Solutions for Preparation Exam
Advanced Econometrics (4.1)

Master Econometrics and Operations Research
Faculty of Economics and Business Administration

Exam:	Advanced Econometrics (4.1)
Code:	E_EORM_AECTR
Coordinator:	dr. F. Blasques
Co-Reader:	Prof. dr. S. J. Koopman
Date:	–
Time:	–
Duration:	2 hours and 45 minutes
Calculator:	Not allowed
Graphical calculator:	Not allowed
Number of questions:	4
Type of questions:	Open
Answer in:	English
Credit score:	100 credits counts for a 10
Grades:	Made public within 10 working days
Inspection:	By appointment (send e-mail to f.blasques@vu.nl)
Number of pages:	5, including front page

- Read the entire exam carefully before you start answering the questions.
- Be clear and concise in your statements, but justify every step in your derivations.
- If you think that further information is needed to answer a question or that the question is ill-posed, then explain your reasoning.
- The questions should be handed back at the end of the exam. Do not take it home.

Good luck!

Question 1 [25 points]   Stochastic Properties of Nonlinear Dynamic Models

In economics and finance, time-series may sometimes exhibit time-varying conditional mean and volatility.

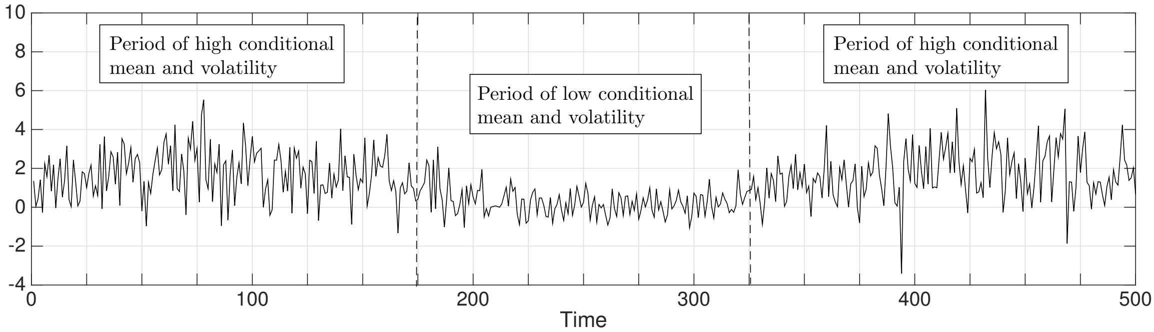

Figure 1: Time-series with time-varying conditional mean and volatility.

Let $\{x_t\}_{t\in\mathbb{Z}}$ be generated according to

$$x_t = \mu_t + \sigma_t \varepsilon_t \quad \text{for every} \quad t \in \mathbb{Z}\ ,$$

where $\{\varepsilon_t\}_{t\in\mathbb{Z}}$ is a sequence of Gaussian iid random variables $\{\varepsilon_t\}_{t\in\mathbb{Z}} \sim \text{NID}(0,1)$. Suppose that the time-varying conditional mean $\{\mu_t\}_{t\in\mathbb{Z}}$ satisfies

$$\mu_t = 0.2\big(x_{t-1} - \mu_{t-1}\big) + 0.7\mu_{t-1} \quad \text{for every} \quad t \in \mathbb{Z}.$$

Furthermore, let the time-varying volatility $\{\sigma_t\}_{t\in\mathbb{Z}}$ be determined by an exogenous sequence $\{z_t\}_{t\in\mathbb{Z}}$, according to

$$\sigma_t = \big(1 + \tanh(z_t)\big) \quad \text{for every} \quad t \in \mathbb{Z}.$$

Finally, let $\{z_t\}_{t\in\mathbb{Z}}$ be generated by the following random coefficient autoregressive model

$$z_{t+1} = \rho_t z_t + v_t \quad \text{for every} \quad t \in \mathbb{Z}\ ,$$

where $\{\rho_t\}_{t\in\mathbb{Z}}$ is a sequence of iid random variables with uniform distribution $\{\rho_t\}_{t\in\mathbb{Z}} \sim \text{UID}(0,\ 1.5)$ taking values in the interval $[0\ ,\ 1.5]$, and $\{v_t\}_{t\in\mathbb{Z}}$ is a sequence of Student-t iid random variables with two degrees of freedom $\{v_t\}_{t\in\mathbb{Z}} \sim TID(2)$.

Note: the acronym *iid* stands for *independent identically distributed.*
Note: the function $1 + \tanh(\cdot)$ is uniformly bounded between 0 and 2.
Note: the random variable v_t satisfies $\mathbb{E}|v_t|^n < \infty$ for $0 < n < 2$.

(a) **(10pts)** Can you show that $\{\sigma_t\}_{t\in\mathbb{Z}}$ is strictly stationary and ergodic?

Answer: Since the tanh function is continuous, it is also measurable with respect to the Borel sigma-algebra. As a result, by Krengel's Theorem, we can conclude that $\{\sigma_t\}_{t\in\mathbb{Z}}$ is an SE sequence as long as $\{z_t\}_{t\in\mathbb{Z}}$ is itself SE.

Since $\{z_t\}_{t\in\mathbb{Z}}$ is generated by a Markov dynamical system,

$$z_{t+1} = \rho_t z_t + v_t \quad \text{for every} \quad t \in \mathbb{Z}\ ,$$

we can show that $\{z_t\}_{t\in\mathbb{Z}}$ is SE by application of Bourgerol's Theorem.

- Condition A1 of Bougerol's Theorem is satisfied since the innovation vector $\{(\rho_t, v_t)\}_{t\in\mathbb{Z}}$ is iid, and hence, it is trivially exogenous and SE.
- Condition A2 of Bougerol's Theorem is satisfied since, for any initialization z_1, we have

$$\begin{aligned}\mathbb{E}\log^+|\rho_t z_1 + v_t| &\le \mathbb{E}|\rho_t z_1 + v_t| \quad \text{(because } \log^+|x| \le |x|\ \forall\, x\text{)}\\ &\le |z_1|\mathbb{E}|\rho_t| + \mathbb{E}|v_t| \quad \text{(by sub-additivity of the absolute value)}\\ &\le |z_1|1.5 + \mathbb{E}|v_t| \quad \text{(since } \{\rho_t\}_{t\in\mathbb{Z}} \sim \text{UID}(0\,,\ 1.5) \text{ and hence } |\rho_t| < 1.5\text{)}\\ &< \infty \quad \text{(because } z_1 \in \mathbb{R}\text{, and } \{v_t\}_{t\in\mathbb{Z}} \sim TID(2) \text{ and hence } \mathbb{E}|v_t|^n < \infty \text{ for any } n < 2\text{)}\end{aligned}$$

- Condition A3 of Bougerol's Theorem (the contraction condition) is satisfied since $\partial(\rho_t z_t + v_t)/\partial z_t = \rho_t$ and hence

$$\begin{aligned}\mathbb{E}\log\sup_z|\rho_t| &= \mathbb{E}\log|\rho_t|\\ &\le \log\mathbb{E}|\rho_t| \quad \text{(Jensen's inequality)}\\ &= \log\mathbb{E}\rho_t \quad (\rho_t \text{ is always positive because } \{\rho_t\}_{t\in\mathbb{Z}} \sim \text{UID}(0\,,\ 1.5))\\ &= \log 0.75 < 0. \quad (\{\rho_t\}_{t\in\mathbb{Z}} \sim \text{UID}(0\,,\ 1.5) \text{ implies that } \mathbb{E}\rho_t = 0.75)\end{aligned}$$

We thus conclude by Bougerol's Theorem that the process $\{z_t(z_1)\}_{t\in\mathbb{N}}$ initialized at any point z_1 converges to a limit SE sequence $\{z_t\}_{t\in\mathbb{Z}}$. Since Bougerol's conditions hold for any initialization $z_1 \in \mathbb{R}$, we can conclude that the sequence starting in the infinite past $\{z_t\}_{t\in\mathbb{Z}}$ is indeed SE.

Finally, as mentioned above, since the process starting in the infinite past $\{z_t\}_{t\in\mathbb{Z}}$ is SE, we conclude that $\{\sigma_t\}_{t\in\mathbb{Z}}$ is also SE.

(b) **(15pts)** Can you show that $\mathbb{E}|x_t|^2 < \infty$? Is $\{x_t\}_{t\in\mathbb{Z}}$ weakly stationary?

Answer: From the previous question (a), we already know that $\{\sigma_t\}_{t\in\mathbb{Z}}$ is SE. Furthermore, since the tanh function is uniformly bounded, we know that σ_t is also uniformly bounded

$$|\sigma_t| = |1 + \tanh(z_t)| \le \sup_z |1 + \tanh(z)| \le 2.$$

We will now show that $\{\mu_t\}_{t\in\mathbb{Z}}$ is SE and has two bounded moments by application of the Power-n Theorem with $n = 2$. First, we note that by substituting x_{t-1} for

$\mu_{t-1} + \sigma_{t-1}\varepsilon_{t-1}$ (observation equation) in the updating equation for μ_t, we find that $\{\mu_t\}_{t\in\mathbb{Z}}$ is generated according to

$$\mu_t = 0.2\sigma_{t-1}\varepsilon_{t-1} + 0.7\mu_{t-1} \quad \text{for every} \quad t \in \mathbb{Z}.$$

Since this is a Markov dynamical system, we can simply verify if the Power-n conditions hold for $n = 2$

- Condition A1 of the Power-n Theorem holds because the innovations vector $\{(\sigma_{t-1}, \varepsilon_{t-1})\}_{t\in\mathbb{Z}}$ is clearly exogenous and SE. The SE nature of $\{\sigma_{t-1})\}_{t\in\mathbb{Z}}$ was shown in the previous question (a), and the SE nature of $\{\varepsilon_{t-1}\}_{t\in\mathbb{Z}}$ follows immediately since this sequence is iid.
- Condition A2 of the Power-n Theorem holds for any initialization μ_1 since

$$\begin{aligned}
\mathbb{E}|0.2\sigma_{t-1}\varepsilon_{t-1} + 0.7\mu_1|^2 &\leq 0.2c\mathbb{E}|\sigma_{t-1}\varepsilon_{t-1}|^2 + 0.7c\mathbb{E}|\mu_1|^2 \quad \text{(by the } c_n\text{-inequality)}\\
&\leq 0.2c\mathbb{E}|\sigma_{t-1}|^2|\mathbb{E}|\varepsilon_{t-1}|^2 + 0.7c|\mu_1|^2 \quad \text{(because } \sigma_{t-1} \text{ and } \varepsilon_{t-1} \text{ are independent)}\\
&\leq 0.2c2\mathbb{E}|\varepsilon_{t-1}|^2 + 0.7c|\mu_1|^2 \quad \text{(because } \sigma_t \text{ is uniformly bounded } \sigma_t < 2 \text{ a.s.)}\\
&< \infty \quad \text{(because } c \in \mathbb{R},\ \mu_1 \in \mathbb{R}, \text{ and } \{\varepsilon_t\}_{t\in\mathbb{Z}} \sim \text{NID}(0,1) \text{ and hence } \mathbb{E}|\varepsilon_{t-1}|^n < \infty\forall n)
\end{aligned}$$

- Condition A3 of the Power-n Theorem (the contraction condition) is satisfied for $n = 2$ since

$$\sup_{\mu}\left|\frac{\partial(0.2\sigma_{t-1}\varepsilon_{t-1} + 0.7\mu)}{\partial\mu}\right|^2 = \sup_{\mu}|0.7|^2 = 0.7^2,$$

  and hence, the degenerate random variable $\rho^n(\varepsilon_{t-1}) = 0.7^2$, which is trivially independent of μ_{t-1}, bounds the uniform derivative of interest, and satisfies the contraction condition

$$\mathbb{E}\rho^2(\varepsilon_{t-1}) = \mathbb{E}0.7^2 < 1.$$

We thus conclude by application the Power-n Theorem, with $n = 2$, that the process $\{\mu_t(\mu_1)\}_{t\in\mathbb{N}}$ initialized at time $t = 1$ with some value μ_1 converges to a limit process $\{\mu_t\}_{t\in\mathbb{Z}}$ which is SE and has two bounded moments $\mathbb{E}|\mu_t|^2 < \infty$. Since, the conditions of the Power-n Theorem were shown to hold for any initialization $\mu_1 \in \mathbb{R}$ we further conclude that the process $\{\mu_t\}_{t\in\mathbb{Z}}$ initialized in the infinite past is indeed SE with two bounded moments.

Finally, we turn to the process $\{x_t\}_{t\in\mathbb{Z}}$. First, we note that $\{x_t\}_{t\in\mathbb{Z}}$ is SE by Krengel's Theorem since it is generated as a continuous function (and hence measurable w.r.t. the Borel sigma-algebra) of the independent SE sequences $\{\mu_t\}_{t\in\mathbb{Z}}$, $\{\sigma_t\}_{t\in\mathbb{Z}}$ and $\{\varepsilon_t\}_{t\in\mathbb{Z}}$. Furthermore, we note that any element x_t of the sequence $\{x_t\}_{t\in\mathbb{Z}}$ has two bounded moments since

$$\begin{aligned}
\mathbb{E}|x_t|^2 &\leq \mathbb{E}|\mu_t + \sigma_t\varepsilon_t|^2\\
&\leq c\mathbb{E}|\mu_t|^2 + c\mathbb{E}|\sigma_t\varepsilon_t|^2 \quad \text{(by the } c_n\text{-inequality)}\\
&\leq c\mathbb{E}|\mu_t|^2 + c\mathbb{E}|\sigma_t|^2\mathbb{E}|\varepsilon_t|^2 \quad \text{(because } \sigma_t \text{ and } \varepsilon_t \text{ are independent)}\\
&\leq c\mathbb{E}|\mu_t|^2 + c4\mathbb{E}|\varepsilon_t|^2 \quad \text{(because } \sigma_t \leq 2 \text{ a.s.)}\\
&< \infty \quad \text{(because } c \in \mathbb{R},\ \mathbb{E}|\mu_t|^2 < \infty \text{ and } \{\varepsilon_t\}_{t\in\mathbb{Z}} \sim \text{NID}(0,1) \text{ and hence } \mathbb{E}|\varepsilon_t|^n < \infty\forall n\ )
\end{aligned}$$

Since $\{x_t\}_{t\in\mathbb{Z}}$ is an SE sequence with two bounded moments $\mathbb{E}|x_t|^2 < \infty$, we can conclude that $\{x_t\}_{t\in\mathbb{Z}}$ is weakly stationary.

Question 2 [25 points] Consistency and Asymptotic Normality of M-Estimators

Let $\mathbf{x}_T := (x_1, ..., x_T)$ be a subset of a fat-tailed strictly stationary and ergodic sequence $\{x_t\}_{t\in\mathbb{Z}}$ satisfying $\mathbb{E}|x_t|^4 < \infty$. It is well known that the least squares estimator is sensitive to the presence of outliers in the data. Let $\hat{\boldsymbol{\theta}}_T$ be a robust M-estimator given by

$$\hat{\boldsymbol{\theta}}_T \in \arg\max_{\boldsymbol{\theta}\in\Theta} -\frac{1}{T}\sum_{t=2}^{T} \frac{u_t(\boldsymbol{\theta})^2}{1+u_t(\boldsymbol{\theta})^2}$$

where $u_t(\boldsymbol{\theta})$ denotes the regression residuals of a nonlinear autoregressive model

$$u_t(\boldsymbol{\theta}) := x_t - \phi(x_{t-1}, \boldsymbol{\theta}) \quad \text{for every } t.$$

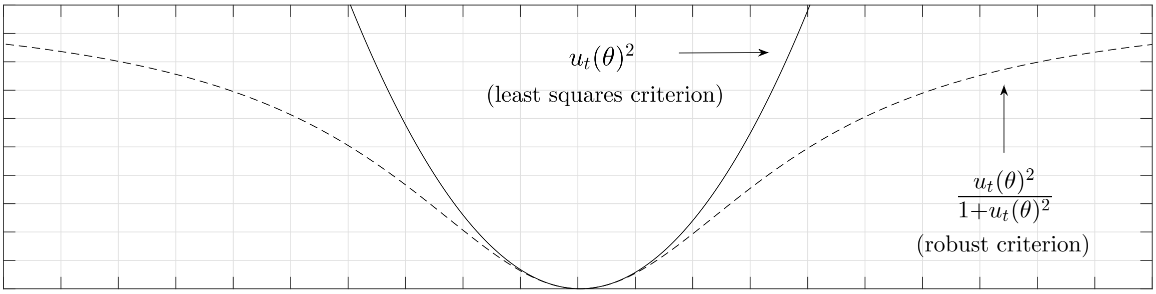

Figure 2: Comparison of the least squares and robust least squares criterion function.

Note: The function $u_t(\boldsymbol{\theta})^2/(1+u_t(\boldsymbol{\theta})^2)$ is uniformly bounded between 0 and 1.

(a) (**10 pts**) Give sufficient conditions for the existence and measurability of the estimator $\hat{\boldsymbol{\theta}}_T$.

Answer: Define $Q(\mathbf{x}_T, \boldsymbol{\theta}) := -\frac{1}{T}\sum_{t=2}^{T} q(x_t, x_{t-1}, \boldsymbol{\theta})$ where

$$q(x_t, x_{t-1}, \boldsymbol{\theta}) := \frac{u_t(\boldsymbol{\theta})^2}{1+u_t(\boldsymbol{\theta})^2} = \frac{(x_t - \phi(x_{t-1}, \boldsymbol{\theta}))^2}{1+(x_t - \phi(x_{t-1}, \boldsymbol{\theta}))^2}.$$

Sufficient conditions for the existence and measurability of the estimator $\hat{\boldsymbol{\theta}}_T$ are:

(i) Compactness of the parameter space Θ.

(ii) Continuity of the criterion function $Q_T(\cdot, \boldsymbol{\theta}) : \mathbb{R}^T \to \mathbb{R}$ for every $\boldsymbol{\theta} \in \Theta$. (which is ensured <u>if</u> the function $\phi(\cdot, \boldsymbol{\theta}) : \mathbb{R} \to \mathbb{R}$ is continuous for every $\boldsymbol{\theta} \in \Theta$).

(iii) Continuity of the criterion function $Q_T(\mathbf{x}_T, \cdot) : \Theta \to \mathbb{R}$ for every sample point $\mathbf{x}_T \in \mathbb{R}^T$ (which is ensured <u>if</u> the function $\phi(x_{t-1}, \cdot) : \Theta \to \mathbb{R}$ is continuous for every data point $x_{t-1} \in \mathbb{R}$).

The compactness of Θ and the continuity of ϕ in both arguments are thus the two additional sufficient conditions needed for the existence and measurability of the robust least squares estimator.

(b) (**15 pts**) Give sufficient conditions for $\hat{\boldsymbol{\theta}}_T$ to be consistent for some point $\boldsymbol{\theta}_0 \in \Theta$. In other words, give conditions that ensure $\hat{\boldsymbol{\theta}}_T \xrightarrow{p} \boldsymbol{\theta}_0$ as $T \to \infty$.
Note: you can assume that certain functions are well behaved and continuously differentiable.

Answer: Define again $Q_T(\mathbf{x}_t, \boldsymbol{\theta}) := -\frac{1}{T}\sum_{t=2}^{T} q(x_t, x_{t-1}, \boldsymbol{\theta})$ where

$$q(x_t, x_{t-1}, \boldsymbol{\theta}) := \frac{u_t(\boldsymbol{\theta})^2}{1 + u_t(\boldsymbol{\theta})^2}.$$

Under the compactness of Θ and the continuity conditions mentioned in the previous question, we know first of all that $\hat{\boldsymbol{\theta}}_T$ exists and is measurable. Hence, by the classical consistency theorem for M-estimators, we can obtain the consistency $\hat{\boldsymbol{\theta}}_T \xrightarrow{p} \boldsymbol{\theta}_0$ as $T \to \infty$ from the following two conditions

C1. The uniform convergence of the criterion function

$$\sup_{\boldsymbol{\theta} \in \Theta} \left| Q_T(\mathbf{x}_T, \boldsymbol{\theta}) - Q_\infty(\boldsymbol{\theta}) \right| \xrightarrow{p} 0 \quad \text{as} \quad T \to \infty.$$

C2. The identifiable uniqueness of $\boldsymbol{\theta}_0 \in \Theta$,

$$\sup_{\boldsymbol{\theta} \in S^c(\boldsymbol{\theta}_0, \delta)} Q_\infty(\boldsymbol{\theta}) < Q_\infty(\boldsymbol{\theta}_0) \quad \text{for every } \delta > 0,$$

where $S^c(\boldsymbol{\theta}_0, \delta)$ denotes the complement of an open ball of radius δ centered at $\boldsymbol{\theta}_0$.

The uniform convergence of the criterion function (Condition C1) is implied by the pointwise convergence of the criterion Q_T

C1.1. $Q_T(\mathbf{x}_T, \boldsymbol{\theta}) \xrightarrow{p} Q_\infty(\boldsymbol{\theta}) \ \forall\ \boldsymbol{\theta} \in \Theta$ as $T \to \infty$,

and the stochastic equicontinuity of the criterion Q_T

C1.2. $\sup_T \mathbb{E} \sup_{\boldsymbol{\theta} \in \Theta} \left\| \frac{\partial Q_T(\mathbf{x}_T, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}} \right\| < \infty.$

The pointwise convergence of the criterion (Condition C1.1) can be obtained by the application of a law of large numbers to the criterion function for every $\boldsymbol{\theta} \in \Theta$. In particular, <u>if</u> $\phi(\cdot, \boldsymbol{\theta}) : \mathbb{R} \to \mathbb{R}$ is continuous on $\mathbb{R}$ for every $\boldsymbol{\theta} \in \Theta$ (as assumed in the previous question), then it is also measurable w.r.t. the Borel σ-algebra, and hence, by Krengel's Theorem, $\{q(x_t, x_{t-1}, \boldsymbol{\theta})\}_{t \in \mathbb{N}}$ is SE for every $\boldsymbol{\theta} \in \Theta$, because q is a measurable function of the SE sequence $\{x_t\}_{t \in \mathbb{Z}}$ for every $\boldsymbol{\theta} \in \Theta$. Furthermore, it is easy to show that $q(x_t, x_{t-1}, \boldsymbol{\theta})$ has one bounded moment for every $\boldsymbol{\theta} \in \Theta$ since $q(x_t, x_{t-1}, \boldsymbol{\theta})$ is actually uniformly bounded

$$\begin{aligned}\mathbb{E}|q(x_t, x_{t-1}, \boldsymbol{\theta})| &= \mathbb{E}\Big|\frac{u_t(\boldsymbol{\theta})^2}{1 + u_t(\boldsymbol{\theta})^2}\Big| \\ &\leq \sup_u \Big|\frac{u^2}{1+u^2}\Big| \leq 1 \quad \text{(because } z^2/(1+z^2) \text{ is unformly bounded by 1)}\end{aligned}$$

We thus obtain the pointwise convergence (condition C1.1) by application of the LLN for SE sequences for every $\boldsymbol{\theta} \in \Theta$

$$\frac{1}{T}\sum_{t=1}^{T} q(x_t, x_{t-1}, \boldsymbol{\theta}) \xrightarrow{p} \mathbb{E}q(x_t, x_{t-1}, \boldsymbol{\theta}) \;\; \forall\, \boldsymbol{\theta} \in \Theta \quad \text{as } T \to \infty.$$

Since q is continuously differentiable and well behaved of order 1, the moment bound $\mathbb{E}|q(x_t, x_{t-1}, \boldsymbol{\theta})| < \infty$ for some $\boldsymbol{\theta} \in \Theta$ implies also that

$$\begin{aligned}\sup_T \mathbb{E} \sup_{\boldsymbol{\theta}\in\Theta} \Big\|\frac{\partial Q_T(\mathbf{x}_T, \boldsymbol{\theta})}{\partial \boldsymbol{\theta}}\Big\| &= \mathbb{E} \sup_{\boldsymbol{\theta}\in\Theta} \Big\|\frac{1}{T}\sum_{t=2}^{T} \nabla q(x_t, x_{t-1}, \boldsymbol{\theta})\Big\| \\ &\leq \frac{1}{T}\sum_{t=2}^{T} \mathbb{E} \sup_{\boldsymbol{\theta}\in\Theta} \Big\|\nabla q(x_t, x_{t-1}, \boldsymbol{\theta})\Big\| < \infty\end{aligned}$$

and hence the stochastic equicontinuity condition (C1.2) is also satisfied.

Finally, <u>if</u> $\boldsymbol{\theta}_0 \in \Theta$ is the unique maximizer of the limit criterion function

$$Q_\infty(\boldsymbol{\theta}) = \mathbb{E}q(x_t, x_{t-1}, \boldsymbol{\theta}) \; < \; \mathbb{E}q(x_t, x_{t-1}, \boldsymbol{\theta}_0) = Q_\infty(\boldsymbol{\theta}_0) \;\; \forall\, \boldsymbol{\theta} \neq \boldsymbol{\theta}_0$$

then, given the compactness of Θ and the continuity of the limit criterion Q_∞ on Θ we obtain that $\boldsymbol{\theta}_0$ is the identifiably unique maximizer of Q_∞ (condition C.2).

Since conditions C.1 (uniform convergence) and C.2 (identifiable uniqueness) stated above hold, we can conclude that $\hat{\boldsymbol{\theta}}_T$ is consistent for $\boldsymbol{\theta}_0$, i.e. that $\hat{\boldsymbol{\theta}}_T \xrightarrow{p} \boldsymbol{\theta}_0$. Note that we have imposed the following (non-stated) conditions to obtain this result:

(1) Θ is compact,

(2) ϕ is continuous in x_{t-1} and $\boldsymbol{\theta}$,

(3) $\boldsymbol{\theta}_0$ is the unique maximizer of the limit criterion function Q_∞.

Question 3 [25 points] Nonlinear Dynamic Model of Dutch GDP

Some econometricians claim that the temporal dependence in the growth rate of the *Gross Domestic Product* (GDP) is stronger during economic recession periods and weaker during expansions. Figure 1 plots the *growth rate of quarterly real GDP in The Netherlands.*

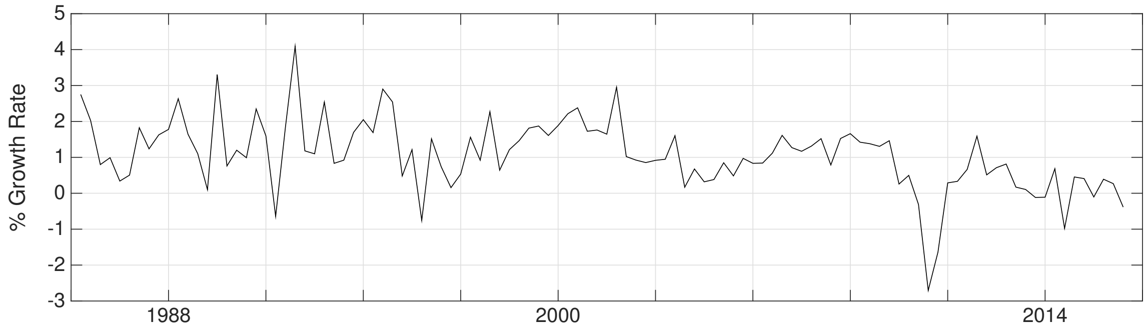

Figure 3: Real GDP growth rate for The Netherlands (in percentage).

Let the sample of GDP growth rates $\{x_t\}_{t=1}^T$ at your disposal be a subset of the realized path of a strictly stationary and ergodic time-series $\{x_t\}_{t\in\mathbb{Z}}$ with bounded moments of fourth order $\mathbb{E}|x_t|^4 < \infty$. Consider the Gaussian logistic *Self Excited Smooth Transition Autoregressive* (SESTAR) model

$$x_t = \alpha + g(x_{t-1};\boldsymbol{\theta})x_{t-1} + \varepsilon_t \quad \text{for every} \quad t \in \mathbb{Z} \quad \text{where} \quad \{\varepsilon_t\}_{t\in\mathbb{Z}} \sim \text{NID}(0,\sigma_\varepsilon^2)$$

$$\text{and} \quad g(x_{t-1};\boldsymbol{\theta}) := \frac{\gamma}{1+\exp(\beta x_{t-1})} \quad \text{for every} \quad t \in \mathbb{Z}.$$

Suppose that the parameters $\boldsymbol{\theta} = (\alpha,\gamma,\beta,\sigma_\varepsilon^2)$ of the model are estimated by maximum likelihood (ML) on a compact parameter space Θ with $\sigma_\varepsilon^2 > 0$. Note also that $g(x;\boldsymbol{\theta})$ is uniformly bounded since $|g(x;\boldsymbol{\theta})| \le |\gamma|$ for every $(x,\boldsymbol{\theta})$.

(a) (**15pts**) Suppose that the ML estimator $\hat{\boldsymbol{\theta}}_T$ is consistent for a parameter $\boldsymbol{\theta}_0$ in the interior of Θ. Can you obtain an approximate distribution for the ML estimator?
Note: you can assume that certain functions are well behaved and continuously differentiable.

Answer: Note first that, according to the SESTAR model, the distribution of x_t conditional on x_{t-1} is given by

$$x_t|x_{t-1} \sim N\Big(\alpha + g(x_{t-1};\boldsymbol{\theta})x_{t-1}\ ,\ \sigma_\varepsilon^2\Big).$$

Hence, the conditional density of x_t given x_{t-1} takes the form,

$$f(x_t|x_{t-1};\boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma_\varepsilon^2}}\exp\left(\frac{-\big(x_t-\alpha-g(x_{t-1};\boldsymbol{\theta})x_{t-1}\big)^2}{2\sigma_\varepsilon^2}\right).$$

and the log likelihood function takes the form

$$Q_T(\boldsymbol{\theta}) := \frac{1}{T}\sum_{t=2}^{T} q(x_t,x_{t-1},\boldsymbol{\theta}) \quad \text{where}$$

$$\begin{aligned} q(x_t, x_{t-1}, \boldsymbol{\theta}) :&= \log f(x_t | x_{t-1}; \boldsymbol{\theta}) \\ &= -\frac{1}{2}\log 2\pi - \frac{1}{2}\log \sigma_\varepsilon^2 - \frac{(x_t - \alpha - g(x_{t-1}; \boldsymbol{\theta})x_{t-1})^2}{2\sigma_\varepsilon^2}. \end{aligned}$$

Since the ML estimator is consistent, i.e. $\hat{\boldsymbol{\theta}}_T \xrightarrow{p} \boldsymbol{\theta}_0$ as $T \to \infty$, to a point $\boldsymbol{\theta}_0$ in the interior of the parameter space Θ, we can apply the classical asymptotic normality theorem. In particular, if the following conditions hold,

C1. Asymptotic normality of the score

$$\sqrt{T}\frac{1}{T}\sum_{t=2}^{T} \nabla q(x_t, x_{t-1}, \boldsymbol{\theta}_0) \xrightarrow{d} N(0, \Sigma) \quad \text{as } T \to \infty.$$

C2. Uniform convergence of the second derivative

$$\sup_{\boldsymbol{\theta}\in\Theta} \|\frac{1}{T}\sum_{t=2}^{T} \nabla^2 q(x_t, x_{t-1}, \boldsymbol{\theta}) - \mathbb{E}\nabla^2 q(x_t, x_{t-1}, \boldsymbol{\theta})\| \xrightarrow{p} 0 \quad \text{as } T \to \infty.$$

C3. Invertibility of $\mathbb{E}\nabla^2 q(x_t, x_{t-1}, \boldsymbol{\theta}_0)$.

Then we can conclude that $\hat{\boldsymbol{\theta}}_T$ is asymptotically Gaussian

$$\sqrt{T}\Big(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\Big) \xrightarrow{d} N\Big(0, \Omega\Sigma\Omega^\top\Big) \quad \text{as } T \to \infty \text{ where } \Omega := \Big(\mathbb{E}\nabla^2 q(x_t, x_{t-1}, \boldsymbol{\theta}_0)\Big)^{-1}.$$

The asymptotic normality of the score (Condition C.1) can be obtained by application of a Central Limit Theorem (CLT). In particular, we note that $\{\nabla q(x_t, x_{t-1}, \boldsymbol{\theta}_0)\}_{t\in\mathbb{Z}}$ satisfies a CLT if:

C1.1. The sequence $\{\nabla q(x_t, x_{t-1}, \boldsymbol{\theta}_0)\}_{t\in\mathbb{Z}}$ is SE.

C1.2. $\nabla q(x_t, x_{t-1}, \boldsymbol{\theta}_0)$ has two bounded moments $\mathbb{E}|\nabla q(x_t, x_{t-1}, \boldsymbol{\theta}_0)|^2 < \infty$.

C1.3. $\{\nabla q(x_t, x_{t-1}, \boldsymbol{\theta}_0)\}_{t\in\mathbb{Z}}$ is either a martingale difference sequence, or it is L_p-approximable.

Condition C1.1 holds because q is continuously differentiable, and hence ∇q is continuous (and also measurable w.r.t. the Borel σ-algebra). This implies, by Krengel's Theorem, that $\{\nabla q(x_t, x_{t-1}, \boldsymbol{\theta}_0)\}_{t\in\mathbb{Z}}$ is SE since ∇q is a measurable function of the SE sequence $\{x_t\}_{t\in\mathbb{Z}}$.

Condition C1.2 holds since $q(x_t, x_{t-1}, \boldsymbol{\theta})$ has two bounded moments and the criterion function q is continuously differentiable and well behaved of order 2. The two bounded moments are obtained as follows

$$
\begin{aligned}
\mathbb{E}|q(x_t, x_{t-1}, \boldsymbol{\theta})|^2 &= \mathbb{E}\Big| -\frac{1}{2}\log 2\pi - \frac{1}{2}\log \sigma_\varepsilon^2 - \frac{(x_t - \alpha - g(x_{t-1};\boldsymbol{\theta})x_{t-1}))^2}{2\sigma_\varepsilon^2}\Big|^2 \\
&\qquad \text{(by the definition of } q(x_t, x_{t-1}, \boldsymbol{\theta})\text{)} \\
&\leq c\mathbb{E}|\frac{1}{2}\log 2\pi|^2 + c\mathbb{E}|\frac{1}{2}\log \sigma_\varepsilon^2|^2 + c\mathbb{E}\Big|\frac{(x_t - \alpha - g(x_{t-1};\boldsymbol{\theta})x_{t-1}))^2}{2\sigma_\varepsilon^2}\Big|^2 \\
&\qquad \text{(by applying the } c_n\text{-inequality with some constant } c\text{)} \\
&\leq c(\frac{1}{2}\log 2\pi)^2 + c(\frac{1}{2}\log \sigma_\varepsilon^2)^2 + c\frac{1}{4\sigma_\varepsilon^4}\mathbb{E}\Big|x_t - \alpha - g(x_{t-1};\boldsymbol{\theta})x_{t-1})\Big|^4 \\
&\qquad \text{(dropping the expectation for all constants)} \\
&\leq A + c\frac{1}{4\sigma_\varepsilon^4}\mathbb{E}\Big|x_t - \alpha - g(x_{t-1};\boldsymbol{\theta})x_{t-1})\Big|^4 \\
&\qquad \text{(defining the constant } A := c(\frac{1}{2}\log 2\pi)^2 + c(\frac{1}{2}\log \sigma_\varepsilon^2)^2\text{)} \\
&\leq A + c'\frac{c}{4\sigma_\varepsilon^4}\mathbb{E}|x_t|^4 + c'\frac{c}{4\sigma_\varepsilon^4}|\alpha|^4 + c'\frac{c}{4\sigma_\varepsilon^4}\mathbb{E}|g(x_{t-1};\boldsymbol{\theta})x_{t-1}|^4 \\
&\qquad \text{(applying again the } c_n\text{-inequality for some constant } c'\text{)} \\
&\leq A' + c'\frac{c}{4\sigma_\varepsilon^4}\mathbb{E}|x_t|^4 + c'\frac{c}{4\sigma_\varepsilon^4}\sup_x |g(x_{t-1};\boldsymbol{\theta})|^4\mathbb{E}|x_{t-1}|^4 \\
&\qquad \text{(because } \sup_x |g(x;\boldsymbol{\theta})| \text{ is a constant it is taken outside the expectation)} \\
&\qquad \text{(defining also } A' := A + c'\frac{c}{4\sigma_\varepsilon^4}|\alpha|^4\text{)} \\
&\leq A' + c'\frac{c}{4\sigma_\varepsilon^4}\mathbb{E}|x_t|^4 + c'\frac{c}{4\sigma_\varepsilon^4}\ \bar{k}\ \mathbb{E}|x_{t-1}|^4 \\
&\qquad (\sup_x |g(x;\boldsymbol{\theta})| \leq \bar{k} \text{ for some } \bar{k} \in \mathbb{R} \text{ because } g(x;\boldsymbol{\theta}) \text{ is uniformly bounded)} \\
&< \infty. \\
&\qquad \text{(because } A' \in \mathbb{R},\ c \in \mathbb{R},\ c' \in \mathbb{R},\ \sigma_\varepsilon^4 > 0 \Rightarrow 1/\sigma_\varepsilon^4 \in \mathbb{R}, \text{ and } \mathbb{E}|x_t|^4 = \mathbb{E}|x_{t-1}|^4 < \infty\text{)}
\end{aligned}
$$

Condition C1.3 holds since: (i) if the model is well specified, then $\{\nabla q(x_t, x_{t-1}, \boldsymbol{\theta}_0)\}_{t\in\mathbb{Z}}$ is a martingale difference sequence; and (ii) if the model is misspecified, then $\{\nabla q(x_t, x_{t-1}, \boldsymbol{\theta}_0)\}_{t\in\mathbb{Z}}$ is L_p-approximable because $\boldsymbol{\theta}_0$ is the unique maximizer of the limit criterion (by assumption) and q is continuously differentiable and well behaved of order 1.

We thus conclude that the score is asymptotically normal, i.e. that

$$\sqrt{T}\frac{1}{T}\sum_{t=2}^{T}\nabla q(x_t, x_{t-1}, \boldsymbol{\theta}_0) \xrightarrow{d} N(0, \Sigma) \quad \text{as } T \to \infty.$$

The uniform convergence of the second derivative (condition C2) can be obtained by ensuring that the following three conditions hold:

C2.1. Θ is compact.

C2.2. The second derivative satisfies a pointwise LLN

$$\frac{1}{T}\sum_{t=2}^{T}\nabla^2 q(x_t,x_{t-1},\boldsymbol{\theta}) \overset{p}{\to} \mathbb{E}\nabla^2 q(x_t,x_{t-1},\boldsymbol{\theta})\ \forall\ \boldsymbol{\theta}\in\Theta \quad \text{as } T\to\infty.$$

C2.3. The second derivative is stochastically equicontinuous

$$\sup_t \mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta}\left|\nabla^3 q(x_t,x_{t-1},\boldsymbol{\theta})\right| < \infty.$$

The compactness of the parameter space (condition C2.1) holds by assumption.

The pointwise LLN (condition C2.2) holds since the following conditions are satisfied

C2.2.1 $\{\nabla^2 q(x_t,x_{t-1},\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ is SE for every $\boldsymbol{\theta}\in\Theta$.

C2.2.2 $\mathbb{E}|\nabla^2 q(x_t,x_{t-1},\boldsymbol{\theta})|<\infty$ for every $\boldsymbol{\theta}\in\Theta$.

The SE nature of the second derivative (condition C2.2.1) holds for every $\boldsymbol{\theta}\in\Theta$ because q is twice continuously differentiable, and hence $\nabla^2 q$ is continuous (and also measurable w.r.t. the Borel σ-algebra) for every $\boldsymbol{\theta}\in\Theta$. This implies, by Krengel's Theorem, that $\{\nabla^2 q(x_t,x_{t-1},\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ is SE for every $\boldsymbol{\theta}\in\Theta$ because $\nabla^2 q$ is a measurable function of the SE sequence $\{x_t\}_{t\in\mathbb{Z}}$ for every $\boldsymbol{\theta}\in\Theta$.

The moment bound for the second derivative (condition C2.2.2) holds since $q(x_t,x_{t-1},\boldsymbol{\theta})$ has one bounded moment (see derivations above), the criterion function q is two times continuously differentiable, and both q and ∇q are well behaved of first order.

The stochastic equicontinuity of the criterion's second derivative (condition C2.3) holds because $q(x_t,x_{t-1},\boldsymbol{\theta})$ has one bounded moment. First, since the criterion function q is three times continuously differentiable, we can conclude that $\{\nabla^3 q(x_t,x_{t-1},\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ is SE by Krengel's Theorem, because $\nabla^3 q$ is continuous (and hence measurable w.r.t. Borel's σ-algebra) on the SE sequence $\{x_t\}_{t\in\mathbb{Z}}$. The SE nature of $\{\nabla^3 q(x_t,x_{t-1},\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ implies also that

$$\sup_t \mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta}\left|\nabla^3 q(x_t,x_{t-1},\boldsymbol{\theta})\right| = \mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta}\left|\nabla^3 q(x_t,x_{t-1},\boldsymbol{\theta})\right|.$$

Second, since q, ∇q and $\nabla^2 q$ are well behaved of first order, the moment bound on the criterion $\mathbb{E}|q(x_t,x_{t-1},\boldsymbol{\theta})|<\infty$ (shown above) ensures that $\nabla^3 q(x_t,x_{t-1},\boldsymbol{\theta})$ has one bounded moment uniformly in Θ

$$\mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta}\left|\nabla^3 q(x_t,x_{t-1},\boldsymbol{\theta})\right| < \infty$$

and hence the second derivative is stochastically equicontinuous.

We have thus verified conditions C2.1, C2.2and C2.3 and are able to conclude that the second derivative converges uniformly (condition C2).

Finally, we turn to the invertibility condition C3. In particular, we obtain immediately that $\mathbb{E}\nabla^2 q(x_t,x_{t-1},\boldsymbol{\theta}_0)$ is invertible because $\boldsymbol{\theta}_0$ is the unique maximizer of the limit criterion function (given by assumption).

Since conditions C1, C2 and C3, of the classical asymptotic normality theorem hold, we conclude that

$$\sqrt{T}\Big(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\Big) \xrightarrow{d} N\Big(0, \Omega\Sigma\Omega^\top\Big) \quad \text{as } T \to \infty \text{ where } \Omega := \Big(\mathbb{E}\nabla^2 q(x_t, x_{t-1}, \boldsymbol{\theta}_0)\Big)^{-1}.$$

This asymptotic result suggests that, in finite samples, we have

$$\sqrt{T}\Big(\hat{\boldsymbol{\theta}}_T - \boldsymbol{\theta}_0\Big) \approx N\Big(0, \Omega\Sigma\Omega^\top\Big)$$

which naturally implies that $\hat{\boldsymbol{\theta}}_T$ is approximately Gaussian, with mean $\boldsymbol{\theta}_0$, and a vanishing variance-covariance matrix given by $\Omega\Sigma\Omega^\top/T$,

$$\hat{\boldsymbol{\theta}}_T \approx N\Big(\boldsymbol{\theta}_0 \ , \ \Omega\Sigma\Omega^\top/T\Big).$$

(b) (**10pts**) Explain how you can use the approximate distribution of $\hat{\boldsymbol{\theta}}_T$ to test the claim that the temporal dependence in the growth rate of Dutch GDP is stronger during economic recession periods and weaker during expansions.

Answer: Testing the hypothesis $\mathrm{H}_0 : \beta_0 = 0$ against the one-sided alternative that β_0 is strictly positive, $\mathrm{H}_1 : \beta > 0$, is probably the most natural way of addressing this question. Indeed, for $\beta = 0$, the SESTAR model reduces to the linear AR(1)

$$x_t = \alpha + \gamma x_{t-1} + \varepsilon_t.$$

This implies that the temporal dependence is the same during expansions and recessions under the null H_0. In contrast, under the alternative hypothesis H_1, the temporal dependence in the growth rate of Dutch GDP is stronger during economic recession periods and weaker during expansions (at least for a positive $\gamma > 0$ which the plotted data suggests). Hence, if we reject the null hypothesis that $\beta_0 = 0$ we favour the claim that *'the growth rate of Dutch GDP is stronger during economic recession periods and weaker during expansions'*.

In practice, we can use the approximate distribution derived in the previous questions as follows. If $H_0 : \beta = 0$ is true, then the asymptotic normality results tells us that the estimator $\hat{\beta}_T$ is approximately Gaussian with mean zero and some variance σ^2_β, i.e.

$$\hat{\beta}_T \approx N(0, \ \sigma^2_\beta).$$

After substituting σ^2_β by some consistent estimator $\hat{\sigma}^2_\beta$, we can finally calculate the *approximate* tail probability of any given point estimate obtained from the data and judge how reasonable the hull hypothesis seems to be. The final decision depends on the adopted significance level.

Note: If the model is misspecified then the estimator $\hat{\sigma}^2_\beta$ should be robust to take into account the potential temporal dependence in the score. Furthermore, in this case, the null $\mathrm{H}_0 : \beta_0 = 0$ is a test on the MLE's pseudo-true parameter. The test then asks essentially if the best approximation to the DGP (in KL divergence, over the space of probability measures for the data) is delivered by a linear AR(1) model (H_0), or by a nonlinear SESTAR (H_1).

Question 4 [25 points]   Time-varying Conditional Volatility in Stock Markets

Financial returns often exhibit 'clusters of volatility' and 'leverage effects'. Figure 2 plots the time-series of *daily percentage returns* for the S&P500 stock market index.

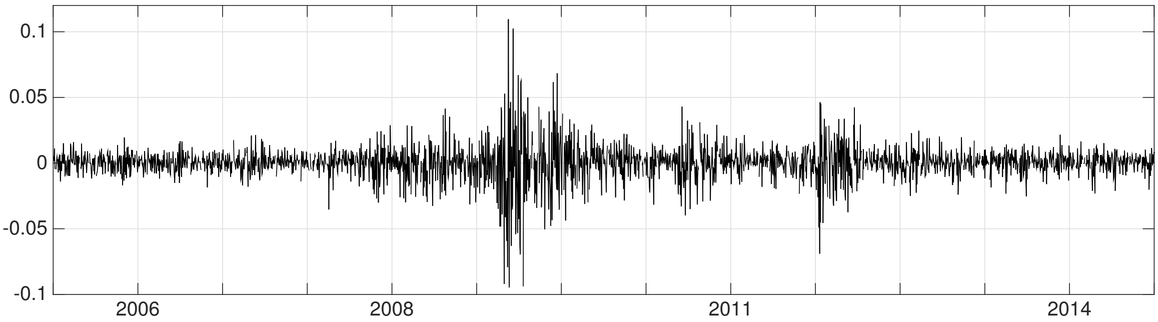

Figure 4: Daily S&P500 percentage returns.

Let the sample of S&P500 returns $\{x_t\}_{t=1}^T$ at your disposal be a subset of the realized path of a strictly stationary and ergodic time-series $\{x_t\}_{t\in\mathbb{Z}}$ satisfying $\mathbb{E}|x_t|^8 < \infty$. Consider the *Asymetric Generalized Autoregressive Conditional Heteroscedasticity* (AGARCH) model

$$x_t = \sigma_t \varepsilon_t \quad \text{for every} \quad t \in \mathbb{Z} \quad \text{where} \quad \{\varepsilon_t\}_{t\in\mathbb{Z}} \sim \text{NID}(0,1),$$

$$\text{where} \quad \sigma_t^2 = \omega + \alpha(x_{t-1} - \delta)^2 + \beta\sigma_{t-1}^2 \quad \text{for every} \quad t \in \mathbb{Z}.$$

Suppose that the parameters $\boldsymbol{\theta} = (\omega, \alpha, \delta, \beta)$ of the model are estimated by maximum likelihood (ML) on a compact parameter space Θ with ω, α and β satisfying

$$\omega > a, \quad \alpha > a, \quad \text{and} \quad a < \beta < 1 \quad \text{for some} \quad a > 0.$$

Note: that the parameter restrictions ensure that $\sigma_t^2 > a > 0$ for every t.

(a) (**7pts**) Give the expression for the log likelihood function.

Answer: Note first that, according to the AGARCH model, the conditional volatility σ_t^2 is known conditional on past data $x_1, x_2, ..., x_{t-1}$ and the initialization point $\sigma_1^2 \in \mathbb{R}^+$. As such, the distribution of x_t conditional on the past $x_{t-1}, x_{t-2}, ...$ is given by

$$x_t | x_{t-1}, x_{t-2}, ... \sim N\left(0 \ , \ \sigma_t^2(\sigma_1^2, \boldsymbol{\theta})\right)$$

where $\sigma_t^2(\sigma_1^2, \boldsymbol{\theta})$ denotes the filtered volatility at time t, obtained under the initialization $\sigma_1^2 \in \mathbb{R}^+$ and the parameter vector $\boldsymbol{\theta} \in \Theta$. Hence, the conditional density of x_t given $x_{t-1}, x_{t-2}, ...$ takes the form,

$$f(x_t|x_{t-1}, x_{t-2}, ...; \boldsymbol{\theta}) = f(x_t|\sigma_t^2(\sigma_1^2, \boldsymbol{\theta})) = \frac{1}{\sqrt{2\pi\sigma_t^2(\sigma_1^2, \boldsymbol{\theta})}} \exp\left(\frac{-x_t^2}{2\sigma_t^2(\sigma_1^2, \boldsymbol{\theta})}\right).$$

Naturally, since the joint density of the data can be factorized as a product of conditional densities, we obtain (ignoring the first marginal)

$$f(x_1, ..., x_T; \boldsymbol{\theta}) = \prod_{t=2}^{T} f(x_t | x_{t-1}, x_{t-2}, ...; \boldsymbol{\theta}) = \prod_{t=2}^{T} f(x_t | \sigma_t^2(\sigma_1^2, \boldsymbol{\theta}))$$

and hence

$$\log f(x_1, ..., x_T; \boldsymbol{\theta}) = \sum_{t=2}^{T} \log f(x_t | \sigma_t^2(\sigma_1^2, \boldsymbol{\theta}))$$

As such the normalized log likelihood function takes the form

$$Q_T(\mathbf{x}_T, \boldsymbol{\theta}) := \frac{1}{T} \sum_{t=2}^{T} q(x_t, \sigma_t^2(\sigma_1^2, \boldsymbol{\theta})) \quad \text{where}$$

$$\begin{aligned} q(x_t, \sigma_t^2(\sigma_1^2, \boldsymbol{\theta})) :&= \log f(x_t | \sigma_t^2(\sigma_1^2, \boldsymbol{\theta})) \\ &= -\frac{1}{2} \log 2\pi - \frac{1}{2} \log \sigma_t^2(\sigma_1^2, \boldsymbol{\theta}) - \frac{x_t^2}{2\sigma_t^2(\sigma_1^2, \boldsymbol{\theta})}. \end{aligned}$$

(b) (**18pts**) Suppose that there exists a $\boldsymbol{\theta}_0 \in \Theta$ that is the unique maximizer of the limit log likelihood function. Can you show that the ML estimator $\hat{\boldsymbol{\theta}}_T$ is consistent for $\boldsymbol{\theta}_0$?

Note: you can assume that certain functions are well behaved and continuously differentiable.

Answer: In preparation for the consistency result, we establish first the stochastic properties of the filter (Part I), and later, turn to the consistency argument (Part II).

PART I

As we shall now see, for every $\boldsymbol{\theta} \in \Theta$, and any initialization $\sigma_1^2 \in \mathbb{R}^+$, the filtered volatility $\{\sigma_t^2(\sigma_1^2, \boldsymbol{\theta})\}_{t\in\mathbb{N}}$ initialized at time $t = 1$, at the value σ_1^2, converges e.a.s. to a unique limit SE sequence $\{\sigma_t^2(\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ with two bounded moments $\mathbb{E}|\sigma_t^2(\boldsymbol{\theta})|^2 < \infty$.

We obtain this result by analyzing the update equation

$$\sigma_t^2 = \omega + \alpha(x_{t-1} - \delta)^2 + \beta\sigma_{t-1}^2$$

which takes the form of a Markov dynamical system with differentiable autoregressive dynamics. In particular, we proceed by verifying that the conditions of the Power-n Theorem hold for any $\boldsymbol{\theta} \in \Theta$, any $\sigma_1^2 \in \mathbb{R}^+$, and $n = 2$:

- Condition A1 of the Power-n Theorem holds because the innovations $\{x_t\}_{t\in\mathbb{Z}}$ are exogenous and SE (given).

- Condition A2 of the Power-n Theorem is satisfied for any $\boldsymbol{\theta}$, any σ_1^2 and $n = 2$ since

$$\begin{aligned}\mathbb{E}|\omega + \alpha(x_{t-1} - \delta)^2 + \beta\sigma_1^2|^2 &\leq c\mathbb{E}|\omega|^2 + c\mathbb{E}|\alpha(x_{t-1} - \delta)^2|^2 + c\mathbb{E}|\beta\sigma_1^2|^2 \\ &\quad \text{(c_n-inequality for some $c \in \mathbb{R}$)} \\ &= c\omega^2 + c\alpha^2\mathbb{E}|x_{t-1} - \delta|^4 + c\beta^2\sigma_1^4 \\ &\quad \text{(dropping expectations of constants)} \\ &= c\omega^2 + c\alpha^2 c'\mathbb{E}|x_{t-1}|^4 + c\alpha^2 c'|\delta|^4 + c\beta^2\sigma_1^4 \\ &\quad \text{(applying again the c_n-inequality for some $c' \in \mathbb{R}$)} \\ &< \infty \\ &\quad \text{(because $c \in \mathbb{R}$, $c' \in \mathbb{R}$, $\omega \in \mathbb{R}$, $\alpha \in \mathbb{R}$, $\beta \in \mathbb{R}$,} \\ &\qquad \text{$\sigma_1^2 \in \mathbb{R}$, and $\mathbb{E}|x_t|^4 < \infty$)}\end{aligned}$$

- Condition A3 of the Power-n Theorem (the contraction condition) is satisfied for any $\boldsymbol{\theta}$ and $n = 2$ since

$$\sup_{\sigma^2}\left|\frac{\partial(\omega + \alpha(x_{t-1} - \delta)^2 + \beta\sigma^2)}{\partial\sigma^2}\right|^2 = \sup_{\sigma^2}|\beta|^2 = \beta^2,$$

  and hence, the degenerate random variable $\rho^2(x_t) = \beta^2$, which is trivially independent of σ_t^2, bounds the uniform derivative of interest, and satisfies the contraction condition

$$\mathbb{E}\rho^2(x_t) = \mathbb{E}\beta^2 = \beta^2 < 1 \quad \text{because } a < \beta < 1.$$

We thus conclude by application the Power-n Theorem, that the process $\{\sigma_t^2(\sigma_1^2, \boldsymbol{\theta})\}_{t\in\mathbb{N}}$ initialized at time $t = 1$ with any given value σ_1^2 converges to a unique limit process $\{\sigma_t^2(\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ which is SE and has two bounded moments $\mathbb{E}|\sigma_t^2(\boldsymbol{\theta})|^2 < \infty$.

PART II

Define again

$$Q_T(\mathbf{x}_T, \boldsymbol{\theta}) := \frac{1}{T}\sum_{t=2}^{T} q(x_t, \sigma_t^2(\sigma_1^2, \boldsymbol{\theta})) \quad \text{where}$$

$$q(x_t, \sigma_t^2(\sigma_1^2, \boldsymbol{\theta})) := -\frac{1}{2}\log 2\pi - \frac{1}{2}\log\sigma_t^2(\sigma_1^2, \boldsymbol{\theta}) - \frac{x_t^2}{2\sigma_t^2(\sigma_1^2, \boldsymbol{\theta})}.$$

By the classical consistency theorem for M-estimators, we can obtain the consistency of $\hat{\boldsymbol{\theta}}_T \xrightarrow{p} \boldsymbol{\theta}_0$ as $T \to \infty$ from the following three conditions:

C0. $\hat{\boldsymbol{\theta}}_T$ exists and is measurable.

C1. The uniform convergence of the criterion function

$$\sup_{\boldsymbol{\theta}\in\Theta}\left|\frac{1}{T}\sum_{t=2}^{T} q(x_t, \sigma_t^2(\sigma_1^2, \boldsymbol{\theta})) - \mathbb{E}q(x_t, \sigma_t^2(\boldsymbol{\theta}))\right| \xrightarrow{p} 0 \quad \text{as } T \to \infty.$$

C2. The identifiable uniqueness of $\boldsymbol{\theta}_0 \in \Theta$,

$$\sup_{\boldsymbol{\theta} \in S^c(\boldsymbol{\theta}_0,\delta)} \mathbb{E}q(x_t, \sigma_t^2(\boldsymbol{\theta})) < \mathbb{E}q(x_t, \sigma_t^2(\boldsymbol{\theta}_0)) \quad \text{for every } \delta > 0,$$

where $S^c(\boldsymbol{\theta}_0, \delta)$ denotes the complement of an open ball of radius δ centered at $\boldsymbol{\theta}_0$.

The existence and measurability of $\hat{\boldsymbol{\theta}}_T$ (condition C0) is ensured by the following conditions

C0.1 Compactness of the parameter space Θ.

C0.2 Continuity of the criterion function on the data $\mathbf{x}_T \in \mathbb{R}^T$ for every $\boldsymbol{\theta} \in \Theta$.

C0.3 Continuity of the criterion function on the parameter $\boldsymbol{\theta}$ for every data point $\mathbf{x}_T \in \mathbb{R}^T$).

The compactness of the parameter space Θ (condition C0.1) is given.

The continuity of the criterion on the data and the parameter (conditions C0.2 and C0.3) hold since the Gaussian density is continuous for $\sigma_t^2 > a > 0$ and since the updating equation for σ_t^2 is continuous on the data and the parameters.

The uniform convergence of the criterion function (Condition C1) is implied by the pointwise convergence of the log likelihood

C1.1. $\frac{1}{T}\sum_{t=2}^T q(x_t, \sigma_t^2(\sigma_1^2, \boldsymbol{\theta})) \overset{p}{\to} \mathbb{E}q(x_t, \sigma_t^2(\boldsymbol{\theta}))\ \forall\ \boldsymbol{\theta} \in \Theta$ as $T \to \infty$,

and the stochastic equicontinuity of the loglikelihood

C1.2. $\sup_t \mathbb{E} \sup_{\boldsymbol{\theta}\in\Theta} \left\| \nabla q(x_t, \sigma_t^2(\sigma_1^2, \boldsymbol{\theta})) \right\| < \infty$.

Since q is uniformly continuous and the filter $\{\sigma_t^2(\sigma_1^2, \boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ converges to a limit SE sequence $\{\sigma_t^2(\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ for every $\boldsymbol{\theta} \in \Theta$ (see Part 1 of proof above),

$$|\sigma_t^2(\sigma_1^2, \boldsymbol{\theta}) - \sigma_t^2(\boldsymbol{\theta})| \overset{p}{\to} 0 \quad \text{as } T \to \infty\ ,$$

the pointwise convergence of the log likelihood can be obtained instead by the application of a LLN for every $\boldsymbol{\theta} \in \Theta$ to the log likelihood evaluated at the limit SE filter $\{\sigma_t^2(\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$,

$$\frac{1}{T}\sum_{t=2}^T q(x_t, \sigma_t^2(\boldsymbol{\theta})) \overset{p}{\to} \mathbb{E}q(x_t, \sigma_t^2(\boldsymbol{\theta}))\ \forall\ \boldsymbol{\theta} \in \Theta \ \text{ as } T \to \infty.$$

We verify that this LLN holds by noting that $\{q(x_t, \sigma_t^2(\boldsymbol{\theta}))\}_{t\in\mathbb{N}}$ is SE for every $\boldsymbol{\theta} \in \Theta$, by Krengel's Theorem, because q is continuous (and hence Borel-measurable) function of the SE sequences $\{x_t\}_{t\in\mathbb{Z}}$ and $\{\sigma_t^2(\boldsymbol{\theta})\}_{t\in\mathbb{Z}}$ for every $\boldsymbol{\theta}\in\Theta$. Furthermore, it is easy to show that $q(x_t,\sigma_t^2(\boldsymbol{\theta}))$ has two bounded moments for every $\boldsymbol{\theta}\in\Theta$ since

$$
\begin{aligned}
\mathbb{E}|q(x_t,\sigma_t^2(\boldsymbol{\theta}))|^2 &= \mathbb{E}\Big| -\frac{1}{2}\log 2\pi - \frac{1}{2}\log\sigma_t^2(\boldsymbol{\theta}) - \frac{x_t^2}{2\sigma_t^2(\boldsymbol{\theta})}\Big|^2 \\
&\quad \text{(by the definition of } q(x_t,\sigma_t^2(\boldsymbol{\theta}))\text{)} \\
&\le A + \frac{1}{4}c\mathbb{E}|\log\sigma_t^2(\boldsymbol{\theta})|^2 + c\mathbb{E}\Big|\frac{x_t^2}{2\sigma_t^2(\boldsymbol{\theta})}\Big|^2 \\
&\quad \text{(by the } c_n\text{-inequality and defining } A := c(\tfrac{1}{2}\log 2\pi)^2\text{)} \\
&\le A + \frac{c}{4}\mathbb{E}|\sigma_t^2(\boldsymbol{\theta})|^2 + c\mathbb{E}\Big|\frac{x_t^2}{2a}\Big|^2 \\
&\quad \text{(since } \mathbb{E}|\log(z_t)|^n<\infty \Leftarrow \mathbb{E}|z_t|^n<\infty \text{ when } z_t\ge a>0 \text{ a.s.} \\
&\qquad \text{and because } \sigma_t^2(\boldsymbol{\theta})>a \text{ a.s. implies that } \frac{x_t^2}{2\sigma_t^2(\boldsymbol{\theta})}\le\frac{x_t^2}{2a} \text{ a.s.)} \\
&\le A + \frac{c}{4}\mathbb{E}|\sigma_t^2(\boldsymbol{\theta})|^2 + \frac{c}{4a^2}\mathbb{E}|x_t|^4 \\
&\quad \text{(bring constant outside)} \\
&< \infty \\
&\quad \text{(because } \mathbb{E}|\sigma_t^2(\boldsymbol{\theta})|^2<\infty \text{ [Part 1 of proof]} \\
&\qquad \text{and } \mathbb{E}|x_t|^4<\infty \text{ [given] and } c\in\mathbb{R} \text{ and } A\in\mathbb{R}\text{)}
\end{aligned}
$$

We thus obtain the pointwise convergence (condition C1.1) by the uniform continuity of q, the convergence of the filter to a limit SE sequence, and the application of a LLN for every $\boldsymbol{\theta}\in\Theta$.

Since the conditions of the Power-n Theorem hold for $n=2$ [Part 1 of the proof], and q is continuously differentiable and well behaved of order 2, the second order moment bound $\mathbb{E}|q(x_t,\sigma_t^2(\sigma_1^2,\boldsymbol{\theta}))|^2<\infty$ obtained above for some $\boldsymbol{\theta}\in\Theta$ implies also that

$$\sup_t \mathbb{E}\sup_{\boldsymbol{\theta}\in\Theta}\Big\|\nabla q(x_t,\sigma_t^2(\sigma_1^2,\boldsymbol{\theta}))\Big\| < \infty.$$

and hence the stochastic equicontinuity condition (C1.2) is also satisfied.

Finally, <u>if</u> $\boldsymbol{\theta}_0\in\Theta$ is the unique maximizer of the limit criterion function

$$Q_\infty(\boldsymbol{\theta}) = \mathbb{E}q(x_t,\sigma_t^2(\sigma_1^2,\boldsymbol{\theta})) \;<\; \mathbb{E}q(x_t,\sigma_t^2(\boldsymbol{\theta})) = Q_\infty(\boldsymbol{\theta}_0) \;\; \forall\,\boldsymbol{\theta}\neq\boldsymbol{\theta}_0$$

then, given the compactness of Θ and the continuity of the limit criterion Q_∞ in $\boldsymbol{\theta}$ we obtain that $\boldsymbol{\theta}_0$ is the identifiably unique maximizer of Q_∞ (condition C.2).

Since conditions C.1 (uniform convergence) and C.2 (identifiable uniqueness) stated above hold, we can conclude that $\hat{\boldsymbol{\theta}}_T$ is consistent for $\boldsymbol{\theta}_0$, i.e. that $\hat{\boldsymbol{\theta}}_T \xrightarrow{p} \boldsymbol{\theta}_0$.