# CPSC 535D: Assignement 2

February 25, 2007

## 1. Stochastic Volatility Model

Consider the following stochastic volatility model

$$\begin{aligned} X_n &= \alpha X_{n-1} + \sigma V_n, \\ Y_n &= \beta \exp(X_n/2) W_n \end{aligned}$$

where $X_1 \sim \mathcal{N}\left(0, \frac{\sigma^2}{1-\alpha^2}\right)$, $V_n \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1)$ and $W_n \overset{\text{i.i.d}}{\sim} \mathcal{N}(0,1)$.

• Given $\theta = (\alpha, \beta, \sigma)$, design a clever SMC method based on the approximation of the optimal importance distribution $p(x_n | x_{n-1}, y_n)$ and compare it to the standard bootstrap filter where one uses the importance distributions $p(x_n | x_{n-1})$.

• Assume now that $\theta$ is unknown, propose and implement an EM algorithm based on SMC to estimate the maximum likelihood estimate

$$\widehat{\theta}_{ML} = \arg\max_{\theta} p(y_{1:T} | \theta)$$

where the observations $y_{1:T}$ ($T = 200$) can be downloaded from the webpage of the course.

• What are the limitations of this algorithm? what happens if $T$ increases?

## 2. Mixture of Poisson Distributions

The data in Table 1 were extracted from The London Times during the years 1910-1912. The two columns labeled "Deaths $i$" refer to the number of deaths to women 80 years and older reported by day. The columns labeled "Frequency $n_i$" referer to the number of days with $i$ deaths.

| Deaths $i$ | Frequency $i$ | Deaths $i$ | Frequency $i$ |
|---|---|---|---|
| 0 | 162 | 5 | 61 |
| 1 | 267 | 6 | 27 |
| 2 | 271 | 7 | 8 |
| 3 | 185 | 8 | 3 |
| 4 | 111 | 9 | 1 |

A Poisson distribution gives a poor fit to these data, possibly because of different patterns of deaths in winter and summer. A mixture of two Poisson distributions provides a much better fit. Under such a model, the likelihood of the $N$ observations is given by

$$\begin{aligned} p(y_1, \ldots, y_N | \theta) &= \prod_{i=1}^{N} \left[ \alpha \exp(-\lambda_1) \frac{\lambda_1^{y_i}}{y_i!} + (1-\alpha) \exp(-\lambda_2) \frac{\lambda_2^{y_i}}{y_i!} \right] \\ &= \prod_{i=1}^{\max_j y_j} \left[ \alpha \exp(-\lambda_1) \frac{\lambda_1^{y_i}}{y_i!} + (1-\alpha) \exp(-\lambda_2) \frac{\lambda_2^{y_i}}{y_i!} \right]^{n_i} \end{aligned}$$

where $\theta = (\alpha, \lambda_1, \lambda_2)$. $\alpha \in [0,1]$ is the mixture parameter, $\lambda_1$ and $\lambda_2$ are the means of the two components, $N$ is the total numbers of days, $n_i = \sum_{j=1}^{N} 1(y_j = i)$ is the total number of days with $i$ deaths, and $\max_j y_j = 9$ for this example. We use the following prior $\alpha \sim \mathcal{U}[0,1]$ (uniform distribution on $[0,1]$) and $\lambda_i \sim \mathcal{G}(0.1, 0.1)$ (Gamma distribution with parameters 0.1 and 0.1) for $i = 1, 2$.

- Propose and Implement a Metropolis-Hastings algorithm to sample from $p(\theta | y_1, ..., y_N)$.
- Propose a Gibbs sampler to sample from $p(\theta | y_1, ..., y_N)$ [Hint: It requires introducing some missing data].

Note that $p(\lambda_1 | y_1, ..., y_N) = p(\lambda_2 | y_1, ..., y_N)$, this suggests a very simple method to assess the convergence of your algorithms.

- Present meaningful point estimates of $\lambda_1$ and $\lambda_2$.