1 Methods

1.1 Model

Our model assumes that there are T time periods $t = 1, 2, \ldots, T$. A company has a stock return y_t at t time period which is observable. At time period t, there's also a hidden variable $h_t \in \{1, 2\}$ representing whether that company is in a good state or a bad state. The model assumes that y_t is generated from h_t by

$$y_t | h_t \sim N(\mu_{h_t}, \sigma_{h_t}) \tag{1}$$

That means, the return is drawn from two normal distributions, a bad distribution for bad state and a good distribution for good state. Presumably, $\mu_2 > \mu_1$, which means that good state tends to get higher expected return.

The h_t satisfies Markov property such that

$$Pr(h_t | h_1^{t-1}) = Pr(h_t | h_{t-1})$$

That means, h_t is conditional independent with $h_1, h_2, \ldots, h_{t-2}$ given h_{t-1}. Therefore the transition probability can be simply characterized as

$Pr(h_t \| h_{t-1})$	$h_t = 1$	$h_t = 2$
$h_{t-1} = 1$	q_1	$1 - q_1$
$h_{t-1} = 2$	q_2	$1 - q_2$

In summary, our model assumes that a company's stock return is a Hidden Markov Model (HMM) of T time periods, where observable data $\boldsymbol{y}$ (stock return) is drawn from normal distributios with mean and standard deviation $\boldsymbol{\mu}, \boldsymbol{\sigma}$ depending on the hidden state $\boldsymbol{h}$. The hidden state transition probability are $\boldsymbol{q}$. Overall, our model has data $\boldsymbol{y}$, hidden variables $\boldsymbol{h}$ and parameters $\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{q}$. In later sections, we also use $\boldsymbol{\theta}$ to denote all parameters $(\boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{q})$.

1.2 Priors and Full Conditionals for MCMC

To infer the parameters and predict future stock return, we run MCMC with the following priors and full conditionals.

For $\boldsymbol{\mu}, \boldsymbol{\sigma} | y, h, q$, we use the Normal-Gamma as the prior (for $\mu, \phi = 1/\sigma^2$) like what we did in ordinary Gaussian (normal) models:

$$1/\sigma^2 \sim G(\nu_0/2, SS_0/2) \tag{2}$$

$$\mu | \phi \sim N(\mu_0, 1/(\kappa_0 \phi)) \tag{3}$$

The full conditionals of $\boldsymbol{\mu}, \boldsymbol{\sigma}$ are also Normal-Gamma. The only difference is that we need to use $\boldsymbol{h}$ to split $\boldsymbol{y}$ into two sets $Y_{h=1}, Y_{h=2}$. Then we use $Y_{h=k}$ to update μ_k, σ_k:

$$\mu_k | \phi_k, Y_{h=k} \sim N\left(\mu_{n,k}, \frac{1}{\kappa_{n,k}\phi_k}\right) \tag{4}$$

$$\phi_k | Y_{h=k} \sim G(\nu_{n,k}/2, SS_{n,k}/2) \tag{5}$$

where

$$\kappa_{n,k} = \kappa_0 + n_k = \kappa_0 + |Y_{h=k}| \tag{6}$$

$$\mu_{n,k} = \frac{\phi_k n_k \overline{Y_{h=k}} + \phi_k \kappa_0 \mu_0}{\phi_k \kappa_{n,k}} \tag{7}$$

$$\nu_n = \nu_0 + n_k \tag{8}$$

$$SS_n = SS_0 + SS_{h=k} + \frac{n_k \kappa_0}{\kappa_{n,k}} (\overline{Y_{h=k}} - \mu_0)^2 \tag{9}$$

All those priors and full conditionals are similar to what's covered in lecture 6.

For $\boldsymbol{q}$, we use Beta as prior

$$q \sim Beta(a_0, b_0) \tag{10}$$

The full conditionals are also Beta distributions. Suppose n_{ij}^t is the number of transitions in the form $h_{t-1} = i, h_t = j$, we have

$$q_1 | \boldsymbol{h} \sim Beta(a_0 + n_{11}^t, b_0 + n_{12}^t) \tag{11}$$

$$q_2 | \boldsymbol{h} \sim Beta(a_0 + n_{21}^t, b_0 + n_{22}^t) \tag{12}$$

The full conditionals of $\boldsymbol{h}$ can be derived in two ways:

1. **Direct Gibbs (DG)**: In the first approach, we calculate full conditional for each h_t (thus assuming that all other $\boldsymbol{h}_{-t}$ are fixed):

$$Pr(h_t | \boldsymbol{h_{-t}, q, y}) \propto Pr(h_t | h_{t-1}, \boldsymbol{q}) Pr(h_{t+1} | h_t, \boldsymbol{q}) Pr(y_t | h_t, \mu_{h_t}, \sigma_{h_t}) \tag{13}$$

   where $Pr(h_i | h_{i-1}, \boldsymbol{q}) = q_{h_{i-1} h_i}$ if we set $q_{11} = q_1, q_{12} = 1 - q_1, q_{21} = q_2$ and $q_{22} = 1 - q_2$.

2. **Forward Backward (FB) Recursion**: In the second approach, we calculate full conditionals of the whole $\boldsymbol{h}$ using Forward Backward Recursion algorithms (actually, we will only use forward step here) that are well known in previous HMM research. The full conditionals are (recall that $\boldsymbol{\theta = (q, \mu, \sigma)}$ represents all parameters)

$$\begin{aligned} Pr(\boldsymbol{h} | \boldsymbol{y, \theta}) &= Pr(h_T | y_1^T, \boldsymbol{\theta}) \prod_{t=n-1}^{1} Pr(h_t | h_{t+1}^T, \boldsymbol{y, \theta}) \\ &= Pr(h_T | y_1^T, \boldsymbol{\theta}) \prod_{t=n-1}^{1} Pr(h_t | h_{t+1}, y_1^{t+1}, \boldsymbol{\theta}) \end{aligned} \tag{14}$$

   where $Pr(h_t | h_{t+1}, y_1^{t+1}, \boldsymbol{\theta}) \propto Pr(h_t, h_{t+1} | y_1^{t+1}, \boldsymbol{\theta})$. Note that $Pr(h_T | y_t^T, \boldsymbol{\theta}), Pr(h_t, h_{t+1} | y_1^{t+1}, \boldsymbol{\theta})$ can be calculated efficiently by Forward Backward Recursion:

$$Pr(h_t, h_{t+1} | y_1^{t+1}, \boldsymbol{\theta}) \propto Pr(h_t | y_1^t, \boldsymbol{\theta}) Pr(h_{t+1} | h_t, \boldsymbol{q}) Pr(y_{t+1} | h_{t+1}, \boldsymbol{\theta}) \tag{15}$$

$$Pr(h_t | y_1^t, \boldsymbol{\theta}) = \sum_{h_{t-1}} Pr(h_t, h_{t-1} | y_1^t, \boldsymbol{\theta}) \tag{16}$$