Bayesian Dynamic Models for Macro and Finance

Homework Assignment

Topic: MCMC Inference in the AR(1) Plus Noise Model
Due data: April 30th 2026

Overview

This assignment explores two fundamental MCMC strategies for posterior inference on latent states in the **AR(1) plus noise model** (also known as the local level AR(1) model), a canonical dynamic linear model (DLM). You will implement and compare a **single-move Gibbs sampler**, which samples each latent state x_t individually from its full conditional distribution, against the **Forward Filtering Backward Sampling (FFBS)** algorithm, a block-move sampler that draws the entire state sequence jointly from the smoothing distribution. The goal is to develop an operational understanding of the mixing behaviour, computational efficiency, and inferential quality of each approach.

1. The AR(1) Plus Noise Model

The model is defined by the following observation and state equations:

$$y_t = x_t + \varepsilon_t, \qquad \varepsilon_t \sim \mathcal{N}(0, \sigma^2), \tag{1}$$

$$x_t = \alpha + \beta\, x_{t-1} + \omega_t, \quad \omega_t \sim \mathcal{N}(0, \tau^2), \tag{2}$$

for $t = 1, \ldots, n$, with the initial state prior $x_0 \sim \mathcal{N}(m_0, C_0)$. The parameter vector is $\boldsymbol{\theta} = (\alpha, \beta, \sigma^2, \tau^2)$, and the latent state sequence is $\mathbf{x} = (x_1, \ldots, x_n)$.

Parameters used in the reference script:

$n = 100$	Length of the time series
$\alpha = 0.0$	Intercept of the state equation
$\beta = 0.98$	Autoregressive coefficient (near unit root)
$\sigma^2 = 1.0$	Observation noise variance
$\tau^2 = 0.5$	State evolution noise variance

2. Background

2.1 Single-Move Gibbs Sampler

In the single-move (element-wise) Gibbs sampler, the latent states are updated one at a time. The full conditional distribution of x_t given all other states and observations is Gaussian, with mean and variance that depend on the adjacent states x_{t-1} and x_{t+1} (when they exist), and on the current observation y_t. Specifically, the full conditional is:

$$x_t \mid \mathbf{x}_{-t},\, y_{1:n},\, \boldsymbol{\theta} \;\sim\; \mathcal{N}(m_t^*,\, C_t^*),$$

where the conditional mean m_t^* and variance C_t^* are obtained by combining information from the observation equation and the two neighbouring transition densities.

Although conceptually simple and easy to implement, single-move samplers can suffer from **slow mixing** when successive states are highly correlated—a situation common when β is close to 1 (near unit root), as in this exercise.

2.2 Forward Filtering, Backward Sampling (FFBS)

FFBS is a block-move algorithm that samples the entire state sequence $x_{1:n}$ jointly from the smoothing distribution $p(x_{1:n} \mid y_{1:n}, \boldsymbol{\theta})$. It proceeds in two passes:

1. **Forward pass (Kalman filter):** Compute the filtered distributions $p(x_t \mid y_{1:t}, \boldsymbol{\theta})$ for $t = 1, \dots, n$ using the standard Kalman recursions.
2. **Backward pass (Kalman smoother/sampler):** Sample $x_n \sim p(x_n \mid y_{1:n})$, then for $t = n-1, \dots, 1$ draw $x_t \sim p(x_t \mid x_{t+1}, y_{1:t})$. Each backward conditional is Gaussian.

Because FFBS draws all states jointly, it eliminates the between-state correlation that degrades single-move mixing. The Markov chains produced by FFBS typically show **near-zero autocorrelation**, yielding a much larger effective sample size per iteration.

3. Data Simulation

Before implementing either sampler, simulate a dataset using the following R code:

```
set.seed(3141593)
n      <- 100
alpha <- 0.0
beta <- 0.98
sig2   <- 1.0
tau2   <- 0.5
x0     <- alpha / (1 - beta)
x      <- numeric(n);
y <- numeric(n)
x[1]   <- rnorm(1, alpha + beta * x0, sqrt(tau2))
y[1]   <- rnorm(1, x[1], sqrt(sig2))
for (t in 2:n) {
  x[t] <- rnorm(1, alpha + beta * x[t-1], sqrt(tau2))
  y[t] <- rnorm(1, x[t], sqrt(sig2))
}
```

Task: Plot the observed series $y_{1:n}$ and the true latent series $x_{1:n}$ together. Comment briefly on the visual relationship between them and on the difficulty this near-unit-root setting might pose for MCMC.

4. Questions

Question 1 — Single-Move Gibbs Sampler

(a) Derive the full conditional distribution of x_t for each of the three cases: interior states ($1 < t < n$), the first state $t = 1$, and the last state $t = n$. Express the conditional mean and variance in closed form.

(b) Implement the single-move Gibbs sampler in R with the true parameters fixed at $\boldsymbol{\theta} = (\alpha, \beta, \sigma^2, \tau^2) = (0,\ 0.98,\ 1.0,\ 0.5)$. Run the sampler for at least **5,000 iterations** (discarding a suitable burn-in).

(c) Plot the autocorrelation function (ACF) of the Markov chain for at least five representative states: $x_1, x_{25}, x_{50}, x_{75}, x_{100}$. Compute the effective sample size (ESS) for each. What do these diagnostics reveal about the mixing of the chain?

(d) Construct 90% credible bands for $p(x_t \mid y_{1:n})$ using the 5th and 95th posterior percentiles from the sampler. Overlay these on a plot of $y_{1:n}$ and $x_{1:n}$. How many of the true states fall within the credible band?

Question 2 — Forward Filtering, Backward Sampling (FFBS)

(a) Describe, in your own words, the two passes of the FFBS algorithm. What distributional family does each backward conditional belong to, and how are its parameters computed from the Kalman filter output?

(b) Implement FFBS in R (again with parameters fixed at their true values). Run the sampler for the same number of iterations as in Question 1.

(c) Repeat the ACF and ESS analysis of Question 1(c) for the FFBS chains. Compare the ACF plots side by side with those from the single-move sampler. Quantitatively, how much larger is the ESS per iteration under FFBS?

(d) Reconstruct the 90% credible bands as in Question 1(d) using the FFBS draws. Do the two sets of bands differ visually? If so, offer an explanation.

Question 3 — Comparative Analysis

(a) For each of the $n = 100$ time points, compute the posterior mean $\mathbb{E}[x_t \mid y_{1:n}]$ under both samplers. Create a scatter plot of the two sets of posterior means against each other, and against the true latent states x_t^{true}. Report the root mean squared error (RMSE) relative to the truth.

(b) The near-unit-root setting ($\beta = 0.98$) is known to be challenging for single-move Gibbs samplers. Using your ACF results, explain *why* this happens in terms of the conditional independence structure of the model.

(c) Measure the wall-clock runtime (in seconds, using `system.time()` or `proc.time()`) for both samplers over 1,000 iterations. Discuss the trade-off between per-iteration cost and mixing quality. Under what practical circumstances might one prefer the single-move sampler despite its poorer mixing?

(d) [Bonus] Repeat the entire comparison for $\beta \in \{0.5, 0.9, 0.99, 1.0\}$, keeping all other parameters fixed. How does the relative advantage of FFBS over the single-move sampler evolve as β approaches 1? Summarise your findings in a table reporting ESS per iteration for each value of β under both samplers.

Question 4 — Conceptual and Theoretical Questions

(a) Show that the full conditional distribution $p(x_t \mid \mathbf{x}_{-t}, y_{1:n}, \boldsymbol{\theta})$ is Gaussian by combining the log-densities of the relevant factors. Identify which factors in the joint density $p(x_{1:n}, y_{1:n} \mid \boldsymbol{\theta})$ involve x_t.

(b) Explain why both samplers target the same posterior distribution $p(x_{1:n} \mid y_{1:n}, \boldsymbol{\theta})$, even though they proceed very differently. What property of Markov chain Monte Carlo guarantees this?

(c) In the FFBS backward pass, the conditional $p(x_t \mid x_{t+1}, y_{1:t})$ is used. Derive its mean and variance in terms of the Kalman filter quantities m_t, C_t and the one-step-ahead predictive quantities a_{t+1}, R_{t+1}. (Follow the notation used in the reference script.)

(d) The AR(1) plus noise model is a special case of a general Normal DLM. Write down the general DLM system matrices (F_t, G_t, V_t, W_t) that correspond to this model. What are the scalar values of F, G, V, and W for the parametrisation above?

5. Deliverables

Submit the following as a single PDF or compiled R Markdown / Quarto document:

1. **R code** — Well-commented, self-contained scripts for the simulation, both samplers, and all figures.
2. **Mathematical derivations** — Answers to all theoretical sub-questions. You may use standard mathematical notation; handwritten scans are acceptable if legible.
3. **Figures** — All required plots (simulated data, ACF panels, credible band panels, scatter plots). Each figure must have a caption and appropriate axis labels.
4. **Discussion** — For each comparative question, a brief written explanation (3–8 sentences) interpreting the results.

6. Hints and Guidance

- **Full conditional for interior states.** The log full conditional of x_t (for $1 < t < n$) is a sum of three Gaussian log-densities: the observation model, the forward transition, and the backward transition. Complete the square to identify the posterior mean and precision.
- **Kalman filter notation.** Use $m_t = \mathbb{E}[x_t \mid y_{1:t}]$, $C_t = \text{Var}(x_t \mid y_{1:t})$, $a_t = \mathbb{E}[x_t \mid y_{1:t-1}]$, $R_t = \text{Var}(x_t \mid y_{1:t-1})$ to be consistent with the reference script.
- **Effective Sample Size.** You may use `coda::effectiveSize()` or the formula

$$\text{ESS} = \frac{S}{1 + 2\sum_{k=1}^{\infty} \rho_k},$$

where S is the number of samples and ρ_k is the lag-k autocorrelation.
- **Reproducibility.** Set `set.seed(3141593)` at the beginning of your script, as in the reference, to ensure your simulated data matches the reference output.
- **Reference script.** The annotated R script by Hedibert Freitas Lopes is available at `hedibert.org/.../ar1plusnoise.html`. You may use it as a starting point, but your submission must demonstrate your own understanding through clear comments and written explanations.