Homework 2

Due June 1st at 1pm.

For $t = 1, \dots, n$, let us consider the following version of a local-level dynamic linear model:

$$\begin{aligned} y_t &= x_t + v_t \\ x_t &= \phi x_{t-1} + w_t \qquad\qquad w_t \overset{iid}{\sim} N(0, \tau^2), \end{aligned}$$

with $x_0 \sim N(m_0, C_0)$.

We consider two possible structures for $v_1, \dots, v_n$:

$$\mathcal{M}_1 : \; v_t \sim \pi N(0, \sigma^2) + (1 - \pi) N(0, \kappa^2 \sigma^2),$$

$$\mathcal{M}_2 : \; v_t \sim t_\nu(0, \sigma^2),$$

where $\pi \in (0, 1)$, $\kappa > 0$ and $\nu > 0$.

Data augmentation

Notice that $v_t \sim \pi N(0, \sigma^2) + (1 - \pi) N(0, \kappa^2 \sigma^2)$ can be rewritten as

$$v_t | \lambda_t \sim N(0, \sigma^2_{\lambda_t}) \quad \text{and} \quad \lambda_t \overset{iid}{\sim} Ber(\pi)$$

where $\sigma_t^2 = \sigma^2_{\lambda_t}$ with $\sigma_0^2 = \sigma^2$ and $\sigma_1^2 = \kappa^2 \sigma^2$.

Similarly, $v_t \sim t_\nu(0, \sigma^2)$ can be rewritten as

$$v_t | \lambda_t \sim N(0, \lambda_t \sigma^2) \quad \text{and} \quad \lambda_t \sim IG(\nu/2, \nu/2).$$

Conditionally on $\{\lambda_t\}_{t=1}^n$, both models are standard normal dynamic linear models (NDLMs).

Prior

We will assume that

$$\begin{aligned} p(\phi,\sigma^2,\tau^2,\pi,\kappa|\mathcal{M}_1) &= p(\phi)p(\sigma^2)p(\tau^2)p(\pi)p(\kappa) \\ p(\phi,\sigma^2,\tau^2,\nu|\mathcal{M}_2) &= p(\phi)p(\sigma^2)p(\tau^2)p(\nu) \end{aligned}$$

where

$$\sigma^2 \sim IG(\nu_0/2, \nu_0\sigma_0^2/2)$$

$$\tau^2 \sim IG(\eta_0/2, \eta_0\tau_0^2/2)$$

$$\pi \sim U(0,1)$$

$$\kappa^2 \sim IG(a,b)$$

$$\nu \sim \text{uniform on } \{1,2,\ldots,m\}$$

$$p(\phi) \propto 1$$

MCMC

Let $y^n = (y_1, \ldots, y_n)$ and $x^n = (x_1, \ldots, x_n)$.
Let $\theta_1 = (x_0, \phi, \sigma^2, \tau^2, \pi, \kappa)$ and $\theta_2 = (x_0, \phi, \sigma^2, \tau^2, \nu)$.

Derive MCMC schemes to sample from

$$p(x^n, \theta_1 | y^n, \mathcal{M}_1) \quad \text{and} \quad p(x^n, \theta_2 | y^n, \mathcal{M}_2),$$

by taking into account that the latent variables $\{\lambda_t\}_{t=1}^n$ facilitate the derivation of easy-to-sample full conditional distributions.

Simulation exercises

In order to test both algorithms, simulate two sets of $n = 200$ observations, one from $\mathcal{M}_1$ and one from $\mathcal{M}_2$, with the following specifications:

$$\mathcal{M}_1 : \theta_1 = (0.0, 0.9, 1.0, 1.0, 0.9, 2.0)$$

$$\mathcal{M}_2 : \theta_2 = (0.0, 0.9, 1.0, 1.0, 4.0)$$

Run your MCMC schemes for 20,000 draws and discard the first half as burn-in. Use the true values as initial values for the fixed parameters θ_1 and θ_2.

Suggestion

First run your code assuming the fixed parameters are known since

$$p(x^n|\theta, y^n, \mathcal{M}_i), \qquad i = 1, 2,$$

are the more involving full conditionals. Well, not really! You can use the augmented latent variables $(\lambda_1, \ldots, \lambda_n)$ to neatly derive quite standard FFBS schemes.

Then breaking

$$p(\theta|x^n, y^n, \mathcal{M}_i), \qquad i = 1, 2,$$

into univariate full conditionals should be straightforward.

Presentations

Two of the three 3-person groups will be selected at the beginning of the class on June 1st to present the results of the two models (one model each group). The third group is off the hook, but that will be known only right before the presentations.

Be ready to describe your simulation and your code.

Each presentation will last at most 15 minutes.

Sampling x_0, ϕ and τ^2

For both models $\mathcal{M}_1$ and $\mathcal{M}_2$, the full conditional distributions of x_0, ϕ and τ^2 are identical.

$$[\tau^2|\cdots] \sim IG\left(\frac{\eta_0+n}{2}, \frac{\eta_0\tau_0^2+\sum_{t=1}^n(x_t-\phi x_{t-1})^2}{2}\right)$$

$$[x_0|\cdots] \sim N\left\{\left(\frac{1}{C_0}+\frac{\phi^2}{\tau^2}\right)^{-1}\left(\frac{m_0}{C_0}+\frac{\phi x_1}{\tau^2}\right);\left(\frac{1}{C_0}+\frac{\phi^2}{\tau^2}\right)^{-1}\right\}$$

$$[\phi|\cdots] \sim N\left\{\frac{\sum_{t=1}^n x_t x_{t-1}}{\sum_{t=1}^n x_{t-1}^2};\tau^2\frac{1}{\sum_{t=1}^n x_{t-1}^2}\right\}$$

Sampling $x_1, \ldots, x_n$ jointly

For both models $\mathcal{M}_1$ and $\mathcal{M}_2$, the full conditional distribution of $x_1, \ldots, x_n$ is obtained by assuming that σ^2 (the observational variance) is replaced by σ_t^2. Starting with $x_0|D_0 \sim N(m_0, C_0)$, it follows by induction that, given $x_{t-1}|D_{t-1} \sim N(m_{t-1}, C_{t-1})$, then

$$\begin{aligned} x_t|D_{t-1} &\sim N(a_t, R_t), \quad a_t = \phi m_{t-1}, \; R_t = \phi^2 C_{t-1} + \tau^2 \\ y_t|D_{t-1} &\sim N(f_t, Q_t), \quad f_t = a_t, \; Q_t = R_t + \sigma_t^2 \\ x_t|D_t &\sim N(m_t, C_t), \quad m_t = (1 - A_t)a_t + A_t y_t, \; C_t = R_t - A_t^2 Q_t \end{aligned}$$

where $A_t = R_t/Q_t$. Then, backward sampling is performed by first sampling x_n from $N(m_n, C_n)$ and then, for $t = n-1, n-2, \ldots, 3, 2, 1$ sampling

$$x_t|x_{t+1}, D_t \sim N\left\{\left(\frac{\phi^2}{\tau^2} + \frac{1}{C_t}\right)^{-1}\left(\frac{\phi x_{t+1}}{\tau^2} + \frac{m_t}{C_t}\right); \left(\frac{\phi^2}{\tau^2} + \frac{1}{C_t}\right)^{-1}\right\}$$

In model $\mathcal{M}_1$, $\sigma_t^2 = (\kappa^2)^{1-\lambda_t}\sigma^2$ with $\lambda_t = 1$ or $\lambda_t = 0$.
In model $\mathcal{M}_2$, $\sigma_t^2 = \lambda_t \sigma^2$, where $\lambda_t > 0$.
In both cases σ_t^2 equals σ^2 times a function of λ_t.

Sampling σ^2

Using the results from the previous slide, we can easily define

$$\tilde{y}_t = \frac{y_t}{\sqrt{(\kappa^2)^{1-\lambda_t}}} \quad \text{and} \quad \tilde{x}_t = \frac{x_t}{\sqrt{(\kappa^2)^{1-\lambda_t}}}$$

for model $\mathcal{M}_1$, and

$$\tilde{y}_t = \frac{y_t}{\sqrt{\lambda_t}} \quad \text{and} \quad \tilde{x}_t = \frac{x_t}{\sqrt{\lambda_t}}$$

for model $\mathcal{M}_2$. Therefore, under both models, the observation equation become

$$\tilde{y}_t = \tilde{x}_t + u_t \qquad u_t \sim N(0, \sigma^2)$$

whose likelihood for σ^2 is $L(\sigma^2) \propto (\sigma^2)^{-n/2} \exp\left\{-\sum_{t=1}^n (\tilde{y}_t - \tilde{x}_t)^2 / 2\sigma^2\right\}$.

Combining this likelihood with the $IG(\nu_0/2, \nu_0\sigma_0^2/2)$ prior for σ^2, leads to the full conditional

$$[\sigma^2 | \cdots] \sim IG\left(\frac{\nu_0 + n}{2}, \frac{\nu_0\tau_0^2 + \sum_{t=1}^n (\tilde{y}_t - \tilde{x}_t)^2}{2}\right).$$

Sampling π and ν

Under model $\mathcal{M}_1$, the full conditional distribution of π is

$$[\pi|\cdots] \sim Beta\left(\sum_{t=1}^{n} \lambda_t, n - \sum_{t=1}^{n} \lambda_t\right).$$

Under model $\mathcal{M}_2$, the full conditional distribution of ν is

$$Pr(\nu = m|\cdots] \propto \prod_{t=1}^{n} \left[\frac{\left(\frac{m}{2}\right)^{\frac{m}{2}}}{\Gamma\left(\frac{m}{2}\right)} \lambda_t^{-\left(\frac{m}{2}+1\right)} \exp\left\{-\frac{m}{2\lambda_t}\right\}\right]$$

Sampling κ

Without loss of generality, assume that, under model $\mathcal{M}_1$, $\tilde{y}_1, \ldots, \tilde{y}_s$ and $\tilde{x}_1, \ldots, \tilde{x}_s$ are the s observations and corresponding s states where $\lambda_t = 0$.

Let

$$\tilde{\tilde{y}}_t = \frac{\tilde{y}_t}{\sigma} \quad \text{and} \quad \tilde{\tilde{x}}_t = \frac{\tilde{x}_t}{\sigma},$$

for $t = 1, \ldots, s$.
Therefore, the observation equation become

$$\tilde{\tilde{y}}_t = \tilde{\tilde{x}}_t + u_t \qquad u_t \sim N(0, \kappa^2)$$

whose likelihood for κ^2 is $L(\kappa^2) \propto (\kappa^2)^{-s/2} \exp\left\{-\sum_{t=1}^{s}(\tilde{\tilde{y}}_t - \tilde{\tilde{x}}_t)^2/2\kappa^2\right\}$.

Combining this likelihood with the $IG(a, b)$ prior for κ, leads to the full conditional

$$[\kappa^2|\cdots] \sim IG\left(a + s/2, b + (\tilde{\tilde{y}}_t - \tilde{\tilde{x}}_t)^2/2\right).$$

Sampling $\lambda_1, \ldots, \lambda_n$

Under model $\mathcal{M}_1$, it is easy to see that, for $t = 1, \ldots, n$,

$$\begin{aligned} Pr(\lambda_t = 1|\cdots) &\propto \pi p_N(y_t; x_t, \sigma^2) \\ Pr(\lambda_t = 0|\cdots) &\propto (1-\pi) p_N(y_t; x_t, \kappa^2\sigma^2), \end{aligned}$$

so

$$(\lambda_t|\cdots) \sim Ber(\pi_t)$$

where

$$\pi_t = \frac{\pi p_N(y_t; x_t, \sigma^2)}{\pi p_N(y_t; x_t, \sigma^2) + (1-\pi) p_N(y_t; x_t, \kappa^2\sigma^2)}$$

Similarly, under model $\mathcal{M}_2$, it follows that, for $t = 1, \ldots, n$,

$$p(\lambda_t|\cdots) \propto \lambda_t^{-(\nu/2+1)} \exp\{-\nu/2\lambda_t\} p_N(y_t; x_t, \lambda_t\sigma^2),$$

so

$$[\lambda_t|\cdots] \sim IG\left(\frac{\nu+1}{2}, \frac{\nu+(y_t-x_t)^2}{2}\right)$$

MCMC for $\mathcal{M}_1$

Initial values:
$x_0^{(0)} = y_1$, $x_t^{(0)} = y_t$, $\lambda_t^{(0)} = 1$, for $t = 1, \ldots, n$, $\kappa^{(0)} = 1$, $\phi^{(0)} = \hat{\phi}_{ols}$ and $\tau^{2(0)} = \hat{\tau}^2_{ols}$.

Cycle through the following steps:

- $[\sigma^2] \sim IG\left(0.5(\nu_0 + n), 0.5(\nu_0 \tau_0^2 + \sum_{t=1}^n w_t (y_t - x_t)^2)\right)$, with $w_t = \kappa^{\lambda_t - 1}$.
- $[\lambda_t] \sim Ber\left(\frac{\pi p_N(y_t; x_t, \sigma^2)}{\pi p_N(y_t; x_t, \sigma^2) + (1-\pi) p_N(y_t; x_t, \kappa^2 \sigma^2)}\right)$.
- $[\pi] \sim Beta\left(\sum_{t=1}^n \lambda_t, n - \sum_{t=1}^n \lambda_t\right)$.
- $[\tau^2] \sim IG\left(0.5(\eta_0 + n), 0.5(\eta_0 \tau_0^2 + \sum_{t=1}^n (x_t - \phi x_{t-1})^2)\right)$.
- $[x_0] \sim N\{V_0^{-1}\left(C_0^{-1} m_0 + \tau^{-2} \phi x_1\right); V_0^{-1}\}$, where $V_0^{-1} = C_0^{-1} + \phi^2/\tau^2$.
- $[\phi] \sim N\{(\sum_{t=1}^n x_t x_{t-1}) s_x^{-2}; \tau^2 s_x^{-2}\}$, where $s_x^2 = \sum_{t=1}^n x_{t-1}^2$.
- $[\kappa^2] \sim IG\left(a + \frac{s}{2}, b + \frac{\sum_{t=1}^n w_t (y_t - x_t)^2}{2\sigma^2}\right)$, with $w_t = 1 - \lambda_t$ and $s = \sum_{t=1}^n w_t$.