浙江科技学院学报,第27卷第1期,2015年2月
Journal of Zhejiang University of Science and Technology
Vol. 27 No. 1, Feb. 2015

doi: 10.3969/j.issn.1671-8798.2015.01.002

基于ESTAR模型的单位根检验
——Wald统计量的研究

胡俊娟
(浙江科技学院 理学院,杭州 310023)

摘　要: 已有的基于ESTAR模型的单位根检验通常假设$\alpha=0$,但实践研究表明α显著。对$\alpha\neq0$的情况进行单位根检验,研究该情形下的Wald检验统计量。从理论上推导了Wald检验统计量的极限分布,并通过Monte Carlo随机模拟,结果表明该统计量有效。

关键词: 单位根检验;Wald统计量;ESTAR模型
中图分类号: O212;F222　　**文献标志码:** A　　**文章编号:** 1671-8798(2015)01-0006-05

Unit root test based on ESTAR process
—Study on Wald statistics

HU Junjuan
(School of Sciences, Zhejiang University of Science and Technology, Hangzhou 310023, China)

Abstract: The alternative hypothesis of exponential smooth transition autoregressive (ESTAR) nonlinearity usually assumes that $\alpha=0$, while the Existing tests of the unit root hypothesis is against that. However, empirical work indicates that the estimated value of α is always greater than zero and significant. Hence, the paper relaxes the restriction in the test regression and investigate the Wald-type test for a unit root process against ESTAR process. The asymptotic distributions of the test statistic are derived. Results show that the Wald statistic is effective via Monte Carlo simulation.

Key words: unit root test; Wald-type statistic; exponential smooth transition autoregressive

自Teräsvirta[1]提出STAR(smooth transition autoregressive)模型以来,非线性模型一直受到国内外学者的关注。STAR模型很好地描述了模型的状态转化和结构变化,因而被广泛地应用于预测工业产出、实际汇率、失业率等主要宏观时间序列[2-3]。由于模型拟合前需要进行单位根检验,众多学者对指数STAR(ESTAR)模型进行了研究。Kapetanios等[4]提出了Dickey-Fuller型t统计量,Kruse[5]用修正的

收稿日期: 2015-01-21
作者简介: 胡俊娟(1979—　),女,浙江省兰溪人,讲师,博士研究生,主要从事时间序列分析研究。

t 统计量进行了研究，Hanck[6]对带均值的情况进行了研究。这些检验都在假设 $\alpha=0$ 的情况下进行，但实践研究表明 α 显著[7]。笔者在此基础上放宽了条件，在 $\alpha\neq 0$ 的情况下进行单位根检验的研究。

1　基于 ESTAR 模型的单位根检验

ESTAR 模型是 STAR 类模型中最常用的模型之一，首先给出 ESTAR 模型的具体形式：

$$y_t=\beta y_{t-1}+\gamma y_{t-1}G(y_{t-1};\theta)+\varepsilon_t,t=1,\cdots,T, \tag{1}$$

式(1)中：$G(y_{t-1};\theta)=1-\exp\{-\theta y_{t-1}^2\}$；$\varepsilon_t$—独立同分布且均值为零，方差是 σ^2；T—样本量。

该过程也可以写成：

$$\Delta y_t=\alpha y_{t-1}+\gamma y_{t-1}G(y_{t-1};\theta)+\varepsilon_t, \tag{2}$$

式(2)中：$\alpha=\beta-1$，Δ—差分算子。

1.1　Dickey-Fuller 型 t 统计量

Kapetanios 等[4]在假设 $\alpha=0$ 的条件下，考虑 θ 是否为零。即对于方程：

$$\Delta y_t=\gamma y_{t-1}G(y_{t-1};\theta)+\varepsilon_t$$

给出了原假设 $H_0:\theta=0$，备择假设 $H_1:\theta>0$ 来进行单位根检验。考虑到参数 γ 未知，常用的做法是用一阶泰勒展开来逼近非线性函数 $G(y_{t-1};\theta)$，则检验方程就变成：

$$\Delta y_t=\delta y_{t-1}^3+\text{误差项},$$

从而对方程检验 $H_0:\delta=0$。Kapetanios 等给出了 Dickey-Fuller 型 t 统计量：

$$\mathrm{KSS}=\frac{\hat{\delta}}{t(\hat{\delta})}, \tag{3}$$

式(3)中：$\hat{\delta}$—δ 的最小二乘估计量；$t(\hat{\delta})$—$\hat{\delta}$ 的标准差。

1.2　Wald 检验统计量

考虑到 ESTAR 模型(式(1))中 α 不一定为零，则对 ESTAR 模型进行一阶泰勒展开后得到的辅助线性方程是：

$$\Delta y_t=\alpha y_{t-1}+\delta y_{t-1}^3+\text{误差项}。 \tag{4}$$

考虑到当 $\alpha=\delta=0$ 时，模型含有单位根，所以对模型检验：$H_0:\alpha=\delta=0$。

Wald 检验是计量经济学三大检验方法之一，只要求计算无约束模型，在实证分析中得到广泛应用。在原假设 $H_0:\boldsymbol{R\beta}=0$ 情况下的 Wald 统计量[8]为：

$$W=(\hat{\beta}-\boldsymbol{\beta})'R'(\boldsymbol{R}S_T^2(X'X)^{-1}R')(\hat{\beta}-\boldsymbol{\beta}), \tag{5}$$

式(5)中：$\boldsymbol{R}$—限制参数矩阵；$\boldsymbol{\beta}$—未知参数向量；$\hat{\beta}$—$\boldsymbol{\beta}$ 的估计量；$S_T^2(X'X)^{-1}$—参数的方差-协方差矩阵；S_T^2—误差的方差估计量。对于 ESTAR 模型(式(1))，则

$$\boldsymbol{R}=\begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix},\boldsymbol{\beta}=(\alpha,\delta)',X=y_{t-1},y_{t-1}^3{}'。$$

聂巧平等[9]研究了 ADF 检验中联合检验 F 统计量的极限分布，发现它们并不是标准的 F 分布，而是 Wiener 过程的泛函。以上研究表明，在原假设为单位根的检验条件下，检验回归式的各系数估计量及其相关的统计量的极限分布并不是常规的正态分布或 F 分布，这自然会让人想到在单位根检验中，检验参数约束是否正确的 Wald 统计量是否服从 χ^2 分布。在单位根检验的回归模型估计中，$\sqrt{T}(\hat{\beta}-\boldsymbol{\beta})$ 收敛的条件不再满足，因此参数约束的 Wald 统计量的极限分布将不再服从 χ^2 分布。笔者对此进行尝试，推导出在参数约束下 Wald 统计量的极限分布。

2　Wald 统计量的极限分布

为了给出统计量 W 的极限分布，先假设序列 y_t 的六阶矩存在。有关 W 极限分布的定理如下。

定理 1　考虑模型(2)在 $H_0:\alpha=\delta=0$ 的条件下，统计量 W 有如下的极限分布：

$$W \Rightarrow A'B^{-1}A,$$

式中：$A=\begin{pmatrix} \frac{1}{2}(w^2(1)-1) \\ \frac{1}{4}w^4(1)-\frac{3}{2}\int_0^1 w^2(r)\mathrm{d}r \end{pmatrix}$；$B=\begin{pmatrix} \int_0^1 w^2(r)\mathrm{d}r & \int_0^1 w^4(r)\mathrm{d}r \\ \int_0^1 w^4(r)\mathrm{d}r & \int_0^1 w^6(r)\mathrm{d}r \end{pmatrix}$；$w(r)$— 标准的布朗运动；“$\Rightarrow$”—弱收敛。

证明：为了标记方便，将 $\sum_{t=1}^{T}$ 记作 $\sum$。对于辅助线性方程(4)，可以得到

$$\hat{\beta}=\begin{pmatrix}\hat{\alpha}\\ \hat{\delta}\end{pmatrix}=\begin{pmatrix}\sum y_{t-1}\varepsilon_t \\ \sum y_{t-1}^3\varepsilon_t\end{pmatrix}\begin{pmatrix}\sum y_{t-1}^2 & \sum y_{t-1}^4 \\ \sum y_{t-1}^4 & \sum y_{t-1}^6\end{pmatrix}^{-1}。$$

考虑模型(2) 在 $H_0:\alpha=\delta=0$ 的条件下(即模型(2) 是单位根过程)，对于统计量 W，则有：

$$W=(\hat{\beta}-\boldsymbol{\beta})'R'(RS_T^2(X'X)^{-1}R')(\hat{\beta}-\boldsymbol{\beta})=$$

$$\begin{pmatrix}\hat{\alpha}\\ \hat{\delta}\end{pmatrix}'\begin{pmatrix}1&0\\0&1\end{pmatrix}\left(\begin{pmatrix}1&0\\0&1\end{pmatrix}S_T^2\begin{pmatrix}\sum y_{t-1}^2 & \sum y_{t-1}^4 \\ \sum y_{t-1}^4 & \sum y_{t-1}^6\end{pmatrix}^{-1}\begin{pmatrix}1&0\\0&1\end{pmatrix}\right)^{-1}\begin{pmatrix}1&0\\0&1\end{pmatrix}\begin{pmatrix}\hat{\alpha}\\ \hat{\delta}\end{pmatrix}=$$

$$\frac{1}{S_T^2}\begin{pmatrix}\sum y_{t-1}\varepsilon_t \\ \sum y_{t-1}^3\varepsilon_t\end{pmatrix}'\begin{pmatrix}\sum y_{t-1}^2 & \sum y_{t-1}^4 \\ \sum y_{t-1}^4 & \sum y_{t-1}^6\end{pmatrix}^{-1}\begin{pmatrix}\sum y_{t-1}\varepsilon_t \\ \sum y_{t-1}^3\varepsilon_t\end{pmatrix}=$$

$$\frac{1}{S_T^2}\begin{pmatrix}\sum y_{t-1}\varepsilon_t \\ \sum y_{t-1}^3\varepsilon_t\end{pmatrix}'\begin{pmatrix}T^{-1}&0\\0&T^{-2}\end{pmatrix}\begin{pmatrix}T^{-1}&0\\0&T^{-2}\end{pmatrix}^{-1}\begin{pmatrix}\sum y_{t-1}^2 & \sum y_{t-1}^4 \\ \sum y_{t-1}^4 & \sum y_{t-1}^6\end{pmatrix}^{-1}\begin{pmatrix}T^{-1}&0\\0&T^{-2}\end{pmatrix}^{-1}\begin{pmatrix}T^{-1}&0\\0&T^{-2}\end{pmatrix}\begin{pmatrix}\sum y_{t-1}\varepsilon_t \\ \sum y_{t-1}^3\varepsilon_t\end{pmatrix}=$$

$$\frac{1}{S_T^2}\begin{pmatrix}T^{-1}\sum y_{t-1}\varepsilon_t \\ T^{-2}\sum y_{t-1}^3\varepsilon_t\end{pmatrix}'\begin{pmatrix}T^{-2}\sum y_{t-1}^2 & T^{-3}\sum y_{t-1}^4 \\ T^{-3}\sum y_{t-1}^4 & T^{-4}\sum y_{t-1}^6\end{pmatrix}^{-1}\begin{pmatrix}T^{-1}\sum y_{t-1}\varepsilon_t \\ T^{-2}\sum y_{t-1}^3\varepsilon_t\end{pmatrix},$$

式中：S_T^2 为误差项的最小二乘估计量，易证 $S_T^2 \xrightarrow{P} \sigma^2$。由连续映射定理和随机积分的弱收敛定理，则有：

$$\frac{1}{T}\sum\left(\frac{y_{t-1}}{\sqrt{T}}\right)^i \Rightarrow \sigma^i\int_0^1 w^i(r)\mathrm{d}r,\ i=2,4,6。$$

即

$$\frac{1}{T^{\frac{i}{2}+1}}\sum y_{t-1}^i \Rightarrow \sigma^i\int_0^1 w^i(r)\mathrm{d}r,\ i=2,4,6。$$

进一步，有：

$$T^{-2}\sum y_{t-1}^3\varepsilon_t \Rightarrow \sigma^4\int_0^1 w^3(r)\mathrm{d}w(r)=\sigma^4\left(\frac{1}{4}w^4(1)-\frac{3}{2}\int_0^1 w^2(r)\mathrm{d}r\right),$$

并且

$$T^{-1}\sum y_{t-1}\varepsilon_t \Rightarrow \sigma^2\int_0^1 w(r)\mathrm{d}w(r)=\sigma^2\left(\frac{1}{2}w^2(1)-1\right)。$$

所以，可以得到：

$$W\Rightarrow \frac{1}{\sigma^2}\begin{pmatrix}\frac{1}{2}\sigma^2(w^2(1)-1) \\ \sigma^4\left(\frac{1}{4}w^4(1)-\frac{3}{2}\int_0^1 w^2(r)\mathrm{d}r\right)\end{pmatrix}'\begin{pmatrix}\sigma^2\int_0^1 w^2(r)\mathrm{d}r & \sigma^4\int_0^1 w^4(r)\mathrm{d}r \\ \sigma^4\int_0^1 w^4(r)\mathrm{d}r & \sigma^6\int_0^1 w^6(r)\mathrm{d}r\end{pmatrix}^{-1}\begin{pmatrix}\frac{1}{2}\sigma^2(w^2(1)-1) \\ \sigma^4\left(\frac{1}{4}w^4(1)-\frac{3}{2}\int_0^1 w^2(r)\mathrm{d}r\right)\end{pmatrix}=$$

$$\frac{1}{\sigma^2}\left(Q'\begin{pmatrix}\frac{1}{2}\sigma^2(w^2(1)-1) \\ \sigma^4\left(\frac{1}{4}w^4(1)-\frac{3}{2}\int_0^1 w^2(r)\mathrm{d}r\right)\end{pmatrix}\right)'\left(Q\begin{pmatrix}\sigma^2\int_0^1 w^2(r)\mathrm{d}r & \sigma^4\int_0^1 w^4(r)\mathrm{d}r \\ \sigma^4\int_0^1 w^4(r)\mathrm{d}r & \sigma^6\int_0^1 w^6(r)\mathrm{d}r\end{pmatrix}Q\right)^{-1}\left(Q\begin{pmatrix}\frac{1}{2}\sigma^2(w^2(1)-1) \\ \sigma^4\left(\frac{1}{4}w^4(1)-\frac{3}{2}\int_0^1 w^2(r)\mathrm{d}r\right)\end{pmatrix}\right),$$

式中，矩阵 $\boldsymbol{Q} = \begin{pmatrix} \sigma^{-2} & 0 \\ 0 & \sigma^{-4} \end{pmatrix}$，因此**定理 1** 成立。

现对模型(2)进行推广，考虑误差项 ε_t 序列相关。根据 ADF 检验，通过加入 p 阶滞后项来代替误差项的序列相关，则模型(2)变成

$$\Delta y_t = \alpha y_{t-1} + \gamma y_{t-1} G(y_{t-1};\theta) + \sum_{j=1}^{p} \Delta y_{t-j} + \varepsilon_t, \tag{6}$$

式(6)中：$\varepsilon_t \sim \text{i. i. d.}(0, \sigma^2)$。则相应地可以得到辅助方程：

$$\Delta y_t = \alpha y_{t-1} + \delta y_{t-1}^3 + \sum_{j=1}^{p} \Delta y_{t-j} + \text{误差项}。$$

对于统计量 W 有如下分布：

定理 2　考虑模型(6)在 $H_0: \alpha = \delta = 0$ 的条件下，统计量 W 有与**定理 1** 一样的极限分布。

证明：设 $T \times p$ 的数据矩阵 $\boldsymbol{Z} = (\Delta Y_{-1}, \cdots, \Delta Y_{-p})$，其中 $\Delta Y_{-i} = (\Delta y_{-i+1}, \cdots, \Delta y_{-i+T})$，设 $T \times T$ 矩阵 $\boldsymbol{M} = I_T - Z(Z'Z)^{-1}Z'$，向量 $\boldsymbol{\varepsilon} = \varepsilon_1, \cdots, \varepsilon_T{}'$，则

$$S_T^2 = \frac{1}{T}\varepsilon' M \boldsymbol{\varepsilon} = \frac{1}{T}\varepsilon' \boldsymbol{\varepsilon} + o_P(1) \xrightarrow{P} \sigma^2。$$

设 $Y_{-1} = (y_0, \cdots, y_{T-1})'$ 和 $Y_{-1}^i = (y_0^i, \cdots, y_{T-1}^i)'$，$i = 1,2,3$。进一步，有：

$$T^{-2} Y^{3\prime}{}_{-1} \boldsymbol{M\varepsilon} = T^{-2} Y^{3\prime}{}_{-1} \boldsymbol{\varepsilon} + o_P(1) \Rightarrow \sigma^4 \left(\frac{1}{4} w^4(1) - \frac{3}{2} \int_0^1 w^2(r)\,\mathrm{d}r \right),$$

$$T^{-2} Y'_{-1} \boldsymbol{M\varepsilon} = T^{-2} Y'_{-1} \boldsymbol{\varepsilon} + o_P(1) \Rightarrow \sigma^2 \left(\frac{1}{2} w^2(1) - 1 \right),$$

和

$$\frac{1}{T^{i+1}} Y^{i\prime}{}_{-1} \boldsymbol{M} Y_{-1}^i = \frac{1}{T^{i+1}} Y^{i\prime}{}_{-1} Y_{-1}^i + o_P(1) \Rightarrow \sigma^{2i} \int_0^1 w^{2i}(r)\,\mathrm{d}r, i = 1,2,3。$$

所以，类似地可以得到 $W \Rightarrow A'B^{-1}A$，

其中，$A = \begin{pmatrix} \frac{1}{2}(w^2(1) - 1) \\ \frac{1}{4} w^4(1) - \frac{3}{2}\int_0^1 w^2(r)\,\mathrm{d}r \end{pmatrix}$，$B = \begin{pmatrix} \int_0^1 w^2(r)\,\mathrm{d}r & \int_0^1 w^4(r)\,\mathrm{d}r \\ \int_0^1 w^4(r)\,\mathrm{d}r & \int_0^1 w^6(r)\,\mathrm{d}r \end{pmatrix}$。

3　Monte Carlo 模拟

上述分析表明，在单位根检验下，参数约束的 Wald 统计量的极限分布都不是 χ^2 分布。本研究采用蒙特卡罗模拟方法模拟出上述 Wald 统计量的临界值。数据生成过程为：$\Delta y_t = \varepsilon_t$，$y_0 = 0$，$\varepsilon_t \sim \text{i. i. d.}(0,1)$。考察统计量 W 的极限分布，样本容量 T 取 1 000，针对样本容量模拟 20 000次，从而得到检验统计量 W 常用检验水平 0.01、0.05、0.10 的渐进临界值(表 1)。

表 1　统计量 W 的临界值

Table 1　Critical values of statistic W

检验水平/%	W 的渐进临界值
10	6.63
5	8.51
1	12.23

对于小样本而言，分别针对样本容量 T 为 50、100、200，考察检验统计量的检验水平。数据生成过程为：$\Delta y_t = \varepsilon_t$，$\varepsilon_t = \rho\varepsilon_{t-1} + u_t$，$u_t \sim \text{i. i. d.}(0,1)$。为了考察更多情况，$\rho$ 取 0～0.4 的均匀分布。通过表 2 可以看出，统计量 W 的实际检验水平与名义检验水平一致。

表 2　统计量 W 的实际检验水平

Table 2　Size of statistic W

名义水平/%	T=50	T=100	T=200
10	0.102 6	0.103 6	0.103 2
5	0.049 1	0.049 9	0.051 7
1	0.011 0	0.011 2	0.011 4

考虑数据生成过程为 ESTAR 模型，现对式(3)的统计量 KSS 和统计量 W 的势进行对比。不妨假设数据生成过程为：

$$\Delta y_t = \alpha y_{t-1} + \gamma y_{t-1}(1-\exp\{-\theta y_{t-1}^2\}) + \varepsilon_t,$$

式中,α 设为 0.1,γ 为−1.5、−1、−0.5,θ 为 0.001～0.01 的均匀分布。给定名义水平为 0.05,从表 3 可以看出,统计量 W 较统计量 KSS 有更高的势。

表 3 统计量 W 和统计量 KSS 的势

Table 3 Power of statistic W and KSS

参数 γ	T=50		T=100		T=200	
	W	KSS	W	KSS	W	KSS
−1.5	0.301 7	0.151 9	0.675 1	0.586 4	0.979 5	0.877 2
−1.0	0.281 2	0.063 2	0.620 3	0.382 9	0.968 7	0.802 1
−0.5	0.228 2	0.007 3	0.594 6	0.063 1	0.940 1	0.545 3

4 结 语

本研究主要针对 ESTAR 模型提出用于单位根检验的 Wald 统计量,通过研究得出以下结论:无论是否有 p 阶滞后项,含有单位根约束条件的 Wald 统计量的极限分布均不服从 χ^2 分布,而是服从以Wiener 过程泛函表示的分布;蒙特卡罗模拟结果显示,有限样本下 Wald 统计量 W 的势比 KSS 的势更高,故认为统计量 W 更有效。

参考文献:

[1] Teräsvirta T. Specification, estimation and evaluation of smooth transition autoregressive models[J]. Journal of the American Statistical Association,1994,89:208-218.

[2] Lundbergh S, Teräsvirta T. Forecasting with smooth transition autoregressive models[M]. [S. L.]: Clements M P, 2001.

[3] Van Dijk D, Teräsvirta T, Franses P H. Smooth transition autoregressive models: A survey of recent developments [J]. Econometric Reviews,2002,21:1-47.

[4] Kapetanios G, Shin Y, Snell A. Testing for a unit root in the nonlinear STAR framework[J]. Journal of Econometrics,2003,112:359-379.

[5] Kruse R. A new unit root test against ESTAR based on a class of modified statistics[J]. Statistical Papers,2011,52(1):71-85.

[6] Hanck C. On the asymptotic distribution of a unit root test against ESTAR alternatives[J]. Statistics and Probability Letters,2012,82:360-364.

[7] Ubilava D, Helmers C G. Forecasting ENSO with a smooth transition autoregressive model[J]. Environmental Modelling and Software,2013,40:181-190.

[8] 张凌翔,张晓峒.单位根检验中的 Wald 统计量研究[J].数量经济技术经济研究,2009(7):146-158.

[9] 聂巧平,张晓峒.ADF 单位根检验中联合检验 F 统计量研究[J].统计研究,2007,24(2):73-80.