1 Notation

1. $i \in \{1, 2, 3\}$ denotes the state of the Markov process
2. Y_t is an observation of inflation at time t
3. $X_t \in \{1, 2, 3\}$ is the state of the Markov process at time t
4. Θ is a vector of all parameters; state means, state variances, transistion probabilties, and a $T \times 3$ matrix of state probabilties for each period (μ, σ^2, A, π)
5. $\Phi(\cdot)$ is the CDF of the standard Normal
6. F_t is the belief of the agent at time t for the distribution of the parmaters

We have the following by definition of the Markov process

$$Y|i \sim \mathcal{N}(\mu_i, \sigma_i^2)$$

$$\begin{aligned}
Pr_t(Y_t \leq y) &= \int_{\theta \in \Theta} Pr(Y_t \leq y|\theta) \, \mathrm{d}\, F_t(\theta) \\
Pr_t(Y_t \leq y) &= \int_{\theta \in \Theta} \Phi\left(\frac{y - \mu}{\sigma}\right) \mathrm{d}\, F_t(\theta) \\
Pr_t(Y_t \leq y) &= \mathbb{E}_{\theta \sim F_t}\left[\Phi\left(\frac{y - \mu}{\sigma}\right)\right]
\end{aligned}$$

When $t \leq T$, we have

$$\mathbb{E}_{\theta \sim F_t}\left[\Phi\left(\frac{y - \mu}{\sigma}\right)\right] = \mathbb{E}_{\theta \sim F_t} \sum_{i=1}^{3} \Phi\left(\frac{y - \mu_i}{\sigma_i}\right) \pi_t(i)$$

Now we want a function $f : \mathcal{R} \times \Delta(\Theta) \to \mathcal{R}$ such that for a given $F \in \Delta(\Theta)$, f maps a standard normal to Z. Let $Z \sim \mathcal{N}(0, 1)$, then we want a function f such that

$$Y_t = f(Z; F_t)$$

Note that

$$\begin{aligned}
Pr_t(Y_t \leq y) &= Pr(f(Z; F_t) \leq y) \\
Pr_t(Y_t \leq y) &= Pr(Z \leq f^{-1}(y; F_t)) \\
Pr_t(Y_t \leq y) &= \Phi\left(f^{-1}(y; F_t)\right)
\end{aligned}$$

Combining results we have

$$\Phi\left(f^{-1}(y;F_t)\right) = \mathbb{E}_{\theta\sim F_t}\left[\Phi\left(\frac{y-\mu_i}{\sigma_i}\right)\right]$$

Implying

$$\begin{aligned}
f^{-1}(y;F_t) &= \Phi^{-1}\left(\mathbb{E}_{\theta\sim F_t}\left[\Phi\left(\frac{y-\mu}{\sigma}\right)\right]\right) \\
f^{-1}(y;F_t) &= \Phi^{-1}\left(\mathbb{E}_{\theta\sim F_t}\sum_{i=1}^{3}\Phi\left(\frac{y-\mu_i}{\sigma_i}\right)\pi_t(i)\right)
\end{aligned}$$

$$\mathbb{E}_{\theta\sim F_t}\sum_{i=1}^{3}\Phi\left(\frac{y-\mu_i}{\sigma_i}\right)\pi_t(i) \tag{1}$$

can be estimated via gibbs sampling. For each gibb's sweep calculate the term

$$\sum_{i=1}^{3}\Phi\left(\frac{y-\mu_i}{\sigma_i}\right)\pi_t(i)$$

and average over the sample of these draws. The function f^{-1} can then be estimated by calculating equation 1 over a grid of y.

2 Biases

We want to decompose the bias in the forecast for inflation. Following the above notation with y_t denoting inflation at time t and π_{it} denoting the probability of being in state i at time t. The agent has a belief about the probability of being in each state for every date in the data (This is true even within a single Gibb's sweep). The probability vector for a future state is the probability of transistioning to each future state from the current state(rows of A^h), weighted by the probability of being in each current state. The inflation forecast is defined as:

$$\pi_{t+h} = \pi_t A^h \qquad \text{State forecast} \tag{2}$$

$$\mathbb{E}_t[y_{t+h}|\theta \sim F_t] = \mathbb{E}_t[\pi'_{t+h}\mu|\theta \sim F_t] \tag{3}$$

Replacing the time subscript with the conditional distribution for notational reasons

$$\mathbb{E}_{F_t}[y_{t+h}] = \mathbb{E}_{F_t}[\pi'_{t+h}\mu] \tag{4}$$

y_{t+h} is a function of π_{t+h} and μ_t. We can take a first order talyor expansion and then calculate the variance of it to give (similiar to delta method techniques):

$$\mathbb{E}_{F_t}[y_{t+h}] \approx \mathbb{E}_{F_t}[\pi'_{t+h}]\,\mathbb{E}_{F_t}[\mu] + \nabla_\theta(y_{t+h})'\operatorname{Cov}_{F_t}(\theta)\nabla_\theta(y_{t+h}) \tag{5}$$

$$\theta = [\pi'_{t+h}, \mu']' \tag{6}$$

$$\nabla_\theta(y_{t+h})' = [\mu_1, \mu_2, \mu_3, \pi_{1,t+h}, \pi_{2,t+h}, \pi_{3,t+h}] \tag{7}$$

Where the derivatives are evalulated at the true values of of θ

The second term is the bias comming from model uncertainty. This term can be estimated by replacing the terms in $\nabla_\theta(y_{t+h})'$ with the expected values across the Gibbs draws. The covariance term can also be calculated by calculating the covariance of the terms across Gibbs draws.

2.1 Notes

In the above derivation, we collapsed the current state probabilties and transition matrices into a single vector of future state probabilties. We can unwind this and expand the θ vector, but this complicates the $\nabla_\theta(y_{t+h})'$. This is doable just more complicated.

2.2 Questions

1. Does this look correct?

2. What do you mean by "correlate with the convexity of f^{-1}?" Specfically, do you have a convexity measure in mind?