Bayesian Conditional Cointegration: Supplementary

Chris Bracegirdle C.BRACEGIRDLE@CS.UCL.AC.UK
David Barber D.BARBER@CS.UCL.AC.UK
Centre for Computational Statistics and Machine Learning, University College London, Gower Street, London

Here we give the full derivation of the inference recursions for the cointegration switching model.

Filtering

The filtered distribution $\alpha(\phi_t) = p(\phi_t|\epsilon_{1:t})$ will be written as

$$\begin{aligned}\alpha(\phi_t) &= p(\phi_t, i_t = 1|\epsilon_{1:t}) + p(\phi_t, i_t = 0|\epsilon_{1:t}) \\ &= \alpha(\phi_t|i_t = 1)\,\alpha(i_t = 1) + \alpha(\phi_t|i_t = 0)\,\alpha(i_t = 0)\end{aligned}$$

Non-cointegrated case

The easiest case corresponds to $i_t = 1$ so we'll tackle this first.

$$\begin{aligned}p(\phi_t, i_t = 1|\epsilon_{1:t}) &= \frac{p(\phi_t, i_t = 1, \epsilon_t|\epsilon_{1:t-1})}{p(\epsilon_t|\epsilon_{1:t-1})} \\ &= \frac{1}{Z_t} p(\epsilon_t|\epsilon_{t-1}, \phi_t)\, p(\phi_t|i_t = 1)\, p(i_t = 1|\epsilon_{1:t-1})\end{aligned}$$

where $Z_t = p(\epsilon_t|\epsilon_{1:t-1})$. Then

$$\begin{aligned}\alpha(i_t = 1) &= \int_{\phi_t} \frac{1}{Z_t} p(\epsilon_t|\epsilon_{t-1}, \phi_t)\, p^1(\phi_t)\, p(i_t|\epsilon_{1:t-1}) \\ &= \frac{1}{Z_t} \mathcal{N}\left(\epsilon_t|\epsilon_{t-1}, \sigma^2\right) \sum_{i_{t-1}} p(i_t = 1|i_{t-1})\, \alpha(i_{t-1}) \qquad (1)\end{aligned}$$

and the continuous component is trivial,

$$\begin{aligned}\alpha(\phi_t|i_t = 1) &\propto p(\epsilon_t|\epsilon_{t-1}, \phi_t)\, p(\phi_t|i_t = 1) \\ &\Rightarrow \quad \alpha(\phi_t|i_t = 1) = p^1(\phi_t)\end{aligned}$$

Cointegrated case

When $i_t = 0$, the inference is more complicated since the posterior for ϕ_t is non-trivial.

$$\begin{aligned}p(\phi_t, i_t = 0|\epsilon_{1:t}) &= \frac{1}{Z_t} \sum_{i_{t-1}} p(\phi_t, i_t, i_{t-1}, \epsilon_t|\epsilon_{1:t-1}) \\ &= \frac{1}{Z_t} \sum_{i_{t-1}} p(\epsilon_t|\epsilon_{t-1}, \phi_t)\, p(i_t = 0|i_{t-1})\, \alpha(i_{t-1}) \\ &\quad \times \underbrace{\int_{\phi_{t-1}} p(\phi_t|i_t = 0, \phi_{t-1}, i_{t-1})\, \alpha(\phi_{t-1}|i_{t-1})}_{=\begin{cases} p^0(\phi_t) & i_{t-1} = 1 \\ \alpha(\phi_{t-1} = \phi_t|i_{t-1} = 0) & i_{t-1} = 0 \end{cases}}\end{aligned}$$

and the emission term $\mathcal{N}\left(\epsilon_t|\phi_t\epsilon_{t-1}, \sigma^2\right)$

$$= \begin{cases} \mathcal{N}\left(\epsilon_t|0, \sigma^2\right) & \epsilon_{t-1} = 0 \\ \frac{1}{|\epsilon_{t-1}|} \mathcal{N}\left(\phi_t \middle| \frac{\epsilon_t}{\epsilon_{t-1}}, \frac{\sigma^2}{\epsilon_{t-1}^2}\right) & \epsilon_{t-1} \neq 0 \end{cases} \qquad (2)$$

Combining these last two results, we see that an additional component is contributed to the filtering recursion for $i_t = 0$ at each iteration. By further noting that a product of Gaussians is a Gaussian, we see that the posterior is given by a truncated mixture of Gaussians; and therefore we derive a recursion by writing each $\alpha(\phi_t|i_t = 0) = p^0(\phi_t) \sum_{\rho_t} w_{\rho_t} \mathcal{N}(\phi_t|f_{\rho_t}, F_{\rho_t})$. The index ρ_t corresponds to the run-length parameter as set out by Bracegirdle & Barber (2011)—giving the number of timesteps since a switch to cointegration.

Ignoring, for brevity, the simpler cases when $\epsilon_{t-1} = 0$, the new regime $\rho_t = 0$ has the component given in (2); for $\rho_t > 0$ we have a product of Gaussians given by the following (with $\rho_{t-1} = \rho_t - 1$), which we condition

$$\begin{aligned}&\mathcal{N}\left(\phi_t|f_{\rho_{t-1}}, F_{\rho_{t-1}}\right) \mathcal{N}\left(\epsilon_t|\phi_t\epsilon_{t-1}, \sigma^2\right) \\ &= \mathcal{N}\left(\epsilon_t|\epsilon_{t-1} f_{\rho_{t-1}}, \sigma^2 + \epsilon_{t-1}^2 F_{\rho_{t-1}}\right) \\ &\quad \times \mathcal{N}\left(\phi_t \middle| \frac{f_{\rho_{t-1}}\sigma^2 + \epsilon_t\epsilon_{t-1}F_{\rho_{t-1}}}{\sigma^2 + \epsilon_{t-1}^2 F_{\rho_{t-1}}}, \frac{\sigma^2 F_{\rho_{t-1}}}{\sigma^2 + \epsilon_{t-1}^2 F_{\rho_{t-1}}}\right)\end{aligned}$$

After calculating the new Gaussian components for $\alpha(\phi_t | i_t = 0)$, we finish the recursion derivations with

$$\alpha(i_t = 0) = \int_{\phi_t} p(\phi_t, i_t = 0 | \epsilon_{1:t}) \tag{3}$$

Likelihood

The likelihood of the data $p(\epsilon_{1:T})$ is given by

$$p(\epsilon_{1:T}) = p(\epsilon_1) \prod_{t=2}^{T} p(\epsilon_t | \epsilon_{1:t-1}) = p(\epsilon_1) \prod_{t=2}^{T} Z_t$$

where the normalisation Z_t is calculated during the filtering recursion from (1) and (3) since

$$\alpha(i_t = 1) + \alpha(i_t = 0) = 1$$

Whilst it is not necessary to calculate this likelihood, it does permit a common-sense stopping criterion for the EM recursion. The likelihood in practice quickly becomes very small, and it is useful to work in log space to avoid numerical underflow. For the same reason, the weights w_{ρ_t} are also calculated in log space in our implementation.

Smoothing

As with the filtered distribution, we will write the smoothed posterior $\gamma(\phi_t) = p(\phi_t | \epsilon_{1:T})$ as

$$\begin{aligned}\gamma(\phi_t) &= p(\phi_t, i_t = 1 | \epsilon_{1:T}) + p(\phi_t, i_t = 0 | \epsilon_{1:T}) \\ &= \gamma(\phi_t | i_t = 1)\, \gamma(i_t = 1) + \gamma(\phi_t | i_t = 0)\, \gamma(i_t = 0)\end{aligned}$$

The naïve approach, which follows as

$$\gamma(\phi_t, i_t) = \int_{\phi_{t+1}} \sum_{i_{t+1}} p(i_t, \phi_t | \phi_{t+1}, i_{t+1}, \epsilon_{1:t})\, \gamma(\phi_{t+1}, i_{t+1})$$

fails because the first term 'dynamics reversal'

$$\frac{p(\phi_{t+1}, i_{t+1} | i_t, \phi_t)\, \alpha(\phi_t, i_t)}{\int_{\phi_t} \sum_{i_t} p(\phi_{t+1}, i_{t+1} | i_t, \phi_t)\, \alpha(\phi_t, i_t)}$$

has a mixture of components in the denominator for the case $i_{t+1} = 0$. However, exact smoothing can be derived by utilising the interpretation of the run-length index ρ_t from the filtering recursion, as set out by Bracegirdle & Barber (2011).

The filtered distribution was in fact characterised as

$$\begin{aligned}\alpha(\phi_t) = \alpha(\phi_t | i_t = 1)\, \alpha(i_t = 1) \\ + \sum_{\rho_t} \alpha(\phi_t | \rho_t)\, \alpha(\rho_t | i_t = 0)\, \alpha(i_t = 0)\end{aligned}$$

where $\alpha(\phi_t | \rho_t)$ is given by a single truncated Gaussian and $\alpha(\rho_t | i_t = 0)$ is proportional to the weight w_{ρ_t}. The index ρ_t indicates the *number of time-steps since the current regime started.* That is, for fixed ρ_t, we know $i_{t-\rho_t-1} = 1$ and $i_{t-\rho_t:t} = 0$. Between time-steps, $\rho_t = \rho_{t+1} - 1$. Conditioning on ρ_t is equivalent to conditioning on $i_{t-\rho_t-1:t}$, and this extra information already encoded into the components enables us to write a simple recursion for the smoothed posterior.

Non-cointegrated case

We start by first considering $\gamma(\phi_t | i_t = 1) = p^1(\phi_t)$; this can be seen intuitively, or algebraically as

$$\begin{aligned}\gamma(\phi_t | i_t = 1) &\propto p(\phi_t, \epsilon_{t+1:T} | i_t = 1, \epsilon_{1:t}) \\ &\propto \alpha(\phi_t | i_t = 1) = p^1(\phi_t)\end{aligned}$$

since $\phi_t \perp\!\!\!\perp \phi_{t+1} | i_t = 1$.

The discrete component

$$\begin{aligned}\gamma(i_t = 1) = \int_{\phi_{t+1}} p(\phi_{t+1}, \rho_{t+1} = 0, i_{t+1} = 0 | \epsilon_{1:T}) \\ + p(i_t = 1 | i_{t+1} = 1, \epsilon_{1:t})\, \gamma(i_{t+1} = 1)\end{aligned}$$

where the first term is found as the integral of the subset of the components from $t+1$ indexed by $\rho_{t+1} = 0$, and the second

$$p(i_t = 1 | i_{t+1} = 1, \epsilon_{1:t}) \propto p(i_{t+1} = 1 | i_t = 1)\, \alpha(i_t = 1)$$

Cointegrated case

To complete the recursion, it suffices to find the components $\gamma(\phi_t, \rho_t, i_t = 0)$.

$$\begin{aligned}\gamma(\phi_t, i_t = 0) = \int_{\phi_{t+1}} p(\phi_t, i_t = 0, \phi_{t+1}, i_{t+1} = 0 | \epsilon_{1:T}) \\ + p(\phi_t, i_t = 0 | i_{t+1} = 1, \epsilon_{1:t})\, \gamma(i_{t+1} = 1)\end{aligned} \tag{4}$$

and the latter term is easily calculated since

$$\begin{aligned}p(\phi_t, i_t = 0 | i_{t+1} = 1, \epsilon_{1:t}) \\ = \alpha(\phi_t | i_t = 0)\, p(i_t = 0 | i_{t+1} = 1, \epsilon_{1:t})\end{aligned}$$

$$p(i_t = 0 | i_{t+1} = 1, \epsilon_{1:t}) \propto p(i_{t+1} = 1 | i_t = 0)\, \alpha(i_t = 0)$$

Finally, the first term in equation (4) collapses,

$$\begin{aligned}&\int_{\phi_{t+1}} p(\phi_t, i_t = 0, \phi_{t+1}, i_{t+1} = 0 | \epsilon_{1:T}) \\ &= \sum_{\rho_{t+1} > 0} \int_{\phi_{t+1}} p(\phi_t, i_t = 0, \phi_{t+1}, i_{t+1} = 0, \rho_{t+1} | \epsilon_{1:T}) \\ &= \sum_{\rho_{t+1} > 0} \gamma(\phi_{t+1} = \phi_t, \rho_{t+1}, i_{t+1} = 0)\end{aligned}$$

We see that all of the previous components from $\gamma(\phi_{t+1} = \phi_t, \rho_{t+1} > 0, i_{t+1} = 0)$ survive the recursion without change, and that all of the components from $\alpha(\phi_t, i_t = 0)$ are contributed with a prefactor.

References

Bracegirdle, C. and Barber, D. Switch-Reset Models : Exact and Approximate Inference. In *AISTATS*, volume 15. JMLR, 2011.