TVAR Spectral Analysis of WTI Oil Price

Wenwen Ye

March 2019

1 Introduction

WTI crude oil is a type of unrefined petroleum whose price is usually used as a benchmark for oil price. In the financial market, it serves as an important index for investors in futures and options market.

In this project, we are interested in using the AR(p) process to filter out the noise or volatility from the trend of crude oil price data. This will be implemented by finding the frequency that maximize the spectral density function (i.e. the mode of the density function) under time-varying AR coefficients ϕ and v. The coefficients will be estimated by the posterior analysis.

The goal of this application includes making predictions on the price(synthetic future) and trading strategies corresponding to the results of analysis.

2 Method

2.1 Spectral Density

The spectral density of an assumed-stationary AR(p) process is as follow:
$f(\omega) = \frac{v}{2\pi|(1-\phi_1 e^{-i\omega} - \dots - \phi_p e^{-ip\omega})|^2}$ (P&W p.104)
where v and ϕ are estimated by the posterior means from the analysis above.
Given the spectral density with current $\theta = (\phi\ ,\ v)$, there exists an ω_{max} that maximize the spectral density $f(\omega)$, which can be found by bisection searching in interval $[0,2\pi]$. Therefore, the corresponding wavelength $\lambda = \frac{2\pi}{\omega_{max}}$ maximizes the spectral density function.

2.2 Time-varying AR spectral Analysis

We can perform the posterior analysis assuming the following normal-gamma conjugate priors:

$\phi|v, D_{t-1} \sim N(m_{t-1}, \frac{Cv}{s-t-1})$

$v^{-1}|D_{t-1} \sim Ga(\frac{n_{t-1}}{2}, \frac{n_{t-1}s_{t-1}}{2})$

Consider θ as a time-varying variables such that $\theta_t = (\phi_t, v_t) \sim p(\theta_t|D_T)$ and θ_t evolves through $\theta_{t+1} = G_t\theta_t + \epsilon_t$, where ϵ_t follows $N(0, W_t)$. The full architecture of TVAR model refers to Prado and West[1], and is abbreviated here due to length limit.

Plug each posterior mean θ_t, $t = 1, \ldots, n$ into the spectral density function will produce a spectral density for each time point t.

Then combine spectral densities across t, we should have a three dimensional graph that describes how the modes of spectral densities changes across time.

3 Prediction and Conclusion

3.1 Time-Variant Spectrum Density

We use FFBS to draw full posterior of (ϕ_t, v_t) for each time point t, where ϕ_t is a p-by-1 vector. We plug in posterior mean of the samples to the spectrum density and obtain a full posterior of spectrum density. All spectrum density in the plots below have been transformed by logarithm function to be more visible.

Figure 1 shows the logarithm of spectrum density on the time line for TVAR(9) model, and Figure 2 are spectrum density with 1000 MC posterior samples of ϕ at t=50,150,200,300,500 and 900, respectively.

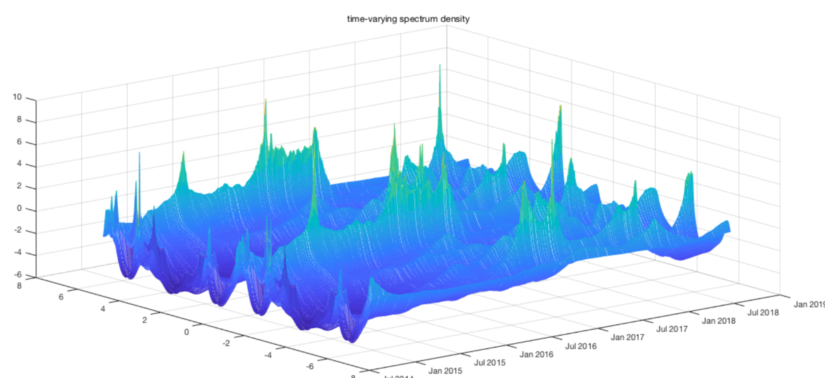


Figure 1: Time-varying spectrum density

In Figure 1, although at some time point the spectrum density becomes rough, the 3d plot show a relatively consistent ridge at the center of frequency

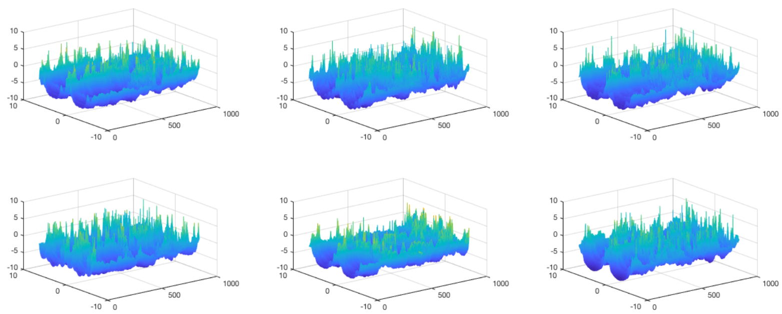


Figure 2: Synthetic future

domain (close to $\omega = 0$). This indicates a strong pattern of periodicity occur at that frequency. Figure 2 shows that the MC sampling looks consistent and stationary.

3.2 Distribution of Peaks

Given 1000 posterior samples of spectrum density at each time point, we compute the distribution of peaks (mode) at selected time point.

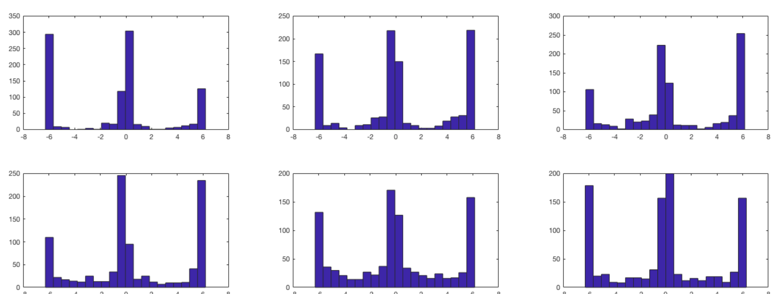


Figure 3: Posterior distribution of spectrum mode at t=50,150,200,300,500,and 900 from left to right, top to bottom

Below is part of table that contains the frequency of most frequent values of all samples and all time point t.

-0.5674	45.0000	4.4687
-0.5253	54.0000	5.3625

Based on the graphs above, it is quite clear that the most of the peaks fall between 0 and -1. The tabulate function tells us that mode based on posterior inference should be somewhere close to ω = -0.54, which then should be transformed into $2\pi/0.54 = 11$ days approximately.

After the full trajectory of $p(\theta_t|D_T)$ is obtained, the we can sample from the predictive distribution. One sample looks like following:

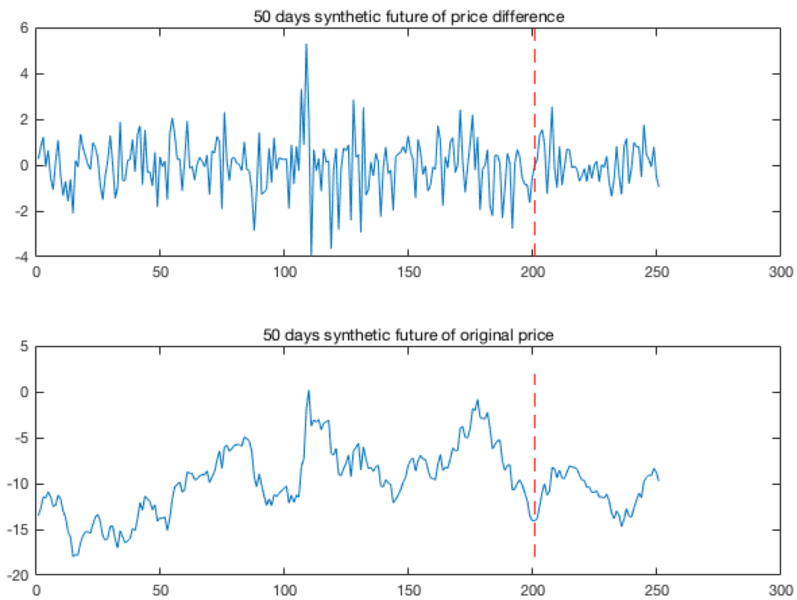


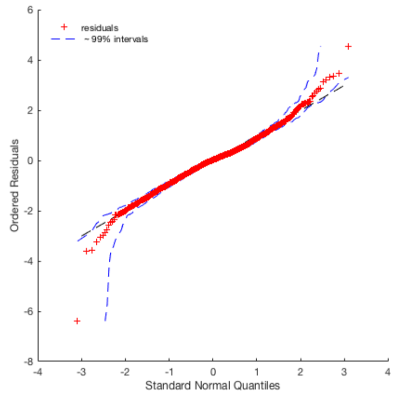
Figure 4: Synthetic future

The synthetic futures looks fairly similar to the original data.

4 Diagnostic and Discussion

The qq-plot for residual shows that he majority of the points are on straight line, while the end of the distribution indicates a fatter tail than normal distribution.

One of the main problem remain in this project is determine the stationarity of the data. Although the posterior draws indicates that the data is stationary, there is still a small percentage of posterior samples of ϕ fall outside of the range of -1 to 1. The trajectory of posterior median ϕ_t shown in 3.3 are not too far from zero.
If the project is to be continued for another semester, it might be interesting to look into those unusual behaviors. Another possible direction is to view model order p as variable, and make inference on its uncertainty[2].



5 Reference

[1] Prado, Raquel, et al. Time Series: Modeling, Computation, and Inference. Chapman & HallCRC, 2018.

[2] Huerta, Gabriel, and Mike West. "Bayesian Inference on Periodicities and Component Spectral Structure in Time Series." Journal of Time Series Analysis, vol. 20, no. 4, 1999, pp. 401–416., doi:10.1111/1467-9892.00145.