Econometric Theory, **18**, 2002, 541–545. Printed in the United States of America.

PROBLEMS AND SOLUTIONS

PROBLEM

02.2.1. *ARMA Representation of Squared Markov Switching Heteroskedastic Models*, proposed by Walter Distaso. Markov switching models have become very popular in macroeconomics, since they were first proposed by Hamilton (1989). More recently, this kind of model has been applied to describe the behavior of the volatility of financial time series (see, e.g., Francq and Zakoian, 2000).

Let the process $\{\Delta_t\}$ be described by a Markov chain with state space $\{0,1\}$ and transition probabilities given by

$$\Pr[\Delta_t = 1 | \Delta_{t-1} = 1] = p,$$

$$\Pr[\Delta_t = 0 | \Delta_{t-1} = 1] = 1 - p,$$

$$\Pr[\Delta_t = 0 | \Delta_{t-1} = 0] = q,$$

$$\Pr[\Delta_t = 1 | \Delta_{t-1} = 0] = 1 - q,$$

with $0 < p, q < 1$. Consider the simple case of a model

$$y_t = \varepsilon_t, \tag{1}$$

where

$$\varepsilon_t = \sigma_t Z_t, \quad \sigma_t^2 = \mu(\Delta_t), \tag{2}$$

$$\mu(\Delta_t) = \sum_{i=1}^{2} \mu_i 1_{\Delta_t = i-1}, \quad 0 < \mu_1 < \mu_2, \tag{3}$$

and $\{Z_t\}$ is i.i.d.(0,1) with an existing fourth moment and independent of $\{\Delta_t\}$.

Show that the process $\{\varepsilon_t^2\}$ admits an ARMA(1,1) representation (see Francq and Zakoian, 2000, Example 7, p. 700) of the form

$$\varepsilon_t^2 = \omega + \alpha \varepsilon_{t-1}^2 + u_t - \beta u_{t-1} \tag{4}$$

and express ω, α, β in terms of p, q, μ_1, μ_2, $E(Z_t^4)$.

REFERENCES

Francq, C. & J.M. Zakoian (2000) Estimating weak GARCH representations. *Econometric Theory* 16, 5, 692–728.

Hamilton, J.D. (1989) A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica* 57, 2, 357–384.

 0266-4666/02 $9.50

https://doi.org/10.1017/S0266466602002141 Published online by Cambridge University Press

SOLUTIONS

01.2.1. *A Determinantal Inequality—First solution*,[1] proposed by Tom Wasbeek and Jos de Berge. Let $C = A_d^{-1/2} A A_d^{-1/2}$; then the question is to show that $|C| \leq |I| = 1$, i.e., the determinant of a correlation matrix is less than (or equal to) unity. Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues of C; then

$$\ln|C| = \ln \prod_{i=1}^{n} \lambda_i = \sum_{i=1}^{n} \ln \lambda_i \leq \sum_{i=1}^{n} (\lambda_i - 1) = \operatorname{tr}(C) - n = 0,$$

which gives the desired result.

Second solution, proposed by Christian Kleiber. This proof uses the arithmetic-geometric mean (AGM) inequality. Let $\lambda_i(\cdot)$ denote the ith eigenvalue of a matrix. Define the diagonal matrix $A_d = \operatorname{diag}(a_{11}, \ldots, a_{nn})$. Then $A_d^{-1/2} A A_d^{-1/2}$ is also positive definite, with a diagonal of ones.

Now, by the AGM inequality,

$$n = \operatorname{tr}(A_d^{-1/2} A A_d^{-1/2}) = \operatorname{tr}(A A_d^{-1}) = \sum_{i=1}^{n} \lambda_i (A A_d^{-1})$$

$$\geq n \left(\prod_{i=1}^{n} \lambda_i (A A_d^{-1}) \right)^{1/n} = n (\det(A A_d^{-1}))^{1/n} = n \left(\det(A) \cdot \prod_{i=1}^{n} a_{ii}^{-1} \right)^{1/n},$$

which implies $\det(A) \leq \det(A_d)$.

Third solution, proposed by Christian Kleiber. This proof considers the problem as a constrained optimization problem: Because A is positive definite, we can write $A = \tilde{R}^\top \tilde{R}$, in particular $a_{jj} = \sum_{i=1}^{n} \tilde{r}_{ij}^2$. Consider the problem

$$\max |\det(R)| \qquad \text{s.t.} \sum_{i=1}^{n} r_{ij}^2 = a_{jj}, \quad \text{for } j = 1, \ldots, n.$$

The objective function is continuous and to be maximized over a compact set, so a solution exists. Let R be a solution. Thus, expanding $\det(R)$ by the jth column,

$$\det(R) = \sum_{i=1}^{n} r_{ij} c_{ij},$$

where $C = (c_{ij})_{i,j}$ is the matrix of (signed) cofactors of R. Clearly

$$|\det R| \geq \prod_{j=1}^{n} a_{jj}^{1/2} = \det(A_d^{1/2}) > 0,$$

https://doi.org/10.1017/S0266466602002141 Published online by Cambridge University Press

because $A_d^{1/2}$ is an element of the feasible region. Now the Cauchy–Schwarz inequality implies

$$|\det(R)| = \left|\sum_{i=1}^{n} r_{ij} c_{ij}\right| \leq \|r^{(j)}\| \cdot \|c^{(j)}\| = a_{jj}^{1/2} \cdot \|c^{(j)}\|,$$

where $r^{(j)} = (r_{1j}, \dots, r_{nj})^\top$ and $c^{(j)} = (c_{1j}, \dots, c_{nj})^\top$, with equality if and only if $r^{(j)} = \rho_j c^{(j)}$, for some $\rho_j \in \mathbb{R}$. Because R solves the maximum problem, $r^{(j)}$ is necessarily of this form. Now consider, for $k \neq j$,

$$r^{(k)\top} r^{(j)} = r^{(k)\top} \rho_j c^{(j)} = \rho_j \cdot \sum_{i=1}^{n} r_{ik} c_{ij},$$

which is seen to be an expansion, by column j, of the determinant of R with column j replaced by a copy of column k. Thus, $r^{(k)\top} r^{(j)} = 0$, for all $k \neq j$; hence R has orthogonal columns. This gives

$$\det(R^\top) \cdot \det(R) = \det(R^\top R) = \det((r^{(i)\top} r^{(j)})_{i,j}) = \det(A_d)$$

and, R being a solution of the maximum problem, $\det(A) = \det(\tilde{R}^\top \tilde{R}) \leq \det(R^\top R) = \det(A_d)$.

Fourth solution, proposed by Heinz Neudecker, the poser of the problem. Consider $X = A_d + \sum_{i<j} x_{ij}(E_{ij} + E_{ji})$, where x_{ij} is the (i,j) element of X and E_{ij} is a basis matrix, i.e., $E_{ij} = e_i e_j'$, where e_i is a unit vector with ith element equal to 1. Then $d|X| = |X| \operatorname{tr} X^{-1} dX = |X| \operatorname{tr} X^{-1} \sum_{i<j}(E_{ij} + E_{ji}) dx_{ij}$. Necessary for a maximum is

$$|X| \operatorname{tr} X^{-1}(E_{ij} + E_{ji}) = 0 \qquad \forall i < j.$$

As $|X| \neq 0$ this yields $x^{ij} = 0$, where x^{ij} is the (i,j) element of X^{-1}. Hence $x_{ij} = 0$. This implies that the stationary point is $X_0 = A_d$. Further at the stationary point

$$\begin{aligned} d^2|X| &= (d|X|) \operatorname{tr} X^{-1} dX + |X| \operatorname{tr}(dX^{-1}) dX \\ &= -|X| \operatorname{tr} X^{-1}(dX) X^{-1} dX < 0 \end{aligned}$$

as $|X| > 0$ and $\operatorname{tr} X^{-1}(dX) X^{-1} dX = (d \operatorname{vec} X)'(X^{-1} \otimes X^{-1}) d \operatorname{vec} X > 0$. This shows that a maximum has been found.

NOTE

1. Christian Kleiber also reported that an alternative proof based on majorization theory can be found in Marshall and Olkin (1979, p. 223).

REFERENCE

Marshall, A.W. & I. Olkin (1979) *Inequalities: Theory of Majorization and Its Applications*. Orlando, Florida: Academic Press.

01.2.2. *The R/S Statistics as a Unit Root Test*—Solution, proposed by Giuseppe Cavaliere.

Part (a). Because $\hat{\mu} = T^{-1}(X_T - X_0) = T^{-1}(\mu T + S_T) = \mu + T^{-1}S_T$, the equality $\sum_{i=1}^{t}(\Delta X_i - \hat{\mu}) = S_t - (t/T)S_T$ holds and we can follow Lo (1991, Theorem 3.1) to get $(\hat{\lambda}/\lambda)R/S \Rightarrow Z_1 := \sup_{s\in[0,1]} V_1(s) - \inf_{s\in[0,1]} V_1(s)$, where V_1 is the Brownian bridge $V_1(s) = B(s) - sB(1)$, B is a standard Brownian motion, and $\Rightarrow$ denotes weak convergence. Finally, R/S has the same asymptotic distribution provided that λ is bounded away from 0 and $\hat{\lambda}_T^2 - \lambda^2 = o_p(1)$. But this fact follows directly from the conditions on $\{u_t\}$ and $\hat{\lambda}_T^2$ (see de Jong, 2000).

Part (b). First, simple calculations allow us to state that $\sum_{i=1}^{t}(\Delta X_i - \hat{\mu}) = X_t - X_0 - \hat{\mu}t =: \hat{X}_t$, from which it follows that the numerator of R/S has the representation $\max_t \hat{X}_t - \min_t \hat{X}_t$. Under the alternative hypothesis, $\hat{X}_t$ is given by $\hat{X}_t = X_t - X_0 - (t/T)(X_T - X_0) = \sum_{i=1}^{t} S_i - t\bar{S}$, where $S_t := \sum_{i=1}^{t} u_i$ and $\bar{S} := T^{-1}\sum_{t=1}^{T} S_t$; note that $\hat{X}_t$ depends neither on μ nor on X_0. By standard $I(2)$ asymptotics and the continuous mapping theorem,

$$\frac{\hat{X}_{[sT]}}{\lambda T^{3/2}} = \frac{1}{\lambda T^{3/2}} \sum_{t=1}^{[sT]} S_t - \frac{[sT]}{\lambda T^{3/2}} \bar{S} \Rightarrow \int_0^s B(r)\,dr - s\int_0^1 B(s)\,ds =: V_2(s)$$

$$\frac{1}{\lambda T^{3/2}}\left(\max_t \hat{X}_t - \min_t \hat{X}_t\right) \Rightarrow Z_2 := \sup_{s\in[0,1]} V_2(s) - \inf_{s\in[0,1]} V_2(s). \qquad \textbf{(1)}$$

Now consider the denominator. Because X_t is $I(2)$, ΔX_t is $I(1)$, and consequently $\hat{\lambda}_T^2$ is based on the sample autocovariance function of the demeaned integrated process $\Delta X_t - \hat{\mu} = S_t - \bar{S}$. In this case Phillips (1991) has proved that

$$\frac{1}{\lambda^2}\frac{\hat{\lambda}_T^2}{Kq_T T} \Rightarrow Z := \int_0^1 (B(s) - \bar{B})^2 ds, \qquad K := \int_{\mathbb{R}} k(s)\,ds \qquad \textbf{(2)}$$

provided that q_T is $O(T^\gamma)$, $\gamma < \frac{1}{2}$. By combining (1) and (2) one gets $(q_T/T)^{1/2}R/S \Rightarrow Z_2/Z^{1/2}$. Finally, the condition $q_T/T \to 0$ implies that R/S diverges to $+\infty$ and a right tail test based on R/S is consistent against $I(2)$.

Part (c). Under the alternative hypothesis,

$$\hat{\mu} = \mu + (\mu_0 - \mu)(1 - \alpha) + T^{-1}S_T + c_T,$$

where $S_t := \sum_{i=1}^{t} u_i$ and $c_T := T^{-1}(\mu_0 - \mu)(\alpha T - [\alpha T])$; therefore

$$\hat{X}_t := \sum_{i=1}^{t} \Delta X_i - \hat{\mu}t = (\mu_0 - \mu)\{(t - [\alpha T])\mathbb{I}(t > [\alpha T]) - (1 - \alpha)t\}$$

$$+ S_t - (t/T)S_T - tc_T.$$

Because $T^{-1}(S_t - (t/T)S_T - tc_T) = o_p(1)$, all t,

$$T^{-1}\hat{X}_{[sT]} = (\mu_0 - \mu)\{(s-\alpha)\mathbb{I}(s > \alpha) - (1-\alpha)s\} + o_p(1).$$

Hence, $(1/T)$ times the numerator of the R/S statistics, i.e., the range of $T^{-1}\hat{X}_{[sT]}$, $s \in [0,1]$, converges in probability to the range of $(\mu_0 - \mu)((s-\alpha) \times \mathbb{I}(s > \alpha) - s(1-\alpha))$, which is given by $|\mu_0 - \mu|\alpha(1-\alpha)$; note that in the case of no trend breaks $T^{-1}(\max_t \hat{X}_t - \min_t \hat{X}_t) \xrightarrow{p} 0$.

Finally, consider the behavior of the long-run variance estimator. Using arguments similar to those of Phillips (1991) (see part (b) in the preceding discussion), it can be proved that

$$\frac{\hat{\lambda}_T^2}{Kq_T} = \kappa + o_p(1), \qquad \kappa := \alpha(1-\alpha)(\mu_0 - \mu)^2 > 0$$

for $\mu_0 \neq \mu$. By combining this result with the probability limit of the numerator it follows that R/S diverges to $+\infty$ and a right tail test based on R/S is consistent against trend breaks.

REFERENCES

de Jong, R.M. (2000) A strong consistency proof for heteroskedasticity and autocorrelation consistent covariance matrix estimators. *Econometric Theory* 16, 262–268.

Lo, A.W. (1991) Long-term memory in stock market prices. *Econometrica* 59, 1279–1313.

Phillips, P.C.B. (1991) Spectral regression for cointegrated time series. In W.A. Barnett, J. Powell, & G.E. Tauchen (eds.), *Nonparametric and Semiparametric Methods in Economics and Statistics*. Cambridge: Cambridge University Press.

https://doi.org/10.1017/S0266466602002141 Published online by Cambridge University Press