Exam for 2010-2011: SOLUTIONS

Exam Question 1:
Question 1.1:

1. Clearly the $|x_t|$ series show autocorrelation and hence "ARCH" effects. Note that one has used $|x|$ rather than x_t^2 to "save moments" and hence the used std. deviations are more plaussible.

2. $\hat{\rho} \pm 2std$ contains the true value $\rho_0 = 0.4$ - Hence if standard (asymptotic) inference holds this is fine.

3. Normality: accepted and ARCH test: accepted - hence well-specified model (as expected of course as the data are simulated after all).

4. One can rewrite the model as

$$\sigma_t^2 = c\left(x_{t-1}^2\right)^{\rho}, \quad \text{where } c = \exp(\mu),$$

and hence it is quite different from a classic linear ARCH model. Rewriting gives,

$$\log\sigma_t^2 = \mu + \rho\log\sigma_{t-1}^2 + \rho\log z_{t-1}^2,$$

and it resembles the classic log SV model of the form,

$$\log\sigma_t^2 = \mu + \rho\log\sigma_{t-1}^2 + \eta_t.$$

A key difference is that η_t in the SV is (typically) assumed to be $\text{iidN}\left(0, \sigma_\eta^2\right)$ (and independent of z_t), while here $\log z_{t-1}^2$ is not Gaussian (far from in fact as known from the discussions on log SV models) and $\left(\log z_{t-1}^2\right)_{t=1,2,\ldots,T}$ is not independent (as a sequence) of (z_t).

Question 1.2:
Set $y_t = \log x_t^2$ then

$$\begin{aligned} y_t &= \log\sigma_t^2 + \log z_t^2 \\ &= \mu + \rho y_{t-1} + \log z_t^2. \end{aligned}$$

This is indeed a Markov chain on the classic AR(1) form with innovations $\eta_t = \log z_t^2$. Thus one can use directly results from the AR(1) process, from where standard drift criterion arguments with for example $\delta(y) = 1 + y^2$ give the condition, $\rho^2 < 1$ for weak mixing of y_t and $Ey_t^2 < \infty$.

Note that one may want to define $\tilde{\eta}_t = \eta_t - E\log z_t^2 = \eta_t - c_1$ and $\tilde{\mu} = \mu + c_1$ such that,

$$y_t = \tilde{\mu} + \rho y_{t-1} + \tilde{\eta}_t,$$

with $\tilde{\eta}_t$ mean-zero iid.

The details of the drift criterion application are:

$$\begin{aligned} & E\left(\delta\left(y_t\right)|y_{t-1} = y\right) \\ & = 1 + \tilde{\mu}^2 + V\left(\log z_t^2\right) + \rho^2 y^2 + 2\tilde{\mu}\rho y \\ & = a + by + \rho y^2. \end{aligned}$$

Hence, using $\rho^2 < 1$,

$$\frac{a + by + \rho y^2}{1 + y^2} \to 0$$

as $y^2 \to \infty$, and as $a + by + \rho y^2 \leq d$ for $y^2 \leq M$ for any $M > 0$ the result follows.

Question 1.3: Procedure to obtain the MLE $\hat{\rho}$: Numerical optimization using e.g. the Newton-Raphson algorithm (BFGS in *ox*) maximzing the log-likelihood function $\ell_T(\rho)$. Theorem III.1 in notes state under which regularity conditions a local optimum can be found - the reg. conditions are similar to the ones for asymptotic normality of the MLE. These may be briefly discussed and/or mentioned. In fact, the result in the next Question 1.4 is one of such reg. conditions.

Outline 1.4: Simple calculus gives,

$$\partial\ell_T(\rho)/\partial\rho = -\frac{1}{2}\sum\left(\log x_{t-1}^2 - \frac{x_t^2\log x_{t-1}^2}{\sigma_t^2}\right) = \frac{1}{2}\sum\left(\frac{x_t^2}{\sigma_t^2} - 1\right)\log x_{t-1}^2$$

With $v_t = \left(\frac{x_t^2}{\sigma_t^2} - 1\right)\log x_{t-1}^2$. Use $\left(\frac{x_t^2}{\sigma_t^2} - 1\right) = z_t^2 - 1$, and rules for conditional expectations give:

$$E\left(v_t \mid x_{t-1}\right) = \left(\log x_{t-1}^2\right) E\left(\left(\frac{x_t^2}{\sigma_t^2} - 1\right)|x_{t-1}\right) = \left(\log x_{t-1}^2\right) E\left(z_t^2 - 1\right) = 0,$$

such that v_t is a martingale difference sequence.

By the weakly mixing established in Question 1.2 with $\delta(y) = 1 + y^2$, we can apply the CLT (Theorem II.1) from the notes since $E\left(y_t^2\right) = E\left(\log x_{t-1}^2\right)^2 < \infty$. Observe that,

$$E\left(v_t^2\right) = E\left(E\left(v_t^2|x_{t-1}\right)\right) = E\left(\left(\log x_{t-1}^2\right)^2 E\left(z_t^2 - 1\right)^2\right) = 2E\left(\log x_{t-1}^2\right)^2$$

since $E\left(z_t^2-1\right)^2 = 2$. We conclude that $\frac{1}{\sqrt{T}}\,\partial\ell_T\left(\rho\right)/\partial\rho\big|_{\rho=\rho_0} \xrightarrow{D} N\left(0,\Sigma\right)$ with

$$\Sigma=\frac{2}{4}E\left(\log x_{t-1}^2\right)^2=\frac{1}{2}E\left(\log x_{t-1}^2\right)^2.$$

One can actually compute $E\left(\log x_{t-1}^2\right)^2$: From Question 1.2, with $y_t=\log x_t^2$,

$$y_t=\tilde{\mu}+\rho y_{t-1}+\tilde{\eta}_t.$$

Hence, if $\rho^2<1$, $V\left(y_t\right)=\frac{\pi^2}{2}\left(1-\rho^2\right)^{-1}$ and $\left(Ey_t\right)^2=\tilde{\mu}^2\left(1-\rho\right)^{-2}$. Thus,

$$\begin{aligned}E\left(\log x_{t-1}^2\right)^2&=Ey_t^2=V\left(y_t\right)+\left(Ey_t\right)^2\\&=\frac{\pi^2}{2}\left(1-\rho^2\right)^{-1}+\left(\mu+1.2\right)^2\left(1-\rho\right)^{-2}.\end{aligned}$$

Question 1.5:

From the definition of x_t (see also Question 1.2),

$$\log x_t^2=\mu+\rho\log x_{t-1}^2+\log z_t^2,$$

and as $\log z_t^2$ is iid with $E\left(\log z_t^2\right)^2$ finite, the OLS estimator $\hat{\phi}_{\text{ols}}$ is consistent. One may (time consuming though) also argue by using the LLN as we have that $\log x_t^2=y_t$ is weakly mixing and have second order moments from Question 1.2.

As the innovations $\log z_t^2$ have mean $\neq 0$, then the OLS of δ will converge to $\mu-E\log z_t^2$. That is, the $\hat{\delta}_{\text{ols}}$ is not consistent. This is also clear from rewriting as in Question 1.2,

$$y_t=\tilde{\mu}+\rho y_{t-1}+\tilde{\eta}_t,$$

such that $\hat{\delta}_{\text{ols}}\rightarrow\tilde{\mu}=\mu-E\log z_t^2$.

Exam Question 2:

Question 2.1: We would **not** expect IGARCH model to fit. We can see from the table that the GARCH model seems to remove ARCH effects but that normality is strongly rejected. One way we can accomodate for the misspecification is the much observed IGARCH: $\hat{\alpha} + \hat{\beta} = 1$ as seen in many financial series. It does not mean that the IGARCH is a good model for these data by any standards. In particular, as the plot is a log-plot, the scale of x_t is of such magnitude that one would believe an explosive variance may be needed ($\alpha + \beta > 1$). It is not likely that standard inference applies and hence that the reported p-values are meaningful. More diagnostics are needed.

Question 2.2: One should see that V_t is a VAR(1) of dimension 2, and hence *conclude from the notes directly* that as V_t is a markov chain, and the condition for weakly mixing and $E||V_t||^2 < \infty$ is,

$$\rho(\Phi) < 1.$$

Equivalently this can be stated as the maximal eigenvalue of Φ is smaller than one in absolute value.

Question 2.3:
Simple calculations give,

$$\begin{aligned} \log|x_t| &= (1,1)\, V_t + \log|z_t| \\ V_t &= \mu + \Phi V_{t-1} + \eta_t \end{aligned}$$

Hence setting $E\log|z_t| = \delta$, this can be written as,

$$\begin{aligned} \log|x_t| &= (1,1)\, V_t + \delta + e_t \\ V_t &= \mu + \Phi V_{t-1} + \eta_t, \end{aligned}$$

where $e_t = (\log|z_t| - \delta)$ are iid mean zero (with variance $\pi^2/8$). Thus the system is directly on State Space Form, and parameters can be estimated by the *Kalman Filter* (ignoring that e_t are not Gaussian - so this is QMLE).

One may want to write the system as in the form suitable for the *ssf-package in ox (*or any other form of course as there is no unique state space form):

$$\begin{pmatrix} V_{t+1} \\ \log|x_t| \end{pmatrix} = \begin{pmatrix} \Phi \\ (1,1) \end{pmatrix} V_t + \begin{pmatrix} \mu \\ \delta \end{pmatrix} + \begin{pmatrix} \eta_{t+1} \\ e_t \end{pmatrix}.$$

Either way this is not MLE as $\log|z_t|$ is not Gaussian. Hence this is QMLE and does not have optimal properties, why estimation methods based on simulations have been developed in the SV literature.

Question 2.4:
From the parameters shown,

$$\log \sigma_{1t} = \log \sigma_{1t-1} + \eta_{1t} = \sum_{i=1}^{t} \eta_{1i} + \log \sigma_{10}$$

and,

$$\begin{aligned} \log \sigma_{2t} &= \varrho \log \sigma_{2t-1} + \eta_{2t} \\ &= \sum_{i=0}^{\infty} \varrho^i \eta_{2t-i} \end{aligned}$$

since $\varrho^2 = 0.25^2 < 1$. Hence,

$$\begin{aligned} \log |x_t| &= \log \sigma_{1t} + \log \sigma_{2t} + \log |z_t| \\ &= \sum_{i=1}^{t} \eta_{1i} + \log \sigma_{10} + \sum_{i=0}^{\infty} \varrho^i \eta_{2t-i} + \log |z_t| \end{aligned}$$

as desired. Thus there is a stochastic trend (Random walk) creating persistence, and a stationary process driving $\log |x_t|$ - this is exactly what one would expect from the figure.