Financial Econometrics A
January 14, 2019
SOLUTION KEY

Please note there is a total of **8** questions that you should provide answers to. That is, **3** questions under *Question A*, and **5** under *Question B*.

Question A:

Consider the model for $x_t \in \mathbb{R}$ given by

$$x_t = \sigma_t z_t, \quad \sigma_t^2 = \omega + \alpha x_{t-1}^2$$

with (z_t) an *i.i.d.* process, and z_t is scaled t-distributed and given by

$$z_t = \sqrt{\frac{v-2}{v}} \tau_t,$$

where τ_t is t-distributed with $v = 5$ degrees of freedom. In particular, we have

$$E\tau_t = 0,\ V(\tau_t) = E\left(\tau_t^2\right) = \frac{v}{v-2} = \frac{5}{3} \text{ and } E\left(\tau_t^4\right) = \frac{3v^2}{(v-4)(v-2)} = 25.$$

Moreover, the model parameters satify that $\omega > 0$ and $\alpha \geq 0$.
Question A.1: Show that x_t is weakly mixing with $Ex_t^4 < \infty$ if $\alpha < 1/3$.
Solution: *With drift function $\delta_4(x) = 1 + x^4$, we find*

$$\begin{aligned} E\left(\delta_4(x_t) | x_{t-1} = x\right) &= 1 + \left(\omega + \alpha x^2\right)^2 \left(\frac{v-2}{v}\right)^2 \frac{3v^2}{(v-4)(v-2)} \\ &= 1 + \left(\omega + \alpha x^2\right)^2 \frac{3(v-2)}{(v-4)} \\ &= 1 + \left(\omega^2 + \alpha^2 x^2 + 2\alpha\omega x\right) 9 \end{aligned}$$

and we find the condition $\alpha^2 < 1/9$.

Question A.2: Consider the extended ARCH(2) model,

$$x_t = \sigma_t z_t, \quad \sigma_t^2 = \omega + \alpha x_{t-1}^2 + \beta x_{t-2}^2,$$

with $\omega > 0$ and $\alpha, \beta \geq 0$, and z_t still *i.i.d.* scaled t-distributed with $v = 5$. The Gaussian QMLE $\hat{\theta} = \left(\hat{\omega}, \hat{\alpha}, \hat{\beta}\right)'$ is obtained by maximizing $L(\theta)$ given by

$$L(\theta) = -\frac{1}{T}\sum_{t=1}^{T}\left(\log \sigma_t^2(\theta) + \frac{x_t^2}{\sigma_t^2(\theta)}\right),$$
$$\sigma_t^2(\theta) = \omega + \alpha x_{t-1}^2 + \beta x_{t-2}^2.$$

It follows that,

$$S_\beta(\theta) = \partial L(\theta)/\partial\beta = \frac{1}{T}\sum_{t=1}^{T}\left(\frac{x_t^2}{\sigma_t^2(\theta)} - 1\right)\frac{x_{t-2}^2}{\sigma_t^2(\theta)}.$$

With $\theta_0 = (\omega_0, \alpha_0, 0)$ (that is, the true value of β is zero), show that if $0 < \alpha_0 < 1/3$ then

$$\sqrt{T}S_\beta(\theta_0) \xrightarrow{d} N(0, \Omega)$$

with

$$\Omega = 8E\left(\frac{x_{t-2}^2}{\omega_0 + \alpha_0 x_{t-1}^2}\right)^2.$$

Solution:

$$\sqrt{T}S_\beta(\theta_0) = \frac{1}{\sqrt{T}}\sum_{t=1}^{T} m_t, \quad m_t = \left(z_t^2 - 1\right)\frac{x_{t-2}^2}{\sigma_t^2(\theta)}$$

Noting that $E[m_t|x_{t-1}, x_{t-2}] = 0$, *by the CLT for weakly mixing processes the result holds provided*

$$Em_t^2 = E\left(z_t^2 - 1\right)^2 E\left(\frac{x_{t-2}^2}{\omega_0 + \alpha_0 x_{t-1}^2}\right)^2 < \infty.$$

Now $E\left(z_t^2 - 1\right)^2 = 9 - 1 = 8 < \infty$ *(as* $v = 5 > 4$*) and next*

$$E\left(\frac{x_{t-2}^2}{\omega_0 + \alpha_0 x_{t-1}^2}\right)^2 < \frac{1}{\omega_0^2}E\left(x_{t-2}^4\right)$$

such that a sufficient condition is $Ex_t^4 < \infty$ *which is* $\alpha_0 < 1/3$.

Question A.3:

In the ARCH(2) model we wish to test the hypothesis $\beta = 0$ by the Quasi likelihood ratio statistic

$$QLR = \left(L(\hat{\theta}) - L\left(\tilde{\theta}\right) \right),$$

where $\tilde{\theta}$ maximizes $L(\theta)$ with $\beta = 0$. The asymptotic distribution of the QLR statistic is not χ_1^2. Explain why it is not χ^2. And discuss which distribution it has.

Solution: *As $\beta_0 = 0$ is on the boundary of the parameter space the limiting distribution is not χ^2. If $\alpha_0 > 0$, the limiting distribution is, up to a scaling constant, a "$\frac{1}{2}\chi^2$" ($= (\max(0, \eta))^2$ with $\eta = N(0,1)$). From A.2 we expect a condition for this to hold is that $Ex_t^4 < \infty$.*

Question B:

Consider the model for $x_t \in \mathbb{R}$ given by

$$x_t = b_{s_t} + a_{s_t} x_{t-1} + \varepsilon_t, \tag{B.1}$$

where the error term ε_t satisfies

$$\varepsilon_t \sim i.i.d.N(0, \sigma^2). \tag{B.2}$$

Moreover,

$$b_{s_t} = 1(s_t = 1)b_1 + 1(s_t = 2)b_2, \tag{B.3}$$

and

$$a_{s_t} = 1(s_t = 1)a_1 + 1(s_t = 2)a_2, \tag{B.4}$$

where s_t is a state variable that takes values in $\{1, 2\}$ according to the transition probabilities

$$P(s_t = j|s_{t-1} = i) = p_{ij} \in [0, 1], \tag{B.5}$$

and $1(s_t = i) = 1$ if $s_t = i$ and $1(s_t = i) = 0$ if $s_t \neq i$ for $i = 1, 2$. We assume throughout that the processes (ε_t) and (s_t) are independent. The model parameters satisfy $b_1, b_2, a_1, a_2 \in \mathbb{R}$ and $\sigma^2 > 0$.

Question B.1: When is the process (s_t) weakly mixing?

Solution: *It is well-known that s_t is weakly mixing if $p_{11}, p_{22} < 1$ (irreducibility) and $p_{11} + p_{22} > 0$ (aperiodicity).*

Question B.2: Let $f(x_t|x_{t-1}, s_t)$ denote the conditional density of x_t given (x_{t-1}, s_t). Provide an expression for $f(x_t|x_{t-1}, s_t)$.

Solution: *For $i = 1, 2$, we note that $(x_t|x_{t-1}, s_t = i) \sim N(b_i + a_i x_{t-1}, \sigma^2)$. Hence, for $i = 1, 2$,*

$$f(x_t|x_{t-1}, s_t = i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x_t - b_i - a_i x_{t-1})^2}{2\sigma^2}\right).$$

We conclude that

$$f(x_t|x_{t-1}, s_t) = \prod_{i=1}^{2} f(x_t|x_{t-1}, s_t = i)^{1(s_t = i)}.$$

Question B.3: In the following we assume that $p_{11} = 1 - p_{22} =: p \in (0,1)$ such that (s_t) is an *i.i.d.* process with $P(s_t = 1) = p$. This implies that $f(x_t|x_{t-1}) = f(x_t|x_{t-1}, s_t = 1)P(s_t = 1) + f(x_t|x_{t-1}, s_t = 2)P(s_t = 2) > 0$. Use this to show that x_t satisfies the drift criterion with drift function $\delta(x) = 1 + x^2$ when

$$a_1^2 p + a_2^2(1-p) < 1.$$

Solution: *We note that x_t is a Markov chain with positive and continuous transition density, $f(x_t|x_{t-1})$. Moreover,*

$$\begin{aligned}
&E[1 + x_t^2|x_{t-1} = x] \\
&= 1 + E[b_{s_t}^2 + a_{s_t}^2 x_{t-1}^2 + \varepsilon_t^2 + 2b_{s_t}a_{s_t}x_{t-1} + 2b_{s_t}\varepsilon_t + 2a_{s_t}\varepsilon_t x_{t-1}|x_{t-1} = x] \\
&= 1 + E[b_{s_t}^2 + a_{s_t}^2 x_{t-1}^2 + \varepsilon_t^2 + 2b_{s_t}a_{s_t}x_{t-1}|x_{t-1} = x].
\end{aligned}$$

Note that

$$\begin{aligned}
E\left[b_{s_t}^2|x_{t-1} = x\right] &= b_1^2 p + b_2^2(1-p), \\
E\left[a_{s_t}^2|x_{t-1} = x\right] &= a_1^2 p + a_2^2(1-p), \\
E\left[a_{s_t}b_{s_t}|x_{t-1} = x\right] &= a_1 b_1 p + a_2 b_2(1-p).
\end{aligned}$$

Hence, with $C \equiv 1 + \sigma^2 + b_1^2 p + b_2^2(1-p)$,

$$E\left[1 + x_t^2|x_{t-1} = x\right] = C + 2\left(a_1 b_1 p + a_2 b_2(1-p)\right)x + \left(a_1^2 p + a_2^2(1-p)\right)x^2.$$

Usual derivations yield that the drift criterion is satisfied, if $[a_1^2 p + a_2^2(1-p)] < 1$. Details should be provided.

Question B.4: Consider the *restricted* version of the model in (B.1)-(B.4), where $b_1 = b_2 = 0$. Maintaining the assumptions from Question B.3, and assuming that (s_t, x_{t-1}) is *observed*, we consider the log-likelihood function (up to a constant),

$$\begin{aligned}
L_T(\theta) = \sum_{t=1}^{T} \Bigg[& 1(s_t = 1)\left\{-\frac{1}{2}\log(\sigma^2) - \frac{(x_t - a_1 x_{t-1})^2}{2\sigma^2} + \log(p)\right\} \\
& + 1(s_t = 2)\left\{-\frac{1}{2}\log(\sigma^2) - \frac{(x_t - a_2 x_{t-1})^2}{2\sigma^2} + \log(1-p)\right\}\Bigg],
\end{aligned}$$

where $\theta = (p, a_1, a_2, \sigma^2)$. Show that the Maximum Likelihood Estimator for a_1 is

$$\hat{a}_1 = \frac{\sum_{t=1}^{T} 1(s_t = 1)x_t x_{t-1}}{\sum_{t=1}^{T} 1(s_t = 1)x_{t-1}^2}.$$

Assume that the joint process (s_t, x_{t-1}) is weakly mixing such that $E[x_{t-1}^2] < \infty$. Argue that $\hat{a}_1 \xrightarrow{p} a_1$ as $T \to \infty$.

Solution: *Solving* $\partial L_T(\theta) / \partial a_1 = 0$ *yields* $\hat{a}_1$. *Moreover,*

$$\begin{aligned}
\hat{a}_1 &= \frac{\sum_{t=1}^{T} 1(s_t = 1)[(1(s_t = 1)a_1 + 1(s_t = 2)a_2)x_{t-1} + \varepsilon_t]x_{t-1}}{\sum_{t=1}^{T} 1(s_t = 1)x_{t-1}^2} \\
&= a_1 + \frac{T^{-1}\sum_{t=1}^{T} 1(s_t = 1)\varepsilon_t x_{t-1}}{T^{-1}\sum_{t=1}^{T} 1(s_t = 1)x_{t-1}^2}.
\end{aligned}$$

By the LLN for weakly mixing process (using that $E[x_t^2] < \infty$*),*

$$T^{-1}\sum_{t=1}^{T} 1(s_t = 1)\varepsilon_t x_{t-1} \xrightarrow{P} E[1(s_t = 1)\varepsilon_t x_{t-1}] = 0$$

and

$$T^{-1}\sum_{t=1}^{T} 1(s_t = 1)x_{t-1}^2 \xrightarrow{P} E[1(s_t = 1)x_{t-1}^2] = pE[x_{t-1}^2] < \infty.$$

We conclude that $\hat{a}_1 \xrightarrow{P} a_1$.

Question B.5: Suppose that the process (s_t) is *unobserved*, but still satisfies the *i.i.d.* assumption, i.e. $p_{11} = (1 - p_{22}) = p \in (0, 1)$. Then the estimator $\hat{a}_1$ derived in Question B.4 is infeasible. Consider instead the function

$$L_T^\dagger(\theta) = E[L_T(\theta)|x_1, ...x_T].$$

It holds (still with $b_1 = b_2 = 0$) that

$$\begin{aligned}
L_T^\dagger(\theta) = \sum_{t=1}^{T} \Bigg[& P_t^\dagger(1) \left\{ -\frac{1}{2}\log(\sigma^2) - \frac{(x_t - a_1 x_{t-1})^2}{2\sigma^2} + \log(p) \right\} \\
& + (1 - P_t^\dagger(1)) \left\{ -\frac{1}{2}\log(\sigma^2) - \frac{(x_t - a_2 x_{t-1})^2}{2\sigma^2} + \log(1 - p) \right\} \Bigg].
\end{aligned}$$

where $P_t^\dagger(1) = P(s_t = 1|x_t)$.
Explain briefly the role of $L_T^\dagger(\theta)$ for the estimation of θ.

Solution: *The function* $L_T^\dagger(\theta)$ *is used as part of the EM Algorithm. Details should be provided.*