Homework 3

AR(1) with a break
Posterior inference via Gibbs sampler

Instructor: Hedibert Freitas Lopes
Course: STP 598 - Time Series (Class # 12767)
Semester: Spring 2022
Due date: 12pm, March 21st, 2022.

Here is our time series data ($n = 100$ observations):

```
y = c(-25.9,-24.2,-22.0,-22.4,-23.8,-22.0,-27.5,-25.6,-26.4,-27.3,-30.0,-30.3,-24.8,-23.7,
-22.2,-24.5,-22.0,-24.7,-25.7,-22.4,-21.5,-19.2,-14.9,-17.0,-21.7,-26.6,-21.1,-22.6,-20.8,
-19.0,-19.6,-17.2,-10.7,-4.7,0.0,0.5,1.7,0.5,-4.8,0.3,0.1,3.2,-4.2,-7.6,-5.0,-2.6,1.5,-3.0,
-1.5,0.0,-3.9,-4.8,2.0,1.8,2.5,0.0,0.8,2.7,4.7,10.2,1.1,1.3,-3.0,-0.7,4.2,1.6,-4.2,-0.7,
4.2,4.9,0.0,0.2,-2.2,-4.9,3.1,6.7,8.2,9.0,4.8,5.2,7.3,7.7,9.3,6.5,12.7,13.0,16.0,14.8,11.0,
10.4,10.8,5.7,2.0,7.3,4.6,0.8,2.7,0.6,7.0,3.8)
```

For this data, let us consider an AR(1) structure with persistence ϕ_1 for the first τ observations and persistence ϕ_2 for the following $n - \tau$ observations:

$$y_t = \begin{cases} \phi_1 y_{t-1} + \epsilon_t & t = 2, \ldots, \tau \\ \phi_2 y_{t-1} + \epsilon_t & t = \tau + 1, \ldots, n, \end{cases}$$

where $\epsilon_t \sim N(0, \sigma^2)$ for $t = 1, \ldots, n$, and parameters $\theta = (\phi_1, \phi_2, \sigma^2, \tau)$. For simplicity, let us assume the following prior specification:

$$\begin{aligned} p(\theta) &= p(\phi_1)p(\phi_2)p(\sigma^2)p(\tau) \\ \phi_i &\sim N(m_0, C_0), \quad i = 1, 2 \\ \sigma^2 &\sim IG(a_0, b_0) \\ \tau &\sim \text{uniform}\{t_0, t_0 + 1, \ldots, t_1\}, \end{aligned}$$

for known hyperparameters $(m_0, C_0, a_0, b_0, t_0, t_1)$.

(1) Run the MLE regressions for a few values of τ, say $\tau = 25, \ldots, 75$, and plot τ versus $\hat{\phi}$ and τ versus $\hat{\sigma}$. Does this exploratory analysis help you figure out a rough guess for the "best" τ?

(2) For $y_{1:n} = (y_1, \ldots, y_n)$, derive the full conditionals:

(2a) $p(\phi_1 | y_{1:n}, \phi_2, \sigma^2, \tau)$

(2b) $p(\phi_2 | y_{1:n}, \phi_1, \sigma^2, \tau)$

(2c) $p(\sigma^2|y_{1:n}, \phi_1, \phi_2, \tau)$

(2d) $p(\tau|y_{1:n}, \phi_1, \phi_2, \sigma^2)$

Notice that "knowing" τ (conditional on τ) reduces the problem to analyzing two AR(1) processes, one based on the first τ observations and the other one based on the last $n-\tau$ observations.

(3) Let the prior hyperparameters be $m_0 = 0$, $C_0 = 1$, $a_0 = 5$, $b_0 = 6$, $t_0 = 25$ and $t_1 = 75$. With 99% probability, the marginal priors for ϕ_i and σ^2 vary between $(-3, 3)$ and $(0, 5)$.

Perform posterior inference for $(\phi_1, \phi_2, \sigma^2, \tau)$, based on a Gibbs sampler that cycles through the above 4 full conditionals, (2a), (2b), (2c) and (2d).

Report your code/implementation and your posterior summaries and findings. For instance, does your MLE guess for τ in (1) is anywhere near the region of high posterior density of $p(\tau|y_{1:n})$?