FINC 520: Time Series
Problem Set #4

Qiushi Huang
Marius Ring

May 15, 2015

1 Question 1

1.1 part (a)

We can rewrite the GARCH(1,1) as

$$\epsilon_t^2 = \omega + (\alpha + \beta)\epsilon_{t-1}^2 - \beta v_{t-1} + v_t \tag{1}$$

$$v_t = \epsilon_t^2 - \sigma_t^2 = (z_t^2 - 1)\sigma_2^2 \tag{2}$$

$$z_t \sim \text{I.I.D. N}(0, 1) \tag{3}$$

We find that

$$\mu \equiv \mathbb{E}\epsilon_t^2 = \frac{\omega}{1 - \alpha - \beta} \tag{4}$$

Thus we may write

$$\epsilon_t^2 - \mu = (\alpha + \beta)(\epsilon_{t-1}^2 - \mu) - \beta v_{t-1} + v_t \tag{5}$$

Then we use the Yule-Walker (not White-Walker as in Game of Thrones!!) method to find the autocorrelations for ϵ_t^2, which will be the same as those of y_t^2 under our assumption.
The first autocorrelation:

$$\rho_1 = \frac{\alpha(1 - \alpha\beta - \beta^2)}{1 - 2\alpha\beta - \beta^2} \tag{6}$$

Which provides us with all the subsequent ones due to p=q=1 in the GARCH(p,q):

$$\rho_k = (\alpha + \beta)^{k-1}\rho_1 \tag{7}$$

1.2 part(b)

The heteroskedasticity consistent estimator for the asymptotic covariance matrix is

$$\hat{V} = \hat{Q}^{-1}\Omega\hat{Q}^{-1} \tag{8}$$

where

$$\hat{\Omega} = \frac{1}{T}\sum \hat{\epsilon}_t^2 y_{t-1}^2 \tag{9}$$

$$\hat{Q} = \frac{1}{T}\sum y_{t-1}^2 \tag{10}$$

We will have that

$$\sqrt{T}(\hat{\phi} - \phi) \to N(0, Q^{-1}\Omega Q^{-1}) \text{ in distribution.} \tag{11}$$

Where we can, asymptotically, replace $Q^{-1}\Omega Q^{-1}$ with $\hat{V}$.

$$\Omega = \text{plim}\frac{1}{T}\sum \hat{\epsilon}_t^2 y_{t-1}^2 = \mathbb{E}\epsilon_t^2\epsilon_{t-1}^2 = \rho_1\gamma_0 + \mu^2 \tag{12}$$

$$Q = \text{plim}\frac{1}{T}\sum y_{t-1}^2 = \mathbb{E}\epsilon_{t-1}^2 = \mu \tag{13}$$

$$V = \frac{\rho_1\gamma_0 + \mu^2}{\mu^2} \tag{14}$$

where

$$\gamma_0 = 3\omega^2(1+\alpha+\beta)[(1-\alpha-\beta)(1-\beta^2-2\alpha\beta-3\alpha^2)]^{-1} \tag{15}$$

$$\mu = \frac{\omega}{1-\alpha-\beta} \tag{16}$$

$$\rho_1 = \frac{\alpha(1-\alpha\beta-\beta^2)}{1-2\alpha\beta-\beta^2} \tag{17}$$

So in the case where $\alpha = \beta = 0$ we have that $\rho_1 \to 0$ and $V \to 1$.

Otherwise, it follows from the terms, and our assumptions on ω, α, and β that $\rho_1 > 0$, and thus

$$V > 1 \tag{18}$$

1.3 part (c)

Parameters	Rejections
A	0.047
B	0.06
C	0.088
D	0.158

$$\hat{\phi}_T = \left(\sum_{t=2}^{T} y_{t-1}^2\right)^{-1} \sum_{t=2}^{T} y_t y_{t-1} \tag{19}$$

The asymptotic variance of $\sqrt{T}\hat{\phi}_T$ will just be 1 under our H_0, and the assumption that error terms are IID Normal. We see this in part (d): just let $\rho_1 = 0$.

$$\tag{20}$$

In A and B $V = 1$ and the IID Normal assumptions on the error term hold, so we would expect to see that that we reject 5% of the time, which we also do.

IN C and D $V > 1$, thus using the IID Normal assumption on the error term, we would be using a variance that is too low, and thus use a test statistic that is too high, and thus reject too often. i.e. **heteroskedasticity that is unaccounted for increases the size of the test beyond our confidence level.**

1.4 part (d) information criterions

The 'approximate' log likelihood function for the AR(p) estimation, assuming that the error terms are IID, is this:

$$logL = -\frac{T-p-1}{2}\log(2\pi) - \frac{T-p-1}{2}\log(\sigma^2) - \frac{1}{2}\sum_{t=p+1}^{T}\frac{\epsilon_t^2}{\sigma^2} \tag{21}$$

$$\epsilon_t = y_t - \sum_{j=1}^{p} \phi_j y_{t-j} \tag{22}$$

$$\hat{\sigma}_p^2 = \frac{1}{T-p-1}\sum_{t=p+1}^{T}\hat{\epsilon}_t^2 = \frac{1}{T-p-1}\sum_{t=p+1}^{T}\left(y_t - \sum_{j=1}^{p}\hat{\phi}_j y_{t-j}\right)^2 \tag{23}$$

Thus for our Information Criterion comparisons we use:

$$\log \hat{L}_p = -\frac{T}{2}\log(\hat{\sigma}_p^2) \tag{24}$$

Since the first terms with p in the logL above artificially increase the logL due to the nature of the conditional/approximate MLE. Also the third term depends only on T and p when we subsitute in for $\hat{\sigma}^2$

The ICs are given by:

$$BIC = -2\log \hat{L}_p + p\log(T) = T\log \hat{\sigma}_p^2 + p\ log(T) \tag{25}$$

$$AIC = -2\log \hat{L}_p + 2p = T\log \hat{\sigma}_p^2 + 2p \tag{26}$$

and $\hat{\phi}_T$ is our usual OLS estimate, regressing y_t on $y_{t-1}, ...y_{t-p}$, with no intercept term.

AIC results

Parameters	p0	p1	p2	p3	p4	p5
A	0.612	0.127	0.102	0.068	0.051	0.04
C	0.885	0.066	0.022	0.015	0.009	0.003
E	0	0.653	0.159	0.08	0.061	0.047
F	0	0.883	0.074	0.024	0.013	0.006
G	0	0.624	0.152	0.103	0.069	0.052
H	0	0.691	0.124	0.067	0.054	0.064

BIC results

Parameters	p0	p1	p2	p3	p4	p5
A	0.931	0.055	0.01	0	0.001	0.003
C	0.978	0.017	0.004	0.001	0	0
E	0	0.921	0.062	0.015	0.002	0
F	0	0.968	0.024	0.005	0.003	0
G	0	0.975	0.025	0	0	0
H	0	0.958	0.026	0.011	0.003	0.002

From theory we would expect that both AIC and BIC would in the limit never pick a p that was too small, and in the limit BIC will also not pick a p that is too big. We also expected AIC to be 'less conservative', i.e., choosing a higher number of lags, and possibly not be consistent towards the true $p = 0$.

The simulation results are well aligned with theory. We see that AIC more often than BIC picks too large p. We also see as $T \to 1000$ that BIC is a lot more consistent towards the true $p = 1$ than AIC for the setups G and H.

2 Question 2

2.1 (a) time series model, covariance stationarity, invertibility

$$y_t^h = \gamma_h(L)\epsilon_t \tag{27}$$

$$\gamma_h(L) = \sum_{j=0}^{h-1} L^j \tag{28}$$

So clearly y_t^h is an MA(h-1) process. For h=1 it is just white noise. For h=0 it is not well defined.

Any MA process is covariance stationary.However, it is **not** invertible, since **all** the roots of $\gamma_h(Z) = 0$, for any $h \geq 2$, lie **on** the unit circle.

2.2 (b) Gain and The Phase

The Gain, $G(w)$, and the Phase, $\theta(w)$, are defined as parameters in the exponential form of $\gamma_h(e^{iw})$:

$$\gamma_h(e^{iw}) = G(w)e^{i\theta(w)} \tag{29}$$

We can write

$$\gamma_h(e^{iw}) = 1 + \sum_{j=1}^{h-1}(e^{iw})^j \tag{30}$$

We know that the gain will be

$$G_h(w) = \mid \gamma_h(e^{iw}) \mid = \left| 1 + \sum_{j=1}^{h-1} e^{ijw} \right| \tag{31}$$

We could write this out in terms of cosines and sines, but it's not very useful, especially not for deriving the phase:

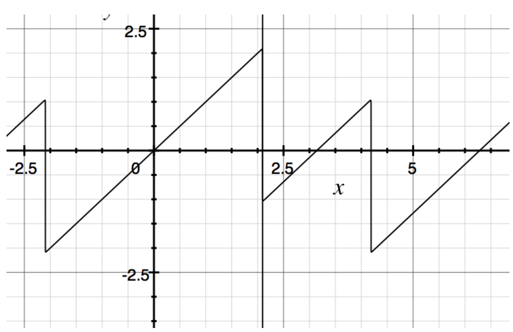
$$\theta_h(w) = \frac{1}{-i}\log\left(\frac{\gamma_h(e^{iw})}{G_h(w)}\right) \tag{32}$$

or more precisely

$$
\begin{aligned}
\theta_h(w) &= \frac{1}{-i}\log\left(\frac{1+\sum_{j=1}^{h-1}e^{ijw}}{\left|1+\sum_{j=1}^{h-1}e^{ijw}\right|}\right) && (33)\\
&= \frac{1}{2i}\log\left(\frac{(1+\sum_{j=1}^{h-1}e^{ijw})^2}{\left|1+\sum_{j=1}^{h-1}e^{ijw}\right|^2}\right) && (34)\\
&= \frac{1}{2i}\log\left(\frac{(1+\sum_{j=1}^{h-1}e^{ijw})^2}{\left(1+\sum_{j=1}^{h-1}e^{ijw}\right)\left(1+\sum_{j=1}^{h-1}e^{-ijw}\right)}\right) && (35)\\
&= \frac{1}{2i}\log\left(\frac{e^{i(h-1)w}}{e^{i(h-1)w}}\frac{1+\sum_{j=1}^{h-1}e^{ijw}}{\left(1+\sum_{j=1}^{h-1}e^{-ijw}\right)}\right) && (36)\\
&= \frac{1}{2i}\log\left(e^{i(h-1)w}\frac{1+\sum_{j=1}^{h-1}e^{ijw}}{\left(1+\sum_{j=1}^{h-1}e^{ijw}\right)}\right) && (37)\\
&= \frac{1}{2i}\log\left(e^{i(h-1)w}\right) && (38)\\
&= \frac{i(h-1)w}{2i} = \frac{h-1}{2}w && (39)\\
& && (40)\\
& && (41)
\end{aligned}
$$

which will be a real-valued function, and well defined for $w \in [0, \frac{2\pi}{h-1}]$.

The general phase looks like this, where it "resets" at every $\frac{2\pi}{h-1}$th frequency. The below plot is for $h = 3$.

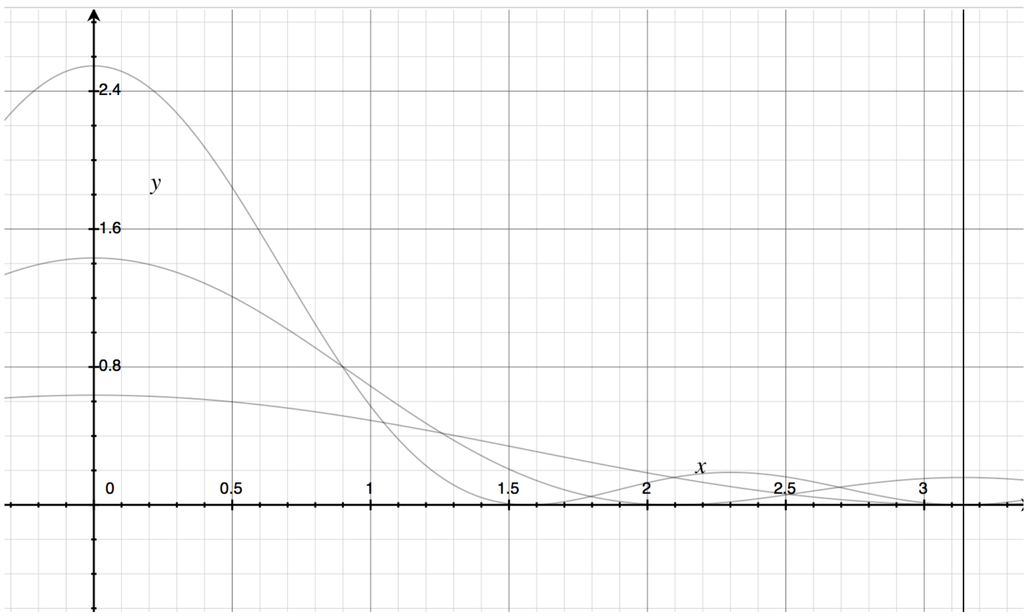


2.3 (c) Spectrum of y_t, h

We use the fact that this is just a filter of a white noise process with spectrum $\frac{\sigma_\epsilon^2}{2\pi}$, thus

$$f_{y_h}(w) = \left|1 + \sum_{j=1}^{h-1} e^{-ijw}\right|^2 \frac{\sigma_\epsilon^2}{2\pi} \tag{42}$$

Again, complicating this by adding sines and cosines does not add any intuition. Instead we plot it for $h = 2, 3, 4$, where the flattest one is h=2, and the waviest is h=4, and we set $\sigma_\epsilon^2 = 1$



3 Question 3

3.1 part (a) graph spectra

$$\begin{aligned} y_t &= c_t + u_t &(43)\\ c_t &= 0.9c_{t-1} + \epsilon_t = ... = \sum_{j=0}^{\infty} 0.9^j \epsilon_{t-j} &(44)\\ u_t &= \eta_t - 0.8\eta_{t-1} &(45)\end{aligned}$$

since c_t and u_t are MA processes with independent errors (of eachother), the covariances of y_t will be the sum of the respective covariances for c_t and u_t, and thus the spectrum for y is the sum of the spectra of c and u.

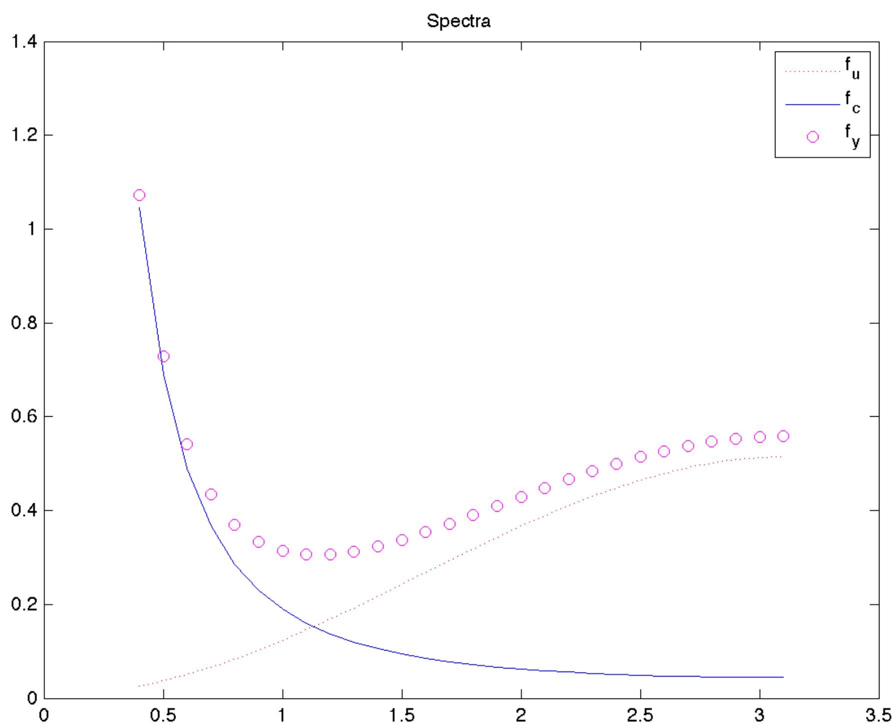
Since both c and u are filters of unit variance white noise processe, with spectra $\frac{1}{2\pi}$, we can write the spectra of c and u as functions of the spectrum of a unit variance white noise process.

Using the general formula

$$f(w) = \left|\Psi(e^{-iw})\right|^2 f_{WN}(w) \tag{46}$$

We obtain

$$\begin{aligned} f_u(w) &= \left|1 - 0.8e^{-iw}\right|^2 \frac{1}{2\pi} \quad (47)\\ f_c(w) &= \left|\frac{1}{1 - 0.9e^{-iw}}\right|^2 \frac{1}{2\pi} \quad (48) \end{aligned}$$



3.2 b

Based on visual inspection, it seems that for $w < 0.5$ the variance almost exclusively comes from c_t, while for $w > 2$, the variance mostly comes from u_t. Thus an initial guess would be to create a filter $\delta(L)$ such that the spectrum of the filtered sequence looks like this, with $\underline{w} = 1$.

$$f_{\hat{c}}(w) = |\delta(e^{-iw})|^2 f_y(w) = \mathbb{1}[w < \underline{w}] f_y(w) \tag{49}$$

The first equality comes from the formula for how the spectrum of a filtered series depends on the raw series. The second equality is what we wish to achieve. Thus we must set:

$$\delta(e^{iw}) = \mathbb{1}[w < \underline{w}] \tag{50}$$

Now we need to find

$$\delta(L) = \sum_{h=-\infty}^{\infty} \delta_h L^h \tag{51}$$

That satisfies

$$\delta(L) = \sum_{h=-\infty}^{\infty} \delta_h \delta(e^{iw})^h \tag{52}$$

For this we use the Fourier Inversion Formula from class:

$$k(h) : \sum_h \mid k(h) \mid < \infty \Rightarrow \left[k(h) = \int_{-\pi}^{\pi} e^{ihw} r(w) dw \iff r(w) = \frac{1}{2\pi} \sum_h e^{-ihw} k(h) \right] \tag{53}$$

$$r(w) \equiv \delta(e^{-iw}), k(h) \equiv 2\pi\delta_h \Rightarrow \delta_h = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ihw} \delta(e^{iw}) dw \tag{54}$$

$$\Rightarrow \delta(L) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \left(\int_{-\pi}^{\pi} e^{ihw} \delta(e^{iw}) dw \right) L^h \tag{55}$$

$$\Rightarrow \delta(L) = \frac{1}{\pi} \sum_{h=-\infty}^{\infty} \frac{\sin(h\underline{w})}{h\underline{w}} L^h \tag{56}$$

But we need to cut it off at some point (L=50). So we use:

$$\Rightarrow \delta(L) = \sum_{h=-\infty}^{\infty} \mathbb{1}[|h| \leq 50] \frac{\sin(h\underline{w})}{\pi h\underline{w}} L^h \tag{57}$$

$$\Rightarrow \delta(L) = \sum_{h=-50}^{50} \frac{\sin(h\underline{w})}{\pi h\underline{w}} L^h \tag{58}$$

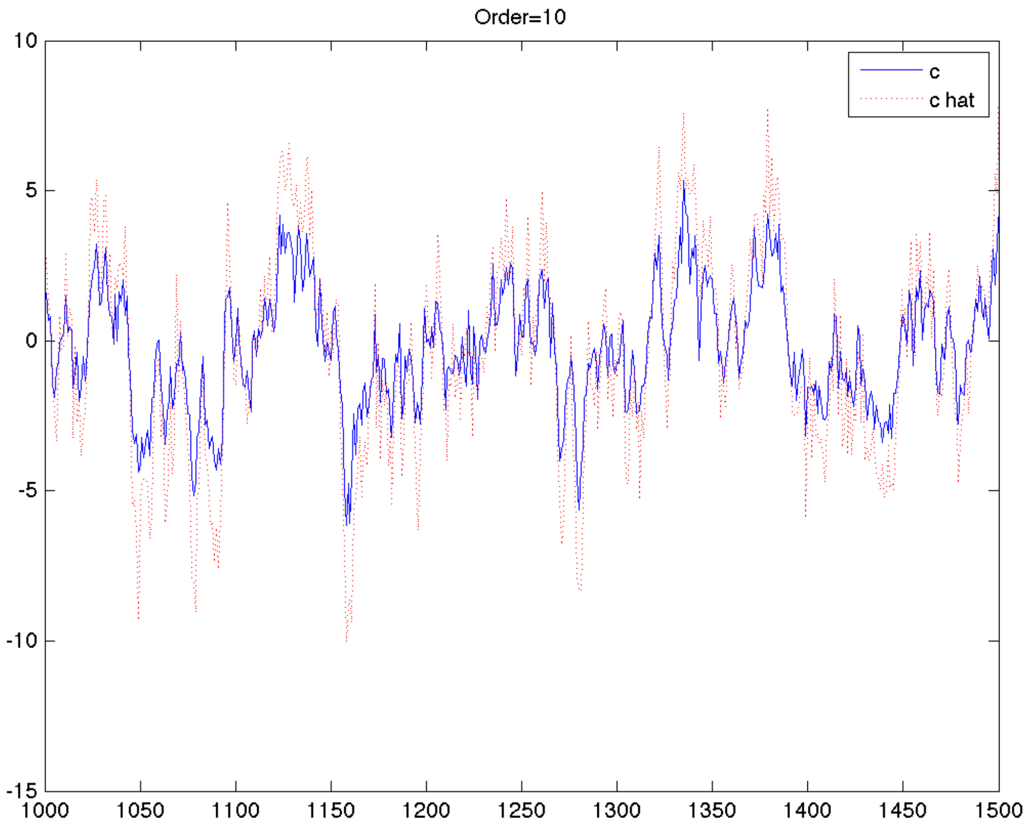
And since we choose $\underline{w} = 1$:

$$\Rightarrow \delta(L) = \sum_{h=-50}^{50} \frac{\sin(h)}{\pi h} L^h \tag{59}$$

For h=0, δ_0=1.

3.3 (c)

Correlation: 0.9559



3.4 (d)

Correlation: 0.9561 The correlation goes up. If we decrease the order further, the correlation goes up further. On one hand we are getting a poorer approximation to our filter, which we don't actually know is optimal, but on the other hand we are getting more data to use.