. You will be given five minutes at the end of the examination to complete the front of any answer books used.

April 2012

Answer Book
Data Sheet
Any bilingual English language dictionary permitted
Only Casio-fx83 calculators are permitted

Final Examination for MSc

Course in Data Assimilation and Inverse Modelling in Geosciences

MTMD03

Unit 2/MT/MD03 : Monte-Carlo methods and Particle filters

Two hours

Answer **ANY TWO** questions

The marks for the individual components of each question are given in [ ] brackets. The total mark for the paper is 100.

1. (a) Describe how one can use the transformation method, also called probability integral transform method, to derive samples from the Rayleigh distribution given by

$$f(z) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}$$

starting from samples from the uniform distribution U(0,1]. Note the bound of the uniform density.

[10 marks]

(b) Explain how the Metropolis-Hastings algorithm works.

[4 marks]

Why does it converge to the correct posterior?

[4 marks]

What is the burn-in period, and how do you know the burn-in period is over?

[6 marks]

Question 1 continued.

© Assume the target pdf is given by figure 1.

Explain why Metropolis-Hastings with one single Markov chain will not work in this case.

[12 marks]

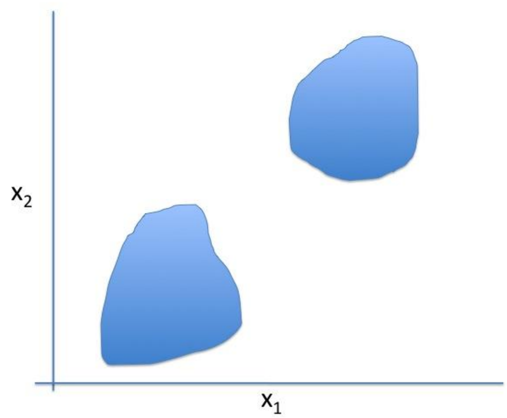

Figure 1

(d) One way to deal with the problem outlined above is to use the Hybrid Monte-Carlo method. Give the steps taken in that algorithm, and explain why it should work in this case.

[12 marks]

Give one other method to solve this problem.

[2 marks]

2.

(a) Consider a 2-dimensional system (x_0, x_1) taking on values *0*, and *1* with probability distribution (0<q<1):

$$\pi = \begin{pmatrix} \pi_{00} & \pi_{01} \\ \pi_{10} & \pi_{11} \end{pmatrix} = \begin{pmatrix} q/2 & (1-q)/2 \\ (1-q)/2 & q/2 \end{pmatrix}$$

Show that this probability distribution is properly normalised.

[2 marks]

Calculate the correlation ρ between x_0 and x_1.

[4 marks]

Calculate the conditional probabilities $p(x_0 \backslash x_1)$, and $p(x_1 \backslash x_0)$ for all values of x_0 and x_1.

[4 marks]

(b) We want to build a Markov Chain that samples from this probability distribution.

Explain how the Gibbs sampler works.

[4 marks]

Sequential samples from the Gibbs sampler will be correlated. Explain how this correlation can be detected, and how one can obtain uncorrelated samples.

[6 marks]

(c) Define $p_n = \Pr(x_0^n = 0)$ for *n=0,1,...* In the following use

$\Pr(x_0^n = 0 \mid x_0^{n-1} = 0) = \Pr(x_0^n = 1 \mid x_0^{n-1} = 1) = q^2 + (1-q)^2$

and

$\Pr(x_0^n = 1 \mid x_0^{n-1} = 0) = \Pr(x_0^n = 0 \mid x_0^{n-1} = 1) = 2q(1-q)$

Show that $p_n = \rho^2 p_{n-1} + 2q(1-q)$ with ρ determined in (a).

[6 marks]

Using this show that $p_n = \rho^{2(n+1)} p_0 + 2q(1-q)\dfrac{1-\rho^{2(n+1)}}{1-\rho^2}$

(Hint: Use $\sum_{j=0}^{n} x = (1-x^{n+1})/(1-x)$

[10 marks]

Finally show that $\lim_{j\to\infty} p_j = \frac{2q(1-q)}{1-\rho^2} = \frac{1}{2}$ so the Gibbs sampler will come up with the correct probability.

[4 marks]

(d) Let us now discuss the convergence rate of the Gibbs sampler.

Show that the marginal chain for x_0 has a transition matrix with eigenvalues 1 and ρ^2, so the convergence rate is $|\rho^2|$.

[6 marks]

Show that if q is close to 0 or 1 the convergence rate will be slow.

[4 marks]

3\. For all items below use that observations are present every time step, and that model and observation errors are Gaussian, and that the measurement operator is linear.

(a) Explain how a proposal transition density can be used in Particle filtering to increase the efficiency of the particles.

[5 marks]

Give the expression for the weights when a proposal density is used, and identify likelihood and proposal density parts.

[4 marks]

Give an example of a proposal transition density, and explain how it works.

[6 marks]

(b) Explain what the so-called 'optimal proposal density' is.

[4 marks]

Calculate the weights using this proposal density using the information given at the start of question 3.

[6 marks]

Explain why it is inefficient when the number of independent observations is large.

[5 marks]

(c) Discuss one method to ensure that all particles have almost equal weight (without resampling the whole ensemble), so that the filter is efficient in high dimensional systems with a large number of independent observations. No need to include equations, but illustrations to explain the ideas should be included.

[20 marks]

End of question paper