Bayesian Analysis of the AR(1) with Markov Switching

by Hedibert Freitas Lopes
Applied Econometrics, Spring 2005
Graduate School of Business
University of Chicago

Model and Prior

The observed values, y_t, of an AR(1) process with switching regime are generated by,

$$y_t = \alpha_{z_t} + \phi y_{t-1} + \varepsilon_t \qquad \varepsilon_t \sim N(0, \sigma^2)$$

where z_t follows a standard homogeneous first order Markov chain with k states

$$P(z_t = j | z_{t-1} = i) = \omega_{ij} \qquad i, j = 1, \ldots, k$$

where P is the transition matrix, and $\sum_{j=1}^{k} \omega_{ij} = 1$, for all $i = 1, \ldots, k$. By defining $\alpha = (\alpha_1, \ldots, \alpha_k)$ and $\omega_i = (\omega_{i1}, \ldots, \omega_{ik})'$, the parameters of the model are $\theta = (\alpha', \phi, \sigma^2, \omega_1', \ldots, \omega_k')'$ and $z = (z_1, \ldots, z_n)'$.

The prior distribution for the components of θ are

$$\begin{aligned}
\alpha &\sim N(\alpha_0 1_k, V_\alpha I_k) 1(\alpha \in A) \\
\phi &\sim N(\phi_0, V_\phi) \\
\sigma^2 &\sim IG(\nu_0/2, \nu_0 s_0^2/2) \\
\omega_i &\sim Dirichlet(\omega_{0i}) \qquad \text{for } i = 1, \ldots, k,
\end{aligned}$$

where $A = \{\alpha \in \Re^k : \alpha_1 \leq \alpha_2 \leq \cdots \leq \alpha_k\}$.

Conditioning on the latent vector z and defining $I_j = \{t : z_t = j\}$, for $j = 1, \ldots, k$ and $t = 1, \ldots, n$, and $p_N(a|b, c)$ is the density of the univariate normal distribution with mean b and variance c evaluated at a the joint posterior distribution of (θ, z) can be written as

$$\pi(\theta, z | y) \propto \pi(\theta)\pi(z|\theta)p(y|\theta, z)$$

with both

$$\pi(\theta | y, z) \quad \propto \quad \pi(\phi)\pi(\sigma^2) \prod_{j=1}^{k} \pi(\omega_j)\pi(z|\omega) \prod_{j=1}^{k} \left[\prod_{t \in I_j} \pi(\alpha_j) p_N(y_t | \alpha_j + \phi y_{t-1}, \sigma^2) \right]$$

and

$$\pi(z | \theta, y) \quad = \quad \pi(z|\omega) \prod_{i=1}^{n} p_N(y_i | \alpha_j + \phi y_{t-1}, \sigma^2)$$

much more straightforward to sample from. Therefore, θ and z can be easily and iteratively sampled within a Gibbs sampler, as described next.

Full conditionals of α, ϕ, σ^2 and ω

Conditionally on $z = (z_1, \ldots, z_n)'$, the full conditional distributions of α, ϕ, σ^2 and ω are relatively simple to sample from. Let $n_{ij} = \sum_{t=2}^{n} 1(z_{t-1} = i, z_t = j)$, $n_i = (n_{i1}, \ldots, n_{ik})'$, $\nu_i = \sum_{j=1}^{k} n_{ij} \equiv \sum_{t=1}^{n} 1(z_t = i)$, and $I_j = \{t : z_t = j\}$, for $j = 1, \ldots, k$.

- For $j = 1, \ldots, k$,
$$[\alpha_j|\phi, \sigma^2, z, y] \sim N(\alpha_1, \tilde{V}_\alpha)1(\alpha_j \in A_j)$$
where $\tilde{V}_\alpha^{-1} = \nu_i\sigma^{-2} + V_\alpha^{-1}$, $\alpha_1 = \tilde{V}_\alpha(\sigma^{-2}\sum_{t\in I_j}(y_t - \phi y_{t-1}) + V_\alpha^{-1}\alpha_0)$, and $A_j = \{\alpha_j \in \Re : \alpha_{j-1} \leq \alpha_j \leq \alpha_{j+1}\}$. Also, $A_1 = \{\alpha_1 \in \Re : \alpha_1 \leq \alpha_2\}$ and $A_k = \{\alpha_k \in \Re : \alpha_{k-1} \leq \alpha_k\}$.
- $[\sigma^2|\alpha, \phi, z, y] \sim IG(0.5(a + n), 0.5(b + \sum_{t=1}^{n}(y_t - \alpha_{z_t} - \phi y_{t-1})^2))$
- $[\phi|\alpha, \sigma^2, z, y] \sim N(\phi_1, \tilde{V}_\phi)$
where $\tilde{V}_\phi^{-1} = V_\phi^{-1} + \sigma^{-2}\sum_{t=1}^{n} y_{t-1}^2$ and $\phi_1 = \tilde{V}_\phi(V_\phi^{-1}\phi_0 + \sigma^{-2}\sum_{t=1}^{n}(y_t - \alpha_{z_t})y_{t-1})$.
- $[\omega_i|z] \sim D(\omega_{0i} + n_i)$

Full conditionals of z

Conditionally on θ and y, and given the Markovian property of z,

$$\begin{aligned} p(z|\theta, y) &= p(z_n|\theta, y)p(z_{n-1}|z_n, \theta, y)p(z_{n-2}|z_{n-1}, \theta, y)\cdots p(z_1|z_2, \theta, y) \\ &= p(z_n|\theta, y_n)\prod_{t=1}^{n-1} p(z_t|z_{t+1}, \theta, y_t) \end{aligned}$$

Therefore, in order to obtain a sample, say z^*, from $p(z|\theta, y)$ one could cycle through the following *Backward Sampling* scheme:

Sample z_n^* from $p(z_n|\theta, y_n)$,

Sample z_{n-1}^* from $p(z_{n-1}|z_n^*, \theta, y_{n-1})$,

Sample z_{n-2}^* from $p(z_{n-2}|z_{n-1}^*, \theta, y_{n-2})$,

$\vdots$

Sample z_2^* from $p(z_2|z_3^*, \theta, y_2)$, and

Sample z_1^* from $p(z_1|z_2^*, \theta, y_1)$.

The crucial question that remains is *For a given time t, how to sample from $p(z_t|z_{t+1}, \theta, y)$?* Let $p(z_0|\theta)$ be the initial probabilities of z prior to observing any y_t. After t observations are collected and given $p(z_t|\theta, y_t)$, both $p(z_{t+1}|\theta, y_t)$ and $p(z_{t+1}|\theta, y_{t+1})$ can be obtained as

$$p(z_{t+1}|\theta, y_t) = \sum_{i=1}^{k} p(z_{t+1}|z_t, \theta)p(z_t|\theta, y_t)$$

and

$$\begin{aligned} p(z_{t+1}|\theta, y_{t+1}) &= cp(y_{t+1}|z_{t+1}, \theta)p(z_{t+1}|\theta, y_t) \\ &= \frac{p(y_{t+1}|z_{t+1}, \theta)p(z_{t+1}|\theta, y_t)}{\sum_{\zeta=1}^{k} p(y_{t+1}|\zeta, \theta)p(\zeta|\theta, y_t)} \end{aligned}$$

Therefore, $p(z_t|\theta, y_t)$, for $t = 1, \ldots, n$ can be easily obtained by cycling through the above *forward filtering* algorithm. Additionally,

$$\begin{aligned} p(z_t|z_{t+1}, \theta, y) &= p(z_t|z_{t+1}, \theta, y_t) = cp(z_{t+1}|z_t, \theta, y_t)p(z_t|\theta, y_t) \\ &= cp(z_{t+1}|z_t, \theta)p(z_t|\theta, y_t) \\ &= \frac{p(z_{t+1}|z_t, \theta)p(z_t|\theta, y_t)}{\sum_{\zeta=1}^{k} p(z_{t+1}|\zeta, \theta)p(\zeta|\theta, y_t)} \end{aligned}$$

Since $p(z_t|\theta, y_t)$ was obtained by implementing the *forward filtering* just described and $p(z_{t+1}|z_t, \theta)$ is the transition equation from the homogeneous first order Markovian process, $z_{n-1}, z_{n=2}, \ldots, z_2$ and z_1 can be sampled by the *backward sampling* we introduced at the beginning of this Section. The combination of these two steps generates the celebrated *Forward-Filtering, Backward-Sampling* algorithm.

Mixture of normal distributions

It is worth noting that the case particular case where

$$P(z_t = j|z_{t-1} = i) = \omega_j \qquad i, j = 1, \ldots, k$$

corresponds to the AR(1) with mixture of normal errors,

$$y_t = \phi y_{t-1} + \sum_{j=1}^{k} \omega_j N(\alpha_j, \sigma^2)$$