CPSC-540 Machine Learning 2005

Homework # 4
Due Monday March 14 in class.

NAME:____

Signature:_____________________________________

STD. NUM: ____________________________________

General guidelines for homeworks:

You are encouraged to discuss the problems with others in the class, but all write-ups are to be done on your own.
Homework grades will be based not only on getting the "correct answer," but also on good writing style and clear presentation of your solution. It is your responsibility to make sure that the graders can easily follow your line of reasoning.
Try every problem. Even if you can't solve the problem, you will receive partial credit for explaining why you got stuck on a promising line of attack. More importantly, you will get valuable feedback that will help you learn the material.
Please acknowledge the people with whom you discussed the problems and what sources you used to help you solve the problem (e.g. books from the library). This won't affect your grade but is important as academic honesty.
When dealing with Matlab exercises, please attach a printout with all your code and show your results clearly.

1. **Metropolis for Logistic Regression**: Repeat question 1 of Homework 3, but this time use the Metropolis algorithm with a random walk proposal distribution. Hand in the code and plots.
2. **Detailed balance**: Show that the MH algorithm satisfies detailed balance.

3. **Particle Filtering**: This question builds on the importance sampling question of the previous homework. However, now we have a dynamical model. That is, the parameters are changing over time. It is therefore necessary to update the posterior distribution as data become available. Examples include tracking an aircraft using radar measurements, estimating a digital communications signal using noisy measurements or estimating the volatility of financial instruments using stock market data.

   The dynamic model consists of three equations: the initial probability, the transition model and the observation model. The unobserved signal (hidden states or unknown parameters) $\{x_t; t \in \mathbb{N}\}$, $x_t \in \mathcal{X}$, is modelled as a *Markov process* of initial distribution $p(x_0)$ and transition equation $p(x_t | x_{t-1})$. The observations $\{y_t; t \in \mathbb{N}\}$, $y_t \in \mathcal{Y}$, are assumed to be conditionally independent given the process $\{x_t; t \in \mathbb{N}\}$ and of marginal distribution $p(y_t | x_t)$. To sum up, the model is described by

$$\begin{aligned} &p(x_0) \\ &p(x_t | x_{t-1}) \quad \text{for } t \geq 1 \\ &p(y_t | x_t) \quad \text{for } t \geq 1 \end{aligned}$$

   (i) Draw the probabilistic graphical model for this problem.

   So we know $p(x_0)$, $p(x_t|x_{t-1})$ and $p(y_t|x_t)$. Our aim is to estimate recursively in time the *posterior distributionp* $(x_{0:t} | y_{1:t})$ and its associated features including the marginal distribution $p(x_t | y_{1:t})$, known as the *filtering distribution*.

   At any time t, the posterior distribution is given by *Bayes' theorem*

$$p(x_{0:t} | y_{1:t}) = \frac{p(y_{1:t} | x_{0:t})\, p(x_{0:t})}{\int p(y_{1:t} | x_{0:t})\, p(x_{0:t})\, dx_{0:t}}$$

   (ii) Show that the following recursive update is true:

$$p(x_{0:t+1} | y_{1:t+1}) = p(x_{0:t} | y_{1:t}) \frac{p(y_{t+1} | x_{t+1})\, p(x_{t+1} | x_t)}{p(y_{t+1} | y_{1:t})}$$

   State your conditional independency assumptions clearly.

   The marginal distribution $p(x_t | y_{1:t})$ also satisfies the following recursion

$$\textit{Prediction: } p(x_t | y_{1:t-1}) = \int p(x_t | x_{t-1})\, p(x_{t-1} | y_{1:t-1})\, dx_{t-1}$$

$$\textit{Updating: } p(x_t | y_{1:t}) = \frac{p(y_t | x_t)\, p(x_t | y_{1:t-1})}{\int p(y_t | x_t)\, p(x_t | y_{1:t-1})\, dx_t}$$

   These expressions and recursions are deceptively simple as one cannot typically compute the integrals. However, if the distributions are Gaussian or discrete, we can solve the integrals. In the Gaussian case, the answer is known as the Kalman filter. In the discrete case, the answer gives rise to the HMM filter. In general, we need to do sampling to solve this problem.

Let's do importance sampling. That is, we sample from a proposal distribution that we choose and then weight the samples appropriately. Let the proposal distribution be

$$q\left(x_{0:t}\middle|\, y_{1:t}\right) = p\left(x_{0:t-1}\middle|\, y_{1:t-1}\right)\; q\left(x_t\middle|\, x_{0:t-1}, y_{1:t}\right)$$

That is, the proposal only affects the state at time t. In other words, we don't modify samples in the past.

(iii) Show that the importance weights are given by

$$w_t \propto \frac{p(y_t|x_t)p(x_t|x_{t-1})}{q(x_t|x_{1:t-1}, y_{1:t})}$$

Hint:

$$w_t = \frac{p(x_{0:t}|y_{1:t})}{q(x_{0:t}|y_{1:t})}$$

(iv) What's the expression for w_t when we go for the choice $q(x_t|x_{1:t-1}, y_{1:t}) = p(x_t|x_{t-1})$?

The idea is to draw N samples (known as particles) at each time step t. We then weight these samples and choose the fittest samples by resampling. The basic algorithm is as follows:

Simple Particle Filter

(a) *Initialization*, $t = 0$.

- For $i = 1, ..., N$, sample $\mathbf{x_0^{(i)}} \sim \mathbf{p}\left(\mathbf{x_0}\right)$ and set $t = 1$.

2. *Importance sampling step*

- For $i = 1, ..., N$, sample $\widetilde{\mathbf{x}}_t^{(i)} \sim p\left(\mathbf{x_t}\middle|\, \mathbf{x_{t-1}^{(i)}}\right)$.
- For $i = 1, ..., N$, evaluate the importance weights $\widetilde{w}_t^{(i)} \propto p\left(\mathbf{y_t}\middle|\, \widetilde{\mathbf{x}}_t^{(i)}\right)$
- Normalise the importance weights.

3. *Selection step*

- Resample with replacement N particles $\left(\mathbf{x_t^{(i)}}; \mathbf{i = 1, \ldots, N}\right)$ from the set $\left(\widetilde{\mathbf{x}}_t^{(i)}; i = 1, \ldots, N\right)$ according to the importance weights.
- Set $t \leftarrow t + 1$ and go to step 2.

The set of N particles at each time step, provides a description of the filtering distribution $p(x_y|y_{1:t})$.

For demonstration purposes, we apply the bootstrap algorithm to the following nonlinear, non-Gaussian model

$$
\begin{aligned}
x_t &= \frac{1}{2}x_{t-1} + 25\frac{x_{t-1}}{1+x_{t-1}^2} + 8\cos(1.2t) + v_t \\
y_t &= \frac{x_t^2}{20} + w_t
\end{aligned}
$$

where $x_1 \sim \mathcal{N}\left(0, \sigma_1^2\right)$, v_t and w_t are mutually independent white Gaussian noises, $v_t \sim \mathcal{N}\left(0, \sigma_v^2\right)$ and $w_t \sim \mathcal{N}\left(0, \sigma_w^2\right)$ with $\sigma_1^2 = 10$, $\sigma_v^2 = 10$ and $\sigma_w^2 = 1$.

(v) The implementation is is the file `pfhomework.m`. Run this file with the data `trackingData.mat` and save the two plots. Finally, replace the model by the following model

$$
\begin{aligned}
x_t &= \frac{1}{2}x_{t-1} + 25\frac{x_{t-1}}{1+x_{t-1}^2} + 5\sin(1.2t) + v_t \\
y_t &= \frac{x_t^2}{11} + w_t
\end{aligned}
$$

where $\sigma_1^2 = 2$, $\sigma_v^2 = 2$ and $\sigma_w^2 = 0.05$. Run the algorithm using the data file `trackingData2.mat`. Generate the two plots. In total, you should hand in four plots and explain the results.