Posterior Sampling in Two Classes of Multivariate Fractionally Integrated Models: Corrigendum to Ravishanker, N. and B. K. Ray (1997) *Australian Journal of Statistics* 39 (3), 295-311

Ross Doppelt and Keith O'Hara*

November 21, 2018

Abstract

We discuss posterior sampling for two distinct multivariate generalizations of the univariate ARIMA model with fractional integration. The existing approach to Bayesian estimation, introduced by Ravishanker and Ray (1997), claims to provide a posterior-sampling algorithm for fractionally integrated vector autoregressive moving averages (FIVARMAs). We show that this algorithm produces posterior draws for vector autoregressive fractionally integrated moving averages (VARFIMAs), a model of independent interest that has not previously received attention in the Bayesian literature.

Keywords: Long Memory, Multivariate Time Series, Bayesian Inference, Posterior Sampling

*Doppelt (corresponding author): Penn State. E-mail: ross.doppelt@psu.edu. O'Hara: New York University. E-mail: keith.ohara@nyu.edu. We thank David Childers and Nalini Ravishanker for discussions and correspondence; any errors are our own.

1. Introduction. Ravishanker and Ray (1997) provide the first attempt at using Bayesian methods to estimate a multivariate fractionally integrated ARIMA model. However, we will show that Ravishanker and Ray's inferential procedure corresponds to a model that's different from the one they specify. Below, we will discuss two distinct multivariate generalizations of the fractional ARIMA model that have substantially different properties: the FIVARMA and the VARFIMA. Ravishanker and Ray seek to estimate the former, but their procedure produces estimates for the latter. In this note, we make two contributions. The first is simply to correct an error in the literature. The second contribution is to show that an algorithm that was intended to estimate FIVARMA processes can be modified to estimate certain VARFIMA processes, a distinct class of models of independent interest. This insight is potentially useful, because we are unaware of other papers that discuss Bayesian VARFIMA estimation. In a companion paper, Doppelt and O'Hara (2018), we develop new methods for Bayesian estimation of certain FIVARMA models.

2. Models. Let $\mathbf{x}_t$ be an $n \times 1$ vector with mean zero. Let L denote the lag operator, and define $\mathbf{\Phi}(L) \equiv \mathbf{I}_n - \sum_{\ell=1}^{p} \mathbf{\Phi}_\ell L^\ell$, $\mathbf{\Theta}(L) \equiv \mathbf{I}_n + \sum_{\ell=1}^{q} \mathbf{\Theta}_\ell L^\ell$, and $\mathbf{D}(L) \equiv \operatorname{diag}\left((1-L)^{d_1}, \cdots, (1-L)^{d_n}\right)$. First, consider the fractionally integrated vector autoregressive moving average (FIVARMA) model:

$$\mathbf{\Phi}(L)\,\mathbf{D}(L)\,\mathbf{x}_t = \mathbf{\Theta}(L)\,\boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \overset{\text{i.i.d.}}{\sim} \mathrm{N}\left(\mathbf{0}_{n\times 1}, \mathbf{\Sigma}\right). \tag{1}$$

Alternatively, consider the vector autoregressive fractionally integrated moving average (VARFIMA) model:

$$\mathbf{D}(L)\,\mathbf{\Phi}(L)\,\mathbf{x}_t = \mathbf{\Theta}(L)\,\boldsymbol{\epsilon}_t, \quad \boldsymbol{\epsilon}_t \overset{\text{i.i.d.}}{\sim} \mathrm{N}\left(\mathbf{0}_{n\times 1}, \mathbf{\Sigma}\right). \tag{2}$$

Assuming that $\mathbf{\Phi}(L)$ is not diagonal, Lobato (1997) shows that the FIVARMA and VARFIMA models will only coincide when $d_1 = \cdots = d_n$. The distinction is important. In the FIVARMA model, each series can have a different order of integration: $x_{i,t} \sim I(d_i)$. But in the VARFIMA model, each series will have the same order of integration, even if the d_i are distinct: $x_{i,t} \sim I\left(\max_{1\le k\le n}\{d_k\}\right)$. Ravishanker and Ray (1997) claim to perform Bayesian estimation of (1). We will show that their sampling algorithm requires $q = 0$, and then, it actually produces draws from the posterior distribution of (2). When $q = 0$, we will refer to (1) as a FIVAR and (2) as a VARFI, following Sela and Hurvich (2009).

3. Representation and Inference. Let $\mathbf{X}_T \equiv (\mathbf{x}_1', \ldots, \mathbf{x}_T')'$ be the observed sample, and let $\boldsymbol{\theta}$ be the parameters. To estimate (1), Ravishanker and Ray (1997) define $\mathbf{a}_t \equiv \mathbf{D}(L)^{-1}\boldsymbol{\epsilon}_t$ and propose a change of variables based on the representation $\mathbf{\Phi}(L)\,\mathbf{x}_t = \mathbf{\Theta}(L)\,\mathbf{a}_t$. This approach is motivated by the fact that the ACF of $\mathbf{a}_t$ is easier to compute than the ACF of $\mathbf{x}_t$. If $\mathbf{\Phi}(L)\,\mathbf{x}_t = \mathbf{\Theta}(L)\,\mathbf{a}_t$, then $\left(\mathbf{X}_{-p}', \mathbf{A}_{-q}', \mathbf{A}_T'\right)'$ is a linear transformation of $\left(\mathbf{X}_{-p}', \mathbf{A}_{-q}', \mathbf{X}_T'\right)'$ with unit Jacobian, where $\mathbf{X}_{-p} \equiv \left(\mathbf{x}_{-p+1}', \ldots, \mathbf{x}_0'\right)'$, $\mathbf{A}_{-q} \equiv \left(\mathbf{a}_{-q+1}', \ldots, \mathbf{a}_0'\right)'$, and $\mathbf{A}_T \equiv (\mathbf{a}_1', \ldots, \mathbf{a}_T')'$. Thus, $\mathbb{P}[\mathbf{X}_{-p}, \mathbf{A}_{-q}, \mathbf{X}_T \mid \boldsymbol{\theta}] = \mathbb{P}[\mathbf{X}_{-p}, \mathbf{A}_{-q}, \mathbf{A}_T \mid \boldsymbol{\theta}]$. Ravishanker and Ray's posterior sampler treats $\mathbf{X}_{-p}$ and $\mathbf{A}_{-q}$ as latent variables. Letting $\mathbf{A}_* \equiv \left(\mathbf{A}_{-q}', \mathbf{A}_T'\right)'$, the posterior kernel therefore takes the form $\mathbb{P}[\boldsymbol{\theta}, \mathbf{X}_{-p}, \mathbf{A}_{-q} \mid \mathbf{X}_T] \propto \mathbb{P}[\mathbf{X}_{-p} \mid \mathbf{A}_*, \boldsymbol{\theta}]\, \mathbb{P}[\mathbf{A}_* \mid \boldsymbol{\theta}]\, \mathbb{P}[\boldsymbol{\theta}]$. It's feasible to compute $\mathbb{P}[\mathbf{X}_{-p} \mid \mathbf{A}_*, \boldsymbol{\theta}]$ and $\mathbb{P}[\mathbf{A}_* \mid \boldsymbol{\theta}]$, because $\left(\mathbf{X}_{-p}', \mathbf{A}_*'\right)'$ are jointly Gaussian, and Ravishanker and Ray provide recursions for computing the covariance matrix. Conditional on the data adhering to the representation $\boldsymbol{\Phi}(L)\,\mathbf{x}_t = \boldsymbol{\Theta}(L)\,\mathbf{a}_t$, with $\mathbf{a}_t \equiv \mathbf{D}(L)^{-1}\,\boldsymbol{\epsilon}_t$, Ravishanker and Ray's calculations are correct.

However, applying $\mathbf{D}(L)^{-1}$ to both sides of equation (1) does not yield $\boldsymbol{\Phi}(L)\,\mathbf{x}_t = \boldsymbol{\Theta}(L)\,\mathbf{a}_t$, because the matrix-valued lag polynomials do not, in general, commute. If we apply the operator $\mathbf{D}(L)^{-1}$ to both sides of equation (2), then we do obtain $\boldsymbol{\Phi}(L)\,\mathbf{x}_t = \boldsymbol{\Theta}(L)\,\mathbf{a}_t$ in the particular case where $\boldsymbol{\Theta}(L) = \mathbf{I}_n$. Hence, one can apply the Ravishanker-Ray algorithm to multivariate models when $q = 0$, as long as one recognizes that the estimates correspond to a VARFI model, not a FIVAR model. Alternatively, one could directly specify $\boldsymbol{\Phi}(L)\,\mathbf{x}_t = \boldsymbol{\Theta}(L)\,\mathbf{D}(L)^{-1}\,\boldsymbol{\epsilon}_t$ as the data-generating process; in that case, the low-frequency properties of the model would resemble a VARFIMA, in the sense that all variables would have the same order of integration, given by the maximal element of $\{d_i\}_{i=1}^n$.

4. Discussion. Whether it's preferable to fit (1) or (2) depends on the application. For example, in Doppelt and O'Hara (2018), we use a Bayesian FIVAR to analyze a dataset with evidence of fractional integration in some series, but not others, so a VARFI would be inappropriate. Alternatively, Sela and Hurvich (2009) examine two inflation indices using frequentist methods and find that a VARFI fits better than a FIVAR. The Bayesian literature has not given as extensive consideration to these classes of models, partly because sampling algorithms for VARFIs have remained an open question. However, by using Ravishanker and Ray's algorithm that was initially developed for FIVARs, one can perform posterior inference for VARFIs.

References

Doppelt, R. and K. O'Hara (2018). Bayesian estimation of fractionally integrated vector autoregressions and an application to identified technology shocks. Technical report, New York University.

Lobato, I. N. (1997). Consistency of the averaged cross-periodogram in long memory series. *Journal of Time Series Analysis 18*(2), 137–155.

Ravishanker, N. and B. K. Ray (1997). Bayesian analysis of vector ARFIMA processes. *Australian Journal of Statistics 39*(3), 295–311.

Sela, R. J. and C. M. Hurvich (2009). Computationally efficient methods for two multivariate fractionally integrated models. *Journal of Time Series Analysis 30*(6), 631–651.