A paper review by[*]

Jacob Rider

December 2020

Abstract

This paper summarizes an efficient jump MCMC method used to model interest-rate dynamics. The motivation is to define several components of the nested model, while also describing the contributions of this research.

Elena Goldman*
Department of Finance and Economics
Lubin School of Business, Pace University
New York, NY 10038
e-mail: egoldman@pace.edu, tel: 212-618-6516

and

Terence D. Agbeyegbe
Department of Economics
Hunter College, CUNY
695 Park Avenue
New York, NY 10021
e-mail: tagbeyeg@hunter.cuny.edu, tel: 212-772-5405

April 7, 2005

[1]

[*]This paper was written as a final report for STAT8674
[1]Original paper author credit

1 Introduction

This review is based on the Estimation of Threshold Time-Series models using Efficient Jump MCMC, by Elena Goldman and Terence D. Agbeyegbe. The aforementioned paper proposes a nested threshold MCMC model that is calibrated to US 3-month interest rate data. The motivation for this model is to explain the heteroskedasticity shown in the interest rate data. To account for this, the model nests five components:

1. Stochastic Differential Equation (SDE)
2. Threshold Autoregressive (TAR)
3. Autoregressive Moving Average (ARMA)
4. Autoregressive Conditionally Heteroscedastic (ARCH)
5. Generalized Autoregressive Conditionally Heteroscedastic (GARCH)

This review will spend some effort to define each of these component models individually (*primarily for the author's education*).

1.1 Motivation

Chan et al. [2] stated:

> " The short-term riskless interest rate is one of the most fundamental and important prices determined in financial markets."

This quote sheds light on the importance of studying interest rate data. The value of entire markets shift in response to slight fluctuations in interest rates. In addition to the importance of studying interest rates, this paper also presents a generalized MCMC method for time-series data with k+1 regimes.

2 Background

2.1 SDE

$$dy_t = (\mu + y_t)dt + \sigma y_t^{\lambda} dW \tag{1}$$

Where the variables are defined as follows: y_t = interest rate at time t ; dW = derivative of brownian motion ; μ = mean drift ; β = mean reversion drift ; σ = volatility parameter ; and λ = elasticity of volatility based on interest rate level.[2]

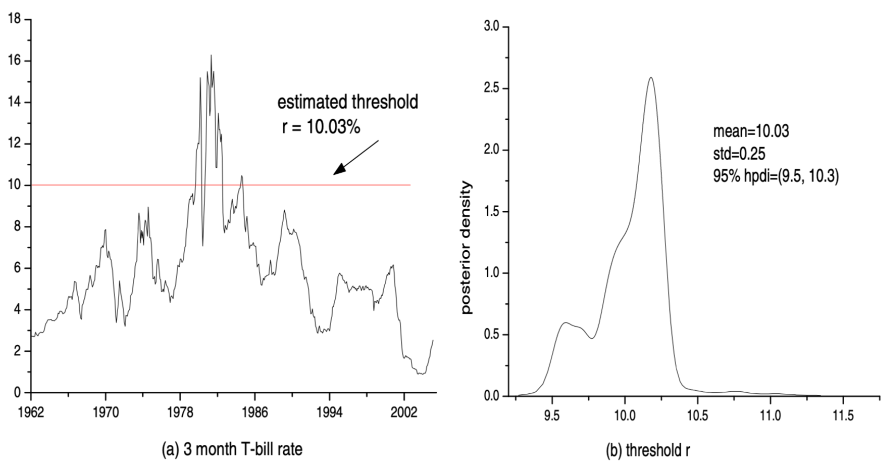
[2]Equations in this paper are sourced from the original paper or its citations.

2.2 TAR

The idea behind the threshold autoregressive component to this model is to separate the time-series data into regimes then calculate each parameter separately across the regimes. The subject paper explains the threshold approach for a generalized case with K+1 regimes, but this review will only focus on the relevant case of two regimes. With two regimes there is only one threshold value, r such that:

$$\begin{cases} j = 1 & y_t < r \\ j = 2 & r < y_t \end{cases} \tag{2}$$

Each parameter is approximated for both regimes. The regimes are visualized by the following result from the paper:



As shown in the graph above, there are clearly two regimes that separate the interest rate data. The upper regime is characterized by high volatility and strong mean reversion, whereas the lower regime is characterized by a moving average process with weak mean reversion.

2.3 ARMA

An Autoregressive Moving Average process is comprised of two polynomials: AR(p) and MA(q). The AR(p) process is structured as follows: [3]

$$R_t = \epsilon_t + \sum_{i=1}^{p} \phi_i R_{t-i} \tag{3}$$

[3]These equations are sourced from Penn State's online statistics course [6]

Similarly, the MA(q) process is structured as follows:

$$R_t = \mu + \epsilon_t + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} \tag{4}$$

Combined, the ARMA(p,q) process is structured as follows:

$$R_t = \mu + \epsilon_t + \sum_{i=1}^{p} \phi_i R_{t-i} + \sum_{i=1}^{q} \theta_i \epsilon_{t-i} \tag{5}$$

ϕ_i are the autoregressive parameters, while θ_i are the moving average parameters. The AR(p) process describes the tendency of a time-series to oscillate around the value of the past p observations. The MA(q) process shows that the expectation of a time series is dependent on the random noise from the past q observations. This random noise at time t is denoted as ϵ_t and is generally assumed to be identically independently distributed from a normal distribution with mean = 0 and variance = σ^2 or,

$$\epsilon_t \sim N(0, \sigma^2) \tag{6}$$

2.4 ARCH

Autoregressive Conditionally Heteroscedastic models are used to account for a time-series with changing variance over time. The ARCH parameter represents the memory of the process. In other words, for ARCH(n), the variance at each time is dependent on the past n interest rate values. In general, the equation for an ARCH(n) process is structured as follows:

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + ... + \alpha_n y_{t-n}^2 \tag{7}$$

2.5 GARCH

The difference between ARCH and Generalized ARCH (GARCH) models is relatively simple. In ARCH models the variance is only dependent on the past n observations of interest rate levels, whereas in GARCH models, the variance is dependent on the past n observations of interest rate and variance. The parameters of the GARCH(n,m) process represent the memory of the process with respect to past interest rate and variance observations. In general, the equation for a GARCH(n,m) process is structured as follows: [4]

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 + ... + \alpha_n y_{t-n}^2 + \beta_m \sigma_{t-m}^2 \tag{8}$$

[4] These equations are sourced from Penn State's online statistics course [6]

3 Nested Model

The ARMA-GARCH parameters were selected by minimizing the modified bayesian information criterion (MBIC) and the result is an ARMA(1,1)-GARCH(1,1) process. These parameters should be estimated for each regime separately, however, the process resulted in the same parameters for both regimes in this case. Recall that this means the process has a memory of one observation for both the interest rate and variance. This results in the following equations:

$$R_t = \mu + \epsilon_t + \phi_1 R_{t-1} + \theta_1 \epsilon_{t-1} \tag{9}$$

$$\sigma_t^2 = \alpha_0 + \alpha_1 y_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \tag{10}$$

In addition to selecting the ARMA-GARCH parameters, parameters for threshold and elasticity were approximated using an efficient jump MCMC algorithm. Essentially this algorithm involves generating new threshold values normally distributed around the past values as shown in the equation below:

$$r_j^i \sim N(r_j^{i-1}, std[r_j^{i-1}]) \tag{11}$$

Then according to the Metropolis-Hastings algorithm, accept the new value (r_j^i) with probability:

$$P(accept) = min[1, \frac{p(\gamma^i, \phi^i, \theta^i, \alpha^i, \beta^i, r^i, \lambda^{i-1}) \mid \text{data}}{p(\gamma^i, \phi^i, \theta^i, \alpha^i, \beta^{i-1}, r^{i-1}, \lambda^{i-1}) \mid \text{data}}] \tag{12}$$

Where p is the posterior distribution referenced in the paper. The paper writes the posterior distribution in a general form for k+1 regimes, however the simpler formula for 2 regimes is not given. This paper review will attempt to derive a simpler formula from the one presented in the paper. Note that x_t and y_t are samples from the data-set. [5]

$$\begin{aligned} \pi(\gamma, \phi, \theta, \alpha, \beta) \quad &= \quad \prod_{i=1}^{k+1} \mathrm{N}(\gamma_{0i}, \Sigma_{\gamma_i}) \times \mathrm{N}(\phi_{0i}, \Sigma_{\phi_i}) \times \mathrm{N}(\theta_{0i}, \Sigma_{\theta_i}) \times \mathrm{N}(\alpha_{0i}, \Sigma_{\alpha_i}) \\ &\quad \times \mathrm{N}(\beta_{0i}, \Sigma_{\beta_i}) \times I\left(0 \leq \lambda^{(i)} \leq \lambda^{up}\right) \times I\left(r_i \in [r_{\text{low}}^{(i)}, r_{\text{up}}^{(i)}]\right) \end{aligned}$$

[5] These formulas are screenshots from the aforementioned paper.

Let us rescale variables y_t^j and x_t^j by the level heteroscedasticity component $|y_{t-1}|^{\lambda^j}$, $(j = 1, ..., k+1)$:

$$\begin{aligned} \tilde{y}_t^j &= y_t^j / |y_{t-1}|^{\lambda^j} \\ \tilde{x}_t^j &= x_t^j / |y_{t-1}|^{\lambda^j} \end{aligned} \tag{7}$$

The posterior distribution is given by

$$p(\gamma, \phi, \theta, \alpha, \beta | data) = \pi(\gamma, \phi, \theta, \alpha, \beta) \prod_{j=1}^{k+1} \prod_{t \in T_j} \frac{1}{\sigma_t |y_{t-1}|^{\lambda^j}} \phi \left(\frac{\tilde{y}_t^j - g(Z_t)}{\sigma_t} \right) \tag{8}$$

where for every $t \in T_j = \{t : r_{j-1} \leq y_{t-d} \leq r_j\}$

$$\begin{aligned} e_t &= \tilde{y}_t^j - \tilde{x}_t^j \gamma^j \\ \epsilon_t &= \tilde{y}_t^j - g(Z_t) \\ g(Z_t) &= \tilde{x}_t \gamma^j - \sum_{i=1}^{p} \phi_i^j e_{t-i} - \sum_{i=1}^{q} \theta_i^j \epsilon_{t-i} \\ \sigma_t^2 &= \alpha_0^j + \sum_{i=1}^{r} \alpha_i^j \epsilon_{t-i}^2 + \sum_{i=1}^{s} \beta_i^j \sigma_{t-i}^2 \end{aligned}$$

For the relevant case with two regimes and ARMA(1,1)-GARCH(1,1) parameters, these formulas become much simpler. Recalling equation 2 $\Rightarrow$ 13:

$$\begin{cases} j = 1 & y_t < r \\ j = 2 & r < y_t \end{cases} \tag{13}$$

The posterior distribution is now given by:

$$p(\gamma, \phi, \theta, \alpha, \beta, |\text{data}) = \pi(\gamma, \phi, \theta, \alpha, \beta) \cdot [\prod_{t \in T_1} \frac{\phi(y_t^{(1)} - g(Z_t))}{\sigma_t^2 |y_{t-1}|^{\lambda^1}}] \cdot [\prod_{t \in T_2} \frac{\phi(y_t^{(2)} - g(Z_t))}{\sigma_t^2 |y_{t-1}|^{\lambda^2}}] \tag{14}$$

where $t \in T_1 = \{t : y_{t-1} \leq r\}$ and $t \in T_2 = \{t : r \leq y_{t-1}\}$
Plugging in the ARMA(1,1)-GARCH(1,1) parameters gives:

$$g(Z_t) = \tilde{x}_t \gamma^{(j)} - \phi_1^{(j)} e_{t-1} - \theta_1^{(j)} \epsilon_{t-1} \tag{15}$$

$$\sigma_t^2 = \alpha_0^{(j)} + \alpha_1^{(j)}\epsilon_{t-1}^2 + \beta_1^{(j)}\sigma_{t-1}^2 \tag{16}$$

$$e_t = \tilde{y}_t^{(j)} - \tilde{x}_t^{(j)}\gamma^{(j)} \tag{17}$$

$$\epsilon_t = \tilde{y}_t^{(j)} - g(Z_t) \tag{18}$$

Naturally, this gives the following recursive equation:

$$g(Z_t) = \tilde{x}_t\gamma^{(j)} - \phi_1^j(\tilde{y_{t-1}} - \tilde{x_{t-1}}\gamma^{(j)}) - \theta_1^j(\tilde{y_{t-1}} - g(Z_{t-1})) \tag{19}$$

Equation 19 gives some intuition for why ϕ is interpreted as the autoregressive parameter, wheareas θ is interpreted as the moving average parameter.

3.1 Results

The resulting parameters for each regime are shown below:

	Regime 1: $(y_{t-1} < r)$			Regime 2: $(y_{t-1} \geq r)$		
	mean	(std)	Corr	mean	(std)	Corr
γ_1	0.047	(0.030)	0.447	2.405	(1.545)	0.491
γ_2	-0.006	(0.007)	0.442	-0.208	(0.131)	0.507
ϕ_1	0.507	(0.138)	0.918	-0.123	(0.432)	0.886
θ_1	-0.129	(0.166)	0.927	0.740	(0.403)	0.876
α_0	0.001	(0.000)	0.771	0.029	(0.024)	0.777
α_1	0.142	(0.030)	0.796	0.328	(0.215)	0.657
β_1	0.835	(0.030)	0.883	0.329	(0.229)	0.826
λ	0.391	(0.130)	0.799	0.638	(0.137)	0.791
ρ=max AR root	0.507	(0.138)	0.918	0.396	(0.212)	0.665
r	10.034	(0.246)	0.315			

It is important to note how the values of these parameters describe qualitative features of each regime. For instance, the high volatility of regime 2 is characterized by the mean of $\lambda = .638$, whereas the strong mean-reversion in regime 2 is characterized by $\phi_1 = -.123$ and $\gamma_1 = 2.405$. In contrast, the persistence of interest rates in the lower regime are shown by: $\beta_1 = .835$.

4 Conclusion

This paper review attempted to layout the general framework of this nested MCMC model with efficient jumping. The original paper presented the model in the general case of K+1 regimes, but determined the time-series data on US-interest rates only required 2 regimes. In an attempt to understand the model more clearly, this review presented the explicit model equations for the case of 2 regimes.

One main insight from this paper is that interest rates become more persistent and less volatile at lower levels. In other words, the interest rate process has a longer memory but behaves less erratically at low levels. In contrast, the upper regime is characterized by high volatility and strong mean-reversion.

5 References

1. Goldman, Elena, and Terence D. Agbeyegbe. "Estimation of threshold time series models using efficient jump MCMC." Bayesian Statistics and Its Applications, edited by Upadhyway, SK, Singh, U. and Dey, DK (2007): 241-253.

2. Chan, Kalok C., et al. "An empirical comparison of alternative models of the short-term interest rate." The journal of finance 47.3 (1992): 1209-1227.

3. Gelman, Andrew, et al. Bayesian data analysis. CRC press, 2013.

4. Pfann, Gerard A., Peter C. Schotman, and Rolf Tschernig. "Nonlinear interest rate dynamics and implications for the term structure." Journal of Econometrics 74.1 (1996): 149-176.

5. Nakatsuma, Teruo. "Bayesian analysis of ARMA–GARCH models: A Markov chain sampling approach." Journal of Econometrics 95.1 (2000): 57-69.

6. "11.1 ARCH/GARCH Models: STAT 510." PennState: Statistics Online Courses, online.stat.psu.edu/stat510/lesson/11/11.1.