

Determining the existence of unit roots based on detrended data

Okyoung Na^{1,a}

^aDepartment of Applied Statistics, Kyonggi University

Abstract

In this paper, we study a method to determine the existence of unit roots by using the adaptive lasso. The previously proposed method that applied the adaptive lasso to the original time series has low power when there is an unknown trend. Therefore, we propose a modified version that fits the ADF regression model without deterministic component using the adaptive lasso to the detrended series instead of the original series. Our Monte Carlo simulation experiments show that the modified method improves the power over the original method and works well in large samples.

Keywords: unit root test, adaptive lasso, detrending

1. 서론

단위근의 존재 여부에 따라 시계열의 특성 및 정상화 방법, 고려할 수 있는 시계열 모형, 예측 등이 달라지므로, 시계열 자료를 분석할 때 단위근의 존재를 파악하는 것은 매우 중요하다. 보통 단위근이 존재하는지 객관적으로 판단하기 위하여 단위근 검정을 사용하는데, Dickey와 Fuller (1979)가 DF 검정을 제안한 이후로 다양한 형태의 단위근 검정들과 전략들이 제안되었으며, 이 중에서 Said와 Dickey (1984)의 ADF 검정과 Phillips와 Perron (1988)의 PP 검정이 대표적이다. 그러나 이 검정들은 단위근 근처에서 검정력이 낮다는 문제점을 가지고 있으며, 이를 극복하기 위해 다양한 방법들이 연구되었다. 예를 들면, Kwiatkowski 등 (1992)은 귀무가설과 대립가설을 기존의 단위근 검정과 반대로 설정하여 새롭게 만든 KPSS 검정을 제안하였고, Elliott 등 (1996), Ng와 Perron (2001) 등은 원 시계열 대신 추세를 제거한 시계열을 사용하여 단위근 검정을 실시하는 방법을 연구하였으며, Enders (2010)는 ADF 검정을 반복적으로 실시하는 전략을 제시하였다.

최근 Na (2019, 2020)는 별점화 추정방법 중 Zou (2006)가 제안한 adaptive lasso를 사용하여 단위근을 식별하는 방법을 제안하였다. 특히 Na (2020)는 ADF 검정에서 고려하는 세 가지 형태의 회귀모형식 중 선형 추세가 포함된 회귀모형식을 adaptive lasso로 추정하였고, 계수에 대한 추정값을 바탕으로 주어진 시계열이 단위근을 가지는지 판단하였다. Na (2020)의 모의실험 결과를 보면, 선형 추세의 기울기가 0인 경우에는 단위근의 존재 여부를 정확하게 판단한 비율이 높다. 그러나 기울기가 0이 아닌 경우에는 단위근이 존재하지 않음에도 불구하고 단위근이 존재한다고 잘못 판단한 비율이 다소 높게 나타났으며, 이는 단위근 검정의 입장에서 볼 때 기울기가 0이 아닌 선형 추세가 존재하는 경우 검정력이 떨어지는 것으로 해석할 수 있다. 이와

¹ Department of Applied Statistics, Kyonggi University, 154-42, Gwanggyosan-Ro, Yeongtong-Gu, Suwon-Si, Gyeonggi-Do 16227, Korea. E-mail: okna@kyonggi.ac.kr

같은 문제점은 ADF 검정에서도 발견되었으며, Elliott 등 (1996)은 ADF 검정에 대한 대안으로 DF-GLS 검정을 제안하였다. DF-GLS 검정은 일반화최소제곱법으로 추세를 추정하여 제거한 후 ADF 검정을 실시하는 방법으로 추세를 제거하지 않았을 때에 비해 검정력이 향상되는 것으로 나타났다. 이에 대한 자세한 내용은 Elliott 등 (1996)을 참고하기 바란다. 본 논문에서도 Elliott 등 (1996)처럼 추세를 제거한 시계열을 가지고 adaptive lasso를 이용하여 ADF-회귀모형식을 추정하는 방법에 대해 연구하였으며, 모의실험을 통해 Na (2020)처럼 원시계열을 이용한 경우에 비해 검정력이 향상되는지 살펴보려 한다.

본 논문에서는 시계열 $\{y_t, t \in \mathbb{N}\}$ 가 다음 모형식을 만족한다고 가정한다:

$$y_t = \beta_0 + \beta_1 t + u_t, \quad \beta_0, \beta_1 \in \mathbb{R}; \quad (1.1)$$

$$u_t = \alpha u_{t-1} + v_t, \quad u_0 = 0, -1 < \alpha \leq 1, \quad (1.2)$$

여기서 β_0, β_1, α 는 미지의 모수이고, $\{v_t, t \in \mathbb{N}\}$ 은 정상시계열이다. 이 모형은 Elliott 등 (1996), Ng와 Perron (2001), Vougas (2007), Na (2020) 등을 포함한 많은 연구에서 고려되었으며, $\alpha = 1$ 인지 여부를 판단하는 것이 주된 관심사이다. 위 모형에는 세 개의 시계열이 존재하는데, 이 중에서 $\{u_t\}$ 와 $\{v_t\}$ 는 관측불가능하며, $\{y_t\}$ 만 관측가능하다. Na (2020)는 관측값 y_1, y_2, \dots, y_n 을 가지고 adaptive lasso를 이용하여 ADF-회귀모형식을 추정하는 방식으로 $\alpha = 1$, 즉 단위근이 존재하는지를 알아보았으나, 본 논문에서는 앞서 이야기한 것처럼 y_t 에서 추세 $d_t = \beta_0 + \beta_1 t$ 를 추정하여 제거한 값

$$y_t^d = y_t - \hat{d}_t = y_t - (\hat{\beta}_0 + \hat{\beta}_1 t), \quad t = 1, 2, \dots, n \quad (1.3)$$

을 사용하여 ADF-회귀모형식을 추정하고자 하며 이에 대한 세부절차는 2장에서 자세하게 설명할 것이다. 그리고 본 논문에서 제안한 방법과 Elliott 등 (1996)의 DF-GLS 검정, Na (2020)가 제안한 방법을 비교하는 모의실험을 진행하고, 그 결과를 3장에 기술할 예정이다.

2. 단위근의 식별

1장에서 간단하게 기술한 것처럼 본 논문에서는 단위근을 식별하기 위해 원 시계열이 아닌 (1) 추세를 제거한 시계열을 가지고, (2) 결정적 추세가 포함되지 않은 ADF-회귀모형식을 adaptive lasso 방법으로 추정한 결과를 사용한다. 즉, 추세 제거와 adaptive lasso 추정이 본 논문에서 제안한 수정안의 핵심 원리이다. 그러므로 두 가지 원리를 각각 적용하여 단위근의 존재를 파악했던 기존의 방법부터 살펴본 이후에 수정된 방법으로 단위근을 식별하는 절차에 대해 설명하고자 한다.

2.1. DF-GLS 검정을 이용한 단위근 식별

ADF 검정을 실시할 때 $\{u_t\}$ 를 관측할 수 있다면, 식 (1.2)에 따라 상수항과 선형 추세를 모두 포함하지 않는 첫 번째 ADF-회귀모형식

$$\Delta u_t = \gamma u_{t-1} + \sum_{i=1}^k \delta_i \Delta u_{t-i} + e_t \quad (2.1)$$

을 이용하여 단위근 검정을 실시할 수 있을 것이다. 여기서 k 는 자연수이고, Δ 은 1차 차분을 나타내는 연산자로 후향연산자 B 에 대한 식으로 표현하면 $\Delta = 1 - B$ 와 같다. 그리고 e_t 는 ADF-회귀모형식의 오차항으로 $\{v_t\}$ 가 k 보다 낮은 차수를 가지는 자기회귀모형일 때 백색잡음이 된다. 일반적으로 $\{v_t\}$ 가 정상시계열인 경우에는 e_t 가 백색잡음이 아닐 수 있는데, 이는 Said와 Dickey (1984)가 설명한 것처럼 정상시계열 $\{v_t\}$ 를 자기회귀모형으로 근사시키기 때문이다.

그러나 안타깝게도 $\{y_t\}$ 만 관측가능하고 결정적 추세 $d_t = \beta_0 + \beta_1 t$ 가 알려져 있지 않다. 하지만 식 (1.1)이 성립할 때 $\Delta u_t = \Delta y_t - \beta_1 t$ 이 성립하므로, $\{u_t\}$ 에 대한 모형식 (2.1)을 다음처럼 변환할 수 있다:

$$\Delta y_t = c_0 + c_1 t + \gamma y_{t-1} + \sum_{i=1}^k \delta_i \Delta y_{t-i} + e_t. \quad (2.2)$$

여기서 상수항 c_0 와 t 의 계수 c_1 은 각각 $c_0 = -\beta_0 \gamma + \beta_1 (1 + \gamma - \sum_{i=1}^k \delta_i)$ 와 $c_1 = -\beta_1 \gamma$ 를 만족한다. 그러므로 ADF 검정을 실시할 때 선형 추세를 포함한 세 번째 ADF-회귀모형식 (2.2)을 사용할 수 있으나, 이 경우 Dickey와 Fuller (1979), Elliott 등 (1996)의 모의실험에서 보듯이 참 모형의 추세가 $d_t = 0$ 이라고 할지라도 검정력이 다소 떨어진다. 이와 같은 문제를 해결하기 위해 Elliott 등 (1996)은 DF-GLS 검정, 즉 일반화최소제곱법을 이용하여 추세를 제거한 후 ADF 검정을 실시하는 방법을 제안하였다.

식 (1.1)의 선형 추세, $d_t = \beta_0 + \beta_1 t$ 를 y_t 에서 제거하기 위해서는 먼저 β_0 와 β_1 을 추정해야 하며, 이 때 Elliott 등 (1996)은 일반화최소제곱법을 사용하였다. 그 이유는 식 (1.1)-(1.2)이 1차 자기회귀 모형을 만족하는 오차항을 갖는 회귀모형식과 유사하고, 다음처럼 변형 가능하기 때문이다:

$$y_1 = \beta_0 + \beta_1 + v_1; \\ y_t - \alpha y_{t-1} = \beta_0(1 - \alpha) + \beta_1(t - \alpha t + \alpha) + v_t, \quad t = 2, 3, \dots, n.$$

그러므로 잔차제곱합

$$S(b_0, b_1; \alpha) = (y_1 - b_0 - b_1)^2 + \sum_{t=2}^n \{(y_t - \alpha y_{t-1}) - b_0(1 - \alpha) - b_1(t - \alpha t + \alpha)\}^2$$

이 최소가 되도록 하는 값을 β_0, β_1 에 대한 추정량으로 사용할 수 있으며, 이 추정량은 $\{v_t\}$ 가 백색잡음일 때 β_0, β_1 에 대한 최량선형불편추정량이 됨을 쉽게 확인할 수 있다. 그러나 α 의 값이 알려져 있지 않으므로, α 대신 다른 값을 사용해야 하며 Elliott 등 (1996)은 단위근 근처에서의 검정력을 고려하여 $\alpha = 1 - 13.5/n$ 을 사용할 것을 추천하였고, Vougas (2007)는 α 에 대한 다양한 추정량의 사용을 검토하였다. Elliott 등 (1996)은 결정적 추세로 $d_t = \beta_0$ 처럼 상수항만 있고 $\beta_1 = 0$ 인 경우도 고려하였으며, 이 때에는 β_0 에 대한 추정값으로 $\hat{\beta}_0 = \arg \min_{b_0} S(b_0, 0; 1 - 7/n)$ 을 사용할 것을 제안하였다.

DF-GLS 검정에서는 위에 설명한 것처럼 β_0, β_1 에 대한 일반화최소제곱추정값을 구한 후, 식 (1.3)을 이용하여 추세를 제거한다. 추세가 제거된 시계열 $\{y_t^d\}$ 는 $\{u_t\}$ 에 대한 추정시계열이므로, 식 (2.1)에서 관측불가능한 $\{u_t\}$ 대신 $\{y_t^d\}$ 를 사용할 수 있다. 즉, 모형식

$$\Delta y_t^d = \gamma y_{t-1}^d + \sum_{i=1}^k \delta_i \Delta y_{t-i}^d + e_t \quad (2.3)$$

을 최소제곱법으로 적합시킨 결과를 바탕으로 ADF 검정을 수행한다.

2.2. Adaptive lasso를 이용한 단위근 식별

시계열 자료 y_1, y_2, \dots, y_n 이 주어졌을 때, 식 (1.2)의 α 가 1인지 여부를 알아보기 위해 Na (2020)는 ADF 검정에서 사용하는 세 가지 형태의 회귀모형식 중 선형 추세가 포함된 모형식 (2.2)에 포함된 모수 ($c_0, c_1, \gamma, \delta_1, \dots, \delta_k$)를 adaptive lasso 방법으로 추정하였다. ADF 검정에서는 모수에 대한 최소제곱추정량을 바탕으로 “ $H_0 : \gamma = 0$ (단위근을 가진다)”를 검정하지만, Na (2020)는 lasso 기법의 특성을 반영하여 adaptive lasso를 이용하여 구한 γ 의 추정값을 그대로 사용하였다. 즉 γ 에 대한 추정값이 0이면, 시계열 $\{y_t\}$ 가 단위근을 갖는 $I(1)$ 과정이라고 판단하였다.

Adaptive lasso 방법을 사용하기 위해서는 가중치와 조절모수를 선택해야하는데, Na (2020)는 모의실험에서 가중치로 모수에 대한 최소제곱추정값의 절대값의 역수를 사용하였고, AIC (Akaike, 1974), HQC (Hannan과 Quinn, 1979), BIC (Schwarz, 1978) 세 가지 정보함수를 사용하여 조절모수를 선택하였다. 그 결과 AIC 보다는 HQC와 BIC를 이용하여 조절모수를 선택할 때 단위근의 존재 여부를 좀 더 잘 판단하는 것을 알 수 있었다.

Na (2020)는 상수항과 기울기에 대한 사전정보가 없는 상태에서 단위근의 존재 여부와 추세의 기울기가 0인지 여부를 동시에 판단하는 문제를 다루었기 때문에 식 (2.2)처럼 세 가지 형태의 회귀모형식 중 가장 일반적인 형태를 고려하였다. 그러나 본 논문처럼 단위근의 존재 여부만 파악할 때는 결정적 추세 $d_t = \beta_0 + \beta_1 t$ 에 대한 정보나 가정에 따라 (2.2)의 모형식뿐만 아니라 $c_0 = c_1 = 0$ 인 모형식과 $c_1 = 0$ 인 모형식도 사용할 수 있다.

2.3. 추세제거와 adaptive lasso를 이용한 단위근 식별

Na (2020)의 모의실험 결과를 보면, $\beta_1 \neq 0$ 이고 $|\alpha| < 1$ 일 때 γ 에 대한 추정값이 0인 비율이 $\beta_1 = 0$ 이고 $|\alpha| < 1$ 일 때에 비해 다소 높게 나타났다. 즉 단위근 검정의 입장에서 볼 때 기울기가 0이 아닌 선형 추세가 존재하는 경우에는 선형 추세가 존재하지 않는 경우에 비해 검정력이 떨어지는 것이다. 본 논문에서는 이와 같은 문제점을 해결하기 위해 Elliott 등 (1996)처럼 원 시계열 $\{y_t\}$ 대신 선형 추세를 제거한 시계열 $\{y_t^d\}$ 을 이용하여 ADF-회귀모형식을 adaptive lasso로 추정하는 방법을 제안하고자 하며, 자세한 절차는 다음과 같다.

1. y_1, y_2, \dots, y_n 을 가지고 식 (1.1)의 β_0, β_1 에 대한 추정값 $\hat{\beta}_0, \hat{\beta}_1$ 을 구한다.
2. 1단계에서 구한 추정값을 이용하여 식 (1.3)처럼 $y_1^d, y_2^d, \dots, y_n^d$ 을 계산한다.
3. ADF-회귀모형식 (2.3)에 사용할 차수 k 와 adaptive lasso 추정에 사용할 조절모수 λ 와 가중치 $\mathbf{w} = (w_0, w_1, \dots, w_k)$ 를 선택한다.
4. 식 (2.3)의 모수 $(\gamma, \delta_1, \dots, \delta_k)$ 에 대한 추정값을 adaptive lasso로 구한다. 즉

$$L(\gamma, \delta_1, \dots, \delta_k) = \frac{1}{2(n-k-1)} \sum_{t=k+2}^n \left(\Delta y_t^d - \gamma y_{t-1}^d - \sum_{i=1}^k \delta_i \Delta y_{t-i}^d \right)^2 + \lambda \left(w_0 |\gamma| + \sum_{i=1}^k w_i |\delta_i| \right)$$

을 최소로 하는 값을 추정값 $(\hat{\gamma}, \hat{\delta}_1, \dots, \hat{\delta}_k)$ 으로 사용한다.

5. 4단계에서 구한 추정값 $\hat{\gamma}$ 가 0이면, $\alpha = 1$, 즉 단위근이 존재한다고 판단한다.

1단계에서 β_0, β_1 을 추정할 때 Vougas (2007)처럼 일반화최소제곱추정량을 사용할 수도 있으나, 본 논문에서는 최소제곱추정량

$$(\hat{\beta}_0^o, \hat{\beta}_1^o) = \arg \min_{b_0, b_1} S(b_0, b_1; 0) = \arg \min_{b_0, b_1} \sum_{t=1}^n (y_t - \beta_0 - \beta_1 t)^2 \quad (2.4)$$

을 $(\hat{\beta}_0, \hat{\beta}_1)$ 으로 사용하였다. 그 이유는 최소제곱추정량이 식 (1.1)과 같은 모형식에서 모수를 구할 때 가장 많이 사용되는 기본적인 추정량일 뿐만 아니라 $\hat{\beta}_1^o$ 은 α 의 값에 관계없이 일치성도 만족하기 때문이다. 그리고 본 논문에서는 식 (1.1)처럼 선형 추세 $d_t = \beta_0 + \beta_1 t$ 를 가정하였으므로 1,2단계에서 $\hat{d}_t = \hat{\beta}_0 + \hat{\beta}_1 t$ 를 구하여 추세를 제거하였으나, 일반적으로 $d_t = \beta_0, d_t = \beta_0 + \beta_1 t + \dots + \beta_p t^p, d_t = \beta_1 x_{1t} + \dots + \beta_p x_{pt}$ 등 다양한 형태의 함수를 사용해서 추세를 추정하여 제거해도 된다. 다만 3장의 모의실험 결과에서 보듯이 결정적 추세의 형태를 선택하는 것이 단위근 식별 결과에 영향을 많이 주므로 잘 선택해야 한다.

3, 4단계에서 추세를 제거한 시계열 $\{y_t^d\}$ 를 가지고 적합시킬 ADF-회귀모형식은 Elliott 등 (1996)처럼 상수항과 선형 추세, 즉 $c_0 + c_1t$ 항이 포함되지 않은 모형식이다. 이는 앞서 기술한 것처럼 $\{y_t^d\}$ 가 $\{u_t\}$ 에 대한 추정시계열이고, $\{u_t\}$ 가 식 (1.2)를 만족하기 때문이다.

3단계에서 차수 k , 가중치 \mathbf{w} , 조절모수 λ 를 선택할 때 다양한 값들을 사용할 수 있다. 그러나 아직 이에 대한 이론적인 연구가 부족한 상태이므로 3장 모의실험에서는 Na (2020)과 동일한 방식으로 값들을 선택하였다. 식 (2.3)의 차수 k 는 Schwert (1989), Na (2019, 2020)에서와 마찬가지로 $12(n/100)^{1/4}$ 을 넘지 않는 최대의 자연수로 선택하였으며, 가중치 \mathbf{w} 는 Zou (2006)처럼 $(\gamma, \delta_1, \dots, \delta_k)$ 에 대한 최소제곱추정값의 역수의 절대값을 사용하였다. 그리고 조절모수 λ 는 최소값과 최대값을 설정, 로그 스케일로 99등분하여 얻은 100개의 값들 중에서 AIC, HQC, BIC 등의 정보함수가 최소가 되도록 하는 값을 선택하였다.

3. 모의실험

Adaptive lasso 방법으로 ADF-회귀모형식을 추정하여 단위근을 식별할 때 원 시계열보다 추세를 제거한 시계열을 사용하는 것이 검정력을 향상시키는지 모의실험을 통해 알아보았다. 또한 추세를 제거할 때 사용하는 추세 함수와 추정방법에 따른 단위근 식별 결과, ADF 검정과 DF-GLS 검정과의 비교를 위한 모의실험도 수행하였다.

모의실험을 수행하기 위해 식 (1.1)–(1.2)에 주어진 모형식을 사용하여 시계열 자료 y_1, y_2, \dots, y_n , $n = 50, 100, 200, 500, 1000$ 을 생성하였다. 기존의 연구 결과에서 보듯이 선형 추세의 상수항보다는 기울기가 단위근 검정 결과에 영향을 더 주므로 본 모의실험에서는 $\beta_0 = 0, \beta_1 = 0, 0.5, 1$ 의 값을 사용하였고, α 의 값은 0.0, 0.4, 0.7, 0.9, 0.91, 0.92, \dots , 1.0처럼 0.9부터 1.0의 구간을 세분하여 살펴보았다. 그리고 정상시계열 $\{v_t\}$ 는

- (WN) $v_t = \epsilon_t$;
- (AR) $v_t = \phi v_{t-1} + \epsilon_t$, $\phi = \pm 0.5$;
- (MA) $v_t = \epsilon_t + \theta \epsilon_{t-1}$, $\theta = \pm 0.8, \pm 0.5$

의 경우를 고려하였다. 여기서 ϵ_t 는 서로 독립적으로 동일하게 표준정규분포를 따르는 확률변수이다.

3.1. 추세 제거 방법에 따른 단위근 식별 결과

우선 2.3절에서 설명한 방식대로 추세를 제거한 뒤 adaptive lasso 방법으로 식 (2.3)의 모형식을 추정하였다. 결정적 추세에 대한 함수로 상수함수 $d_t = \beta_0$ 와 1차함수 $d_t = \beta_0 + \beta_1 t$ 두 가지 형태를 고려하여 추세를 추정하였으며, 추정 방법으로 최소제곱법 이외에 추가로 일반화최소제곱법도 사용하였다. 자세한 추정량은 다음과 같다.

추세	Constant (C) : $d_t = \beta_0$	Linear Trend (LT) : $d_t = \beta_0 + \beta_1 t$
OLS	$\hat{\beta}_0^O = \arg \min_{b_0} S(b_0, 0; 0)$	$(\hat{\beta}_0^O, \hat{\beta}_1^O) = \arg \min_{b_0, b_1} S(b_0, b_1; 0)$
ERS	$\hat{\beta}_0^E = \arg \min_{b_0} S(b_0, 0; 1 - 7/n)$	$(\hat{\beta}_0^E, \hat{\beta}_1^E) = \arg \min_{b_0, b_1} S(b_0, b_1; 1 - 13.5/n)$
GLS	$\hat{\beta}_0^G = \arg \min_{b_0} S(b_0, 0; \hat{\alpha})$	$(\hat{\beta}_0^G, \hat{\beta}_1^G) = \arg \min_{b_0, b_1} S(b_0, b_1; \hat{\alpha})$

OLS는 최소제곱추정량을 뜻하고, ERS는 Elliott 등 (1996)이 DF-GLS 검정에서 사용한 일반화최소제곱추정량이다. 그리고 GLS는 Vougas (2007)이 고려했던 다양한 일반화최소제곱추정량 중 하나로 $\hat{\alpha}$ 는 $\hat{\rho}_{LS} = \sum_{t=2}^n \hat{u}_t \hat{u}_{t-1} / \sum_{t=2}^n \hat{u}_{t-1}^2$ 와 같다. 여기서 $\hat{u}_1, \hat{u}_2, \dots, \hat{u}_n$ 은 OLS 잔차이며, 결정적 추세로 선형 추세를 고려한 경우에는 $\hat{u}_t = y_t - \hat{\beta}_0^O - \hat{\beta}_1^O t$ 이고, 상수함수를 사용한 경우에는 $\hat{u}_t = y_t - \hat{\beta}_0^O$ 이다.

Table 1과 Table 2, Figure 1은 OLS, ERS, GLS 세 가지 방법으로 추세를 제거한 후 (2.3)의 γ 를 adaptive lasso로 추정한 결과 중 $\{v_t\}$ 가 WN인 경우를 정리한 것이다. Table 1의 숫자는 1,000번의 반복 실험 중 추정값

Table 1: Number of cases that $\hat{\gamma} = 0$

α	n	OLS						ERS						GLS					
		C			LT			C			LT			C			LT		
		AIC	HQC	BIC	AIC	HQC	BIC	AIC	HQC	BIC	AIC	HQC	BIC	AIC	HQC	BIC	AIC	HQC	BIC
$\beta_1 = 0$																			
0	50	1	2	3	0	0	0	188	233	314	0	3	6	1	2	3	0	0	0
	100	0	0	0	0	0	0	187	280	379	2	9	32	0	0	0	0	0	0
	200	0	0	0	0	0	0	124	256	377	0	2	31	0	0	0	0	0	0
	500	0	0	0	0	0	0	68	148	312	0	0	19	0	0	0	0	0	0
	1000	0	0	0	0	0	0	30	106	259	0	0	12	0	0	0	0	0	0
0.7	50	20	42	126	7	13	53	96	160	315	7	22	80	24	53	150	6	19	74
	100	0	0	6	1	1	1	38	83	161	1	2	7	0	0	7	0	0	2
	200	0	0	0	0	0	0	11	38	86	0	0	2	0	0	0	0	0	0
	500	0	0	0	0	0	0	0	11	47	0	0	1	0	0	0	0	0	0
	1000	0	0	0	0	0	0	0	1	16	0	0	0	0	0	0	0	0	0
0.9	50	204	338	607	86	182	415	323	503	724	96	218	475	271	429	653	111	239	476
	100	74	205	451	22	89	265	133	325	616	29	135	361	97	251	493	37	130	334
	200	2	18	101	1	5	51	16	80	257	1	13	104	2	27	130	1	9	82
	500	0	0	0	0	0	0	0	2	14	0	0	0	0	0	0	0	0	0
	1000	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
0.95	50	330	484	721	139	278	522	513	667	843	177	342	589	479	625	787	215	370	588
	100	198	455	711	87	251	518	375	646	856	131	366	656	315	566	792	147	358	631
	200	46	235	518	15	133	358	111	406	694	26	209	488	79	315	590	31	198	462
	500	0	1	43	0	0	26	1	20	155	0	0	44	0	4	60	0	0	35
	1000	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0
1	50	507	676	853	161	313	573	719	837	935	217	383	641	680	787	898	259	412	638
	100	483	690	874	162	414	683	776	897	957	277	557	801	718	850	930	305	573	792
	200	470	733	898	170	456	751	804	943	987	334	665	890	766	912	971	344	672	874
	500	485	804	945	194	537	822	838	972	997	406	777	936	820	962	995	408	775	932
	1000	507	820	958	188	541	863	834	970	995	411	804	961	822	964	994	404	797	958
$\beta_1 = 0.5$																			
0	50	948	981	997	0	0	0	808	879	969	4	9	14	259	337	534	0	0	0
	100	983	1000	1000	0	0	0	494	611	831	0	2	20	95	126	226	0	0	0
	200	999	1000	1000	0	0	0	256	379	558	0	2	38	54	88	150	0	0	0
	500	1000	1000	1000	0	0	0	21	87	254	0	0	27	3	17	69	0	0	0
	1000	1000	1000	1000	0	0	0	7	61	269	0	0	9	2	23	157	0	0	0
0.7	50	966	988	999	3	8	62	780	863	945	7	18	88	208	277	459	6	17	80
	100	997	1000	1000	0	0	0	260	382	622	0	0	7	16	30	67	0	0	0
	200	997	1000	1000	0	0	0	7	22	56	0	0	3	0	1	1	0	0	0
	500	1000	1000	1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1000	1000	1000	1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.9	50	920	961	987	106	207	444	739	806	893	127	255	503	276	346	478	153	282	502
	100	987	997	1000	12	67	276	284	368	524	20	103	357	38	50	92	25	110	332
	200	1000	1000	1000	1	1	43	5	8	18	0	7	86	0	0	0	0	5	69
	500	1000	1000	1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	1000	1000	1000	1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.95	50	905	952	986	118	260	534	763	823	889	148	315	581	316	378	495	198	359	591
	100	963	990	1000	70	249	515	336	428	564	125	357	624	98	124	171	131	357	611
	200	993	999	1000	13	117	371	19	27	38	21	193	501	3	4	6	22	189	467
	500	1000	1000	1000	0	0	30	0	0	0	0	0	57	0	0	0	0	0	45
	1000	1000	1000	1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	50	892	944	987	204	360	600	768	820	886	255	419	657	391	446	546	304	450	659
	100	962	995	998	169	434	691	426	507	630	290	571	800	110	158	231	320	588	790
	200	991	1000	1000	188	453	742	51	80	131	365	656	857	10	20	40	376	658	849
	500	998	1000	1000	214	573	828	0	0	0	445	789	941	0	0	0	447	783	930
	1000	999	1000	1000	191	583	873	0	0	0	412	810	958	0	0	0	408	808	955
$\beta_1 = 1$																			
0	50	957	977	991	0	0	1	895	927	973	1	2	11	565	614	692	0	0	1
	100	991	1000	1000	0	0	0	908	957	990	3	7	28	618	716	815	0	0	0
	200	999	1000	1000	0	0	0	882	960	994	0	1	34	710	836	939	0	0	0
	500	1000	1000	1000	0	0	0	871	972	999	0	0	19	743	933	994	0	0	0
	1000	1000	1000	1000	0	0	0	916	990	1000	0	0	13	864	983	999	0	0	0
0.7	50	949	977	993	5	12	64	827	856	889	9	21	83	369	390	412	9	17	74
	100	998	1000	1000	0	0	1	725	786	845	0	1	7	225	257	304	0	1	3
	200	999	1000	1000	0	0	0	418	551	661	0	0	1	86	161	226	0	0	0
	500	1000	1000	1000	0	0	0	37	160	394	0	0	0	3	36	145	0	0	0
	1000	1000	1000	1000	0	0	0	1	52	365	0	0	0	1	14	212	0	0	0
0.9	50	924	961	979	87	187	422	765	796	841	108	230	482	332	357	395	135	253	483
	100	991	997	999	17	92	267	617	691	760	28	129	347	213	239	271	31	134	327
	200	1000	1000	1000	1	3	52	202	288	352	1	10	90	31	51	77	1	6	67
	500	1000	1000	1000	0	0	0	3	9	18	0	0	0	0	1	1	0	0	0
	1000	1000	1000	1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0.95	50	901	940	965	131	273	543	766	786	841	171	343	601	375	397	420	212	372	600
	100	986	997	1000	63	241	533	605	664	721	118	352	645	228	250	281	136	373	626
	200	999	1000	1000	19	110	385	211	278	357	28	202	491	45	63	78	27	191	468
	500	1000	1000	1000	0	0	28	1	5	10	0	3	62	0	1	1	0	1	46
	1000	1000	1000	1000	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	50	896	946	971	186	337	600	769	800	833	239	404	670	377	391	413	279	434	672
	100	972	991	997	163	410	686	628	684	736	274	568	783	274	302	329	307	557	

Table 2: Sample mean of $\hat{\gamma}$ (sample standard deviation of $\hat{\gamma}$)

α	n	OLS				ERS				GLS			
		C		LT		C		LT		C		LT	
		AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC	AIC	BIC
$\beta_1 = 0$													
0	100	-0.967 (0.161)	-0.925 (0.121)	-0.988 (0.178)	-0.940 (0.113)	-0.450 (0.372)	-0.409 (0.383)	-0.763 (0.257)	-0.731 (0.263)	-0.967 (0.161)	-0.924 (0.121)	-0.988 (0.175)	-0.940 (0.113)
	500	-0.987 (0.054)	-0.979 (0.047)	-0.991 (0.053)	-0.982 (0.047)	-0.326 (0.354)	-0.347 (0.372)	-0.546 (0.341)	-0.605 (0.320)	-0.987 (0.054)	-0.979 (0.047)	-0.991 (0.053)	-0.982 (0.047)
	1000	-0.993 (0.037)	-0.990 (0.034)	-0.996 (0.038)	-0.991 (0.034)	-0.286 (0.351)	-0.329 (0.372)	-0.460 (0.348)	-0.558 (0.341)	-0.993 (0.037)	-0.990 (0.034)	-0.996 (0.038)	-0.991 (0.034)
0.7	100	-0.307 (0.096)	-0.272 (0.087)	-0.336 (0.107)	-0.296 (0.088)	-0.220 (0.108)	-0.186 (0.115)	-0.293 (0.092)	-0.264 (0.087)	-0.304 (0.095)	-0.270 (0.088)	-0.328 (0.102)	-0.292 (0.089)
	500	-0.300 (0.037)	-0.290 (0.035)	-0.305 (0.038)	-0.294 (0.035)	-0.201 (0.091)	-0.194 (0.092)	-0.261 (0.056)	-0.254 (0.054)	-0.300 (0.037)	-0.290 (0.035)	-0.304 (0.038)	-0.294 (0.035)
	1000	-0.300 (0.024)	-0.294 (0.022)	-0.302 (0.024)	-0.296 (0.023)	-0.189 (0.095)	-0.196 (0.088)	-0.255 (0.056)	-0.255 (0.048)	-0.300 (0.024)	-0.294 (0.022)	-0.302 (0.024)	-0.296 (0.023)
0.9	100	-0.118 (0.071)	-0.072 (0.074)	-0.156 (0.080)	-0.106 (0.079)	-0.089 (0.062)	-0.046 (0.064)	-0.130 (0.071)	-0.084 (0.075)	-0.109 (0.071)	-0.065 (0.073)	-0.137 (0.076)	-0.092 (0.079)
	500	-0.102 (0.022)	-0.092 (0.022)	-0.107 (0.023)	-0.097 (0.022)	-0.084 (0.025)	-0.075 (0.026)	-0.097 (0.022)	-0.088 (0.022)	-0.101 (0.022)	-0.091 (0.022)	-0.105 (0.023)	-0.095 (0.023)
	1000	-0.100 (0.015)	-0.094 (0.015)	-0.102 (0.016)	-0.096 (0.015)	-0.084 (0.021)	-0.079 (0.021)	-0.094 (0.016)	-0.089 (0.015)	-0.100 (0.015)	-0.094 (0.015)	-0.102 (0.016)	-0.096 (0.015)
0.95	100	-0.075 (0.064)	-0.035 (0.062)	-0.116 (0.079)	-0.065 (0.079)	-0.047 (0.052)	-0.016 (0.042)	-0.090 (0.070)	-0.042 (0.067)	-0.060 (0.062)	-0.025 (0.056)	-0.093 (0.076)	-0.048 (0.072)
	500	-0.053 (0.018)	-0.043 (0.019)	-0.059 (0.019)	-0.049 (0.019)	-0.044 (0.016)	-0.033 (0.019)	-0.052 (0.017)	-0.043 (0.018)	-0.051 (0.018)	-0.041 (0.019)	-0.056 (0.019)	-0.046 (0.019)
	1000	-0.051 (0.012)	-0.045 (0.012)	-0.053 (0.012)	-0.047 (0.012)	-0.044 (0.012)	-0.039 (0.012)	-0.049 (0.012)	-0.043 (0.012)	-0.050 (0.012)	-0.045 (0.012)	-0.052 (0.012)	-0.046 (0.012)
1	100	-0.037 (0.051)	-0.013 (0.037)	-0.087 (0.068)	-0.038 (0.062)	-0.012 (0.032)	-0.004 (0.021)	-0.058 (0.056)	-0.022 (0.048)	-0.016 (0.039)	-0.006 (0.029)	-0.058 (0.060)	-0.024 (0.052)
	500	-0.006 (0.009)	-0.001 (0.005)	-0.015 (0.012)	-0.005 (0.011)	-0.001 (0.004)	0.000 (0.001)	-0.008 (0.010)	-0.002 (0.006)	-0.001 (0.005)	0.000 (0.002)	-0.008 (0.010)	-0.002 (0.007)
	1000	-0.003 (0.004)	0.000 (0.002)	-0.008 (0.006)	-0.002 (0.005)	-0.001 (0.002)	0.000 (0.001)	-0.004 (0.005)	-0.001 (0.003)	-0.001 (0.003)	0.000 (0.001)	-0.004 (0.005)	-0.001 (0.003)
$\beta_1 = 1$													
0	100	0.000 (0.000)	0.000 (0.000)	-1.001 (0.202)	-0.948 (0.117)	0.000 (0.001)	0.000 (0.000)	-0.766 (0.255)	-0.736 (0.260)	0.002 (0.002)	0.001 (0.002)	-1.002 (0.207)	-0.948 (0.117)
	500	0.000 (0.000)	0.000 (0.000)	-0.997 (0.070)	-0.986 (0.046)	0.000 (0.000)	0.000 (0.000)	-0.545 (0.346)	-0.596 (0.328)	0.000 (0.000)	0.000 (0.000)	-0.997 (0.070)	-0.986 (0.046)
	1000	0.000 (0.000)	0.000 (0.000)	-0.994 (0.039)	-0.989 (0.032)	0.000 (0.000)	0.000 (0.000)	-0.452 (0.346)	-0.550 (0.340)	0.000 (0.000)	0.000 (0.000)	-0.994 (0.039)	-0.989 (0.032)
0.7	100	0.000 (0.000)	0.000 (0.000)	-0.346 (0.115)	-0.300 (0.090)	0.001 (0.002)	0.001 (0.002)	-0.303 (0.096)	-0.268 (0.091)	0.005 (0.003)	0.005 (0.004)	-0.340 (0.115)	-0.295 (0.092)
	500	0.000 (0.000)	0.000 (0.000)	-0.306 (0.037)	-0.294 (0.033)	0.000 (0.000)	0.000 (0.000)	-0.260 (0.055)	-0.253 (0.053)	0.000 (0.000)	0.001 (0.000)	-0.305 (0.037)	-0.293 (0.033)
	1000	0.000 (0.000)	0.000 (0.000)	-0.300 (0.024)	-0.294 (0.023)	0.000 (0.000)	0.000 (0.000)	-0.251 (0.058)	-0.252 (0.049)	0.000 (0.000)	0.000 (0.000)	-0.300 (0.024)	-0.294 (0.023)
0.9	100	0.000 (0.000)	0.000 (0.000)	-0.157 (0.081)	-0.108 (0.082)	0.002 (0.003)	0.001 (0.003)	-0.132 (0.072)	-0.088 (0.078)	0.005 (0.004)	0.006 (0.004)	-0.140 (0.081)	-0.095 (0.082)
	500	0.000 (0.000)	0.000 (0.000)	-0.106 (0.024)	-0.096 (0.023)	0.001 (0.000)	0.001 (0.000)	-0.096 (0.023)	-0.087 (0.023)	0.001 (0.000)	0.001 (0.000)	-0.104 (0.024)	-0.095 (0.023)
	1000	0.000 (0.000)	0.000 (0.000)	-0.103 (0.015)	-0.097 (0.015)	0.000 (0.000)	0.000 (0.000)	-0.094 (0.016)	-0.089 (0.015)	0.000 (0.000)	0.000 (0.000)	-0.102 (0.015)	-0.096 (0.015)
0.95	100	0.000 (0.000)	0.000 (0.000)	-0.113 (0.071)	-0.058 (0.071)	0.002 (0.003)	0.002 (0.003)	-0.088 (0.063)	-0.042 (0.062)	0.005 (0.004)	0.006 (0.005)	-0.089 (0.068)	-0.046 (0.066)
	500	0.000 (0.000)	0.000 (0.000)	-0.058 (0.019)	-0.048 (0.019)	0.001 (0.000)	0.001 (0.000)	-0.052 (0.018)	-0.042 (0.019)	0.001 (0.000)	0.001 (0.000)	-0.055 (0.019)	-0.045 (0.019)
	1000	0.000 (0.000)	0.000 (0.000)	-0.053 (0.013)	-0.047 (0.012)	0.000 (0.000)	0.000 (0.000)	-0.049 (0.012)	-0.044 (0.011)	0.000 (0.000)	0.001 (0.000)	-0.052 (0.012)	-0.046 (0.012)
1	100	0.000 (0.001)	0.000 (0.000)	-0.088 (0.066)	-0.038 (0.063)	0.002 (0.003)	0.002 (0.004)	-0.063 (0.059)	-0.025 (0.053)	0.005 (0.004)	0.006 (0.005)	-0.062 (0.063)	-0.027 (0.055)
	500	0.000 (0.000)	0.000 (0.000)	-0.015 (0.013)	-0.004 (0.011)	0.001 (0.000)	0.001 (0.000)	-0.008 (0.010)	-0.001 (0.006)	0.001 (0.000)	0.001 (0.000)	-0.008 (0.010)	-0.002 (0.006)
	1000	0.000 (0.000)	0.000 (0.000)	-0.007 (0.006)	-0.002 (0.005)	0.000 (0.000)	0.000 (0.000)	-0.004 (0.005)	-0.001 (0.003)	0.000 (0.000)	0.000 (0.000)	-0.004 (0.005)	-0.001 (0.003)

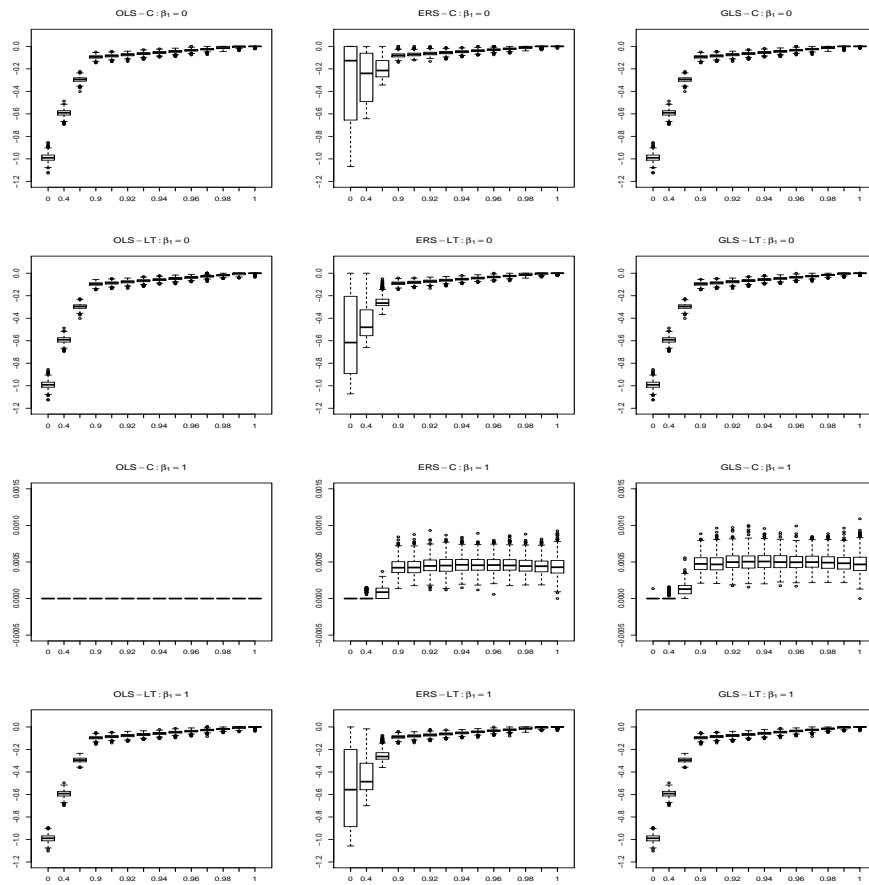


Figure 1: Boxplot of $\hat{\gamma}$ when $n = 1000$.

$\hat{\gamma}$ 이 0인 경우, 즉 단위근이 존재한다고 판단한 경우의 수를 나타내고, Table 2의 숫자는 1,000개의 추정값 $\hat{\gamma}$ 에 대한 표본평균이며, () 안의 숫자는 표본표준편차이다. Figure 1은 $n = 1000$ 일 때 1,000개의 추정값에 대한 상자그림이며, 표와 그림으로부터 다음과 같은 사실을 파악할 수 있다.

- 추세에 대한 함수로 t 에 대한 1차 함수, LT를 사용한 경우에는 참 추세의 기울기가 γ 의 추정에 큰 영향을 주지 않으며, 다음과 같은 공통적인 특징을 가진다.
 - Table 1을 보면, 자료의 개수 n 이 클수록 단위근의 존재를 판단하는 정확도가 높아진다. 그리고 $\alpha < 1$ 인 경우, α 의 값이 1에서 멀리 떨어질수록 단위근이 존재하지 않는다고 올바르게 판단하는 비율이 높아지며, $n \geq 500$ 일 때는 95% 이상의 정확도를 보인다.
 - Table 2와 Figure 1을 보면, OLS와 GLS로 선형 추세를 추정할 경우에는 추정방법에 관계없이 γ 에 대한 추정값 $\hat{\gamma}$ 의 분포가 거의 비슷하다. 그리고 대부분의 경우 $\hat{\gamma}$ 의 평균이 $\alpha - 1$ 과 유사한 값을 가지는데, $\alpha - 1$ 은 $\{v_t\}$ 가 WN인 경우 γ 에 대한 참값임을 식 변환을 통해 쉽게 확인할 수 있다. 또한 n 의 값이 커질수록, adaptive lasso를 이용하여 구한 γ 의 추정값들의 평균이 $\alpha - 1$ 에 가까워지고 표준편차가 0으로 감소한다. 이를 통해 OLS와 GLS 방법으로 선형 추세를 제거한 후 구한 adaptive lasso 추정량은 일치성을 만족한다

고 생각할 수 있다.

- ERS의 경우는 추정 결과가 조금 다르다. $\alpha \geq 0.9$ 일 때는 OLS나 GLS처럼 γ 에 대한 추정값이 참값과 유사하지만, $\alpha \leq 0.7$ 일 때, 특히 $\alpha = 0$ 일 때는 n 이 큰 경우에도 γ 에 대한 추정값이 참값 -1 과 가깝지 않고, 표준편차도 크다. 이는 ERS가 α 가 1 근처일 때 검정력을 높이기 위해 잔차제곱합 $S(b_0, b_1; a)$ 에서 1에 가까운 값을 a 로 사용하기 때문이며, 이로 인해 Table 1에서 보듯이 $\alpha = 0$ 인 경우 단위근이 존재한다고 잘못 판단하는 경우가 2~3% 정도 발생하였다.
- Table 1을 보면, $\hat{\gamma} = 0$ 인 경우, 즉 단위근이 존재한다고 판단한 비율은 선형 추세를 추정하는 방법에 따라 크게 차이가 나지는 않지만, OLS를 사용하는 경우에 비해 ERS나 GLS의 경우의 비율이 조금씩 높다. 즉 단위근 검정의 입장에서 볼 때 OLS로 선형 추세를 제거하는 방법이 다른 추정방법을 사용하는 것에 비해 제1종의 오류를 범할 확률과 검정력이 조금씩 높게 나타난다.
- Adaptive lasso의 조절모수를 선택하는 방법에 따라 $\hat{\gamma} = 0$ 의 비율이 달라진다. AIC를 사용할 때 비율이 가장 낮고, BIC를 사용할 때 비율이 가장 높다. $\alpha = 1$, 즉 단위근이 존재하는 경우 올바르게 판단한 비율을 보면, AIC로 조절모수를 선택한 경우는 20%~40% 정도로 비율이 높지 않다. 이에 비해 BIC의 경우는 정확하게 판단한 비율이 60% 이상으로 높고, n 이 증가할수록 비율이 약 90%까지 증가한다.
- 상수항수 C 를 이용하여 추세를 추정하는 경우에는 참 추세의 기울기와 추정 방법에 따라 판단 결과가 달라진다.
 - 참 추세의 기울기가 0인 경우에는 Table 2와 Figure 1에서 보듯이 $\alpha \leq 0.7$ 이면서 ERS를 사용한 경우를 제외한 대부분의 경우 $\hat{\gamma}$ 가 참값 $\gamma = \alpha - 1$ 를 잘 추정한다. 그 결과 Table 1의 결과처럼 α 가 1에 가까우면서 n 이 작은 때를 제외한 나머지의 경우 단위근의 존재 여부를 판단하는 정확도가 높게 나타난다. 특히 BIC로 adaptive lasso의 조절모수를 선택하고, $n \geq 500$ 으로 큰 경우에는 α 의 값에 관계없이 정확도가 95% 이상으로 매우 높다. 이는 추세를 추정할 때 사용한 함수가 참 추세의 형태와 일치하기 때문인 것으로 보인다.
 - 이에 비해 참 추세의 기울기가 0이 아닌 경우에는 단위근의 존재를 판단한 결과가 추세를 추정한 방법에 따라 다르다. Table 1을 보면, OLS로 추세를 추정할 경우에는 α 와 조절모수를 선택할 때 사용한 정보함수에 관계없이 대부분 단위근이 존재한다고 판단한 비율이 100%에 가깝다. 이는 ADF 검정에서 기울기가 0이 아닌 선형 추세가 존재하는 경우 상수항만 포함된 ADF-회귀모형을 사용하여 단위근 검정을 실시했을 때와 같은 결과이다.
 - 그러나 Table 1에 보듯이 ERS와 GLS의 경우에는 $\hat{\gamma} = 0$ 인 비율이 낮다. 심지어 n 이 크고 $\alpha = 1$ 일 때에도 $\hat{\gamma} = 0$ 인 비율이 0에 가깝다. Table 2와 Figure 1을 통해 좀 더 추정값의 분포를 자세히 살펴보면, $\hat{\gamma}$ 는 비록 0은 아니지만 0에 매우 가까운 양수를 갖는다. 이런 현상은 OLS 때와는 달리 ADF-회귀모형에서 독립변수로 사용되는 $\{y_t^d\}$ 의 표본평균이 0에 가깝지 않고, 종속변수 $\{\Delta y_t^d\}$ 의 표본평균이 참 추세의 기울기 β_1 과 유사하며, ADF-회귀모형 중 (2.3)처럼 상수항을 포함하지 않는 모형식을 사용하기 때문에 나타나는 것으로 보인다. 비록 본 논문에는 결과를 포함시키지는 않았지만, GLS의 경우 상수항이 포함된 ADF-회귀모형을 사용하면 이런 현상이 어느 정도 해결되는 것으로 나타났다.

위에 기술된 결과를 종합적으로 판단할 때, 추세를 추정할 때 사용하는 함수의 형태가 참 추세를 포함할 수 있도록 선택하는 것이 추정 방법을 선택하는 것보다 중요하다. 그리고 세 가지 추정방법 중에는 추세를 잘못 추정했을 때의 현상이나 추정 방법의 편리성과 추정값의 성질, 검정력 등을 고려할 때 OLS를 사용하는 것이 ERS나 GLS를 사용하는 것에 비해 나쁘지 않다.

추가로 실제 자료 분석에서는 $\{u_t\}$ 를 사용하는 것이 불가능하지만, $\{u_t\}$ 대신 $\{y_t^d\}$ 를 사용하는 것이 어떤 영향을 미치는지 알아보기 위해 두 시계열 $\{u_t\}$ 와 $\{y_t^d\}$ 를 가지고 adaptive lasso를 이용하여 γ 의 값을 추정한

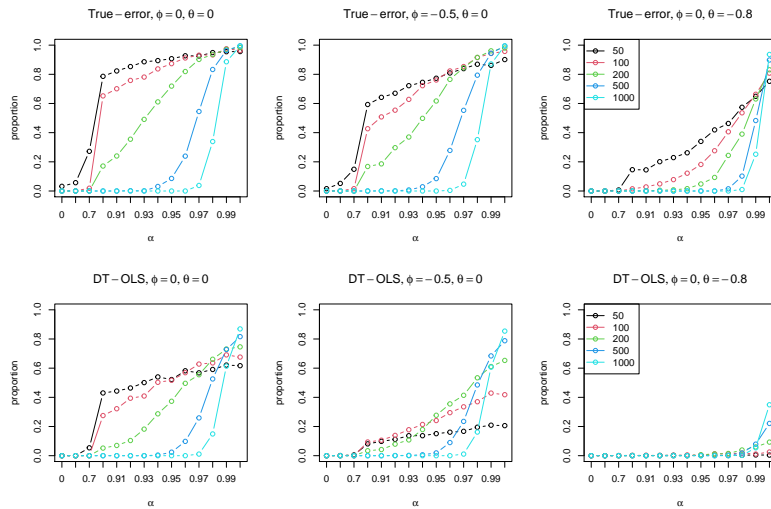


Figure 2: Proportion of $\hat{\gamma} = 0$ when $\beta_1 = 0$.

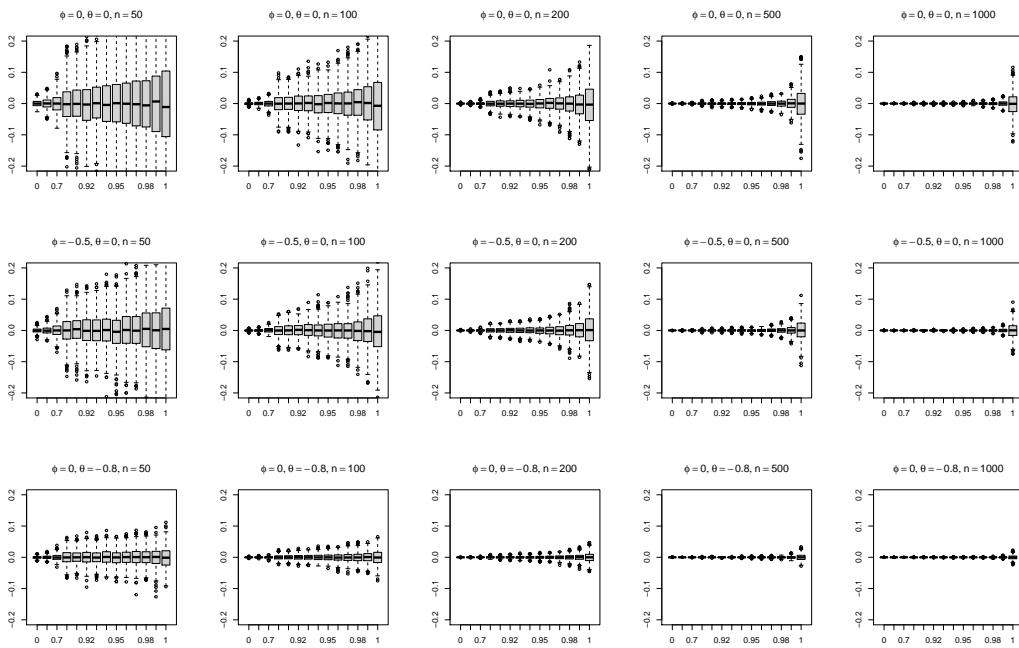


Figure 3: Boxplot of $\hat{\beta}_1^Q$ when $\beta_1 = 0$.

결과를 비교하였다. 모의실험 결과는 OLS 방법으로 추세를 추정하여 제거하고 BIC로 조절모수를 선택한 경우를 대표로 하여 Figures 2-3에 정리하였다. Figure 2의 1행과 2행은 각각 $\{u_t\}$ 와 $\{y_t^d\}$ 를 가지고 구한 γ 의 추정값 1,000개 중 0의 비율, 즉 단위근이 존재한다고 판단한 비율을 나타낸다. 그리고 Figure 3은 1,000번의 반복실험을 통해 구한 최소제곱추정값 $\hat{\beta}_1^Q$ 에 대한 상자그림이다. 이들 그림을 보면, $\{v_t\}$ 에 대한 모형이 동일한 경우에는 자료의 개수 n 이 증가함에 따라 추세를 좀 더 정확히 추정하고 그 결과 $\{u_t\}$ 를 이용하여 단위근을

판단한 결과와 $\{y_t^d\}$ 를 이용한 결과가 더 가까워지는 것을 알 수 있다. 또한 $\{u_t\}$ 를 이용할 때 n 이 증가함에 따라 단위근을 판단하는 정확도가 증가하며, $n \geq 500$ 인 경우에는 정확도가 매우 높다. 이와 같은 영향으로 $\{u_t\}$ 대신 $\{y_t^d\}$ 를 사용할 때에도 n 이 클 때 정확도가 높게 나타났다. 그러나 반드시 추세에 대한 정확한 추정이 단위근 식별의 정확도 향상으로 이어진다고 하기는 어렵다. Figures 2-3에서 $\alpha = 1$ 이고 n 이 동일한 경우를 보면, $\theta = -0.8$ 인 경우가 나머지 경우에 비해 β_1 을 좀 더 정확하게 추정하지만 단위근이 존재한다고 정확하게 판단한 비율은 오히려 낮은 것을 확인할 수 있다. 이는 ADF-회귀모형식에서 자기회귀모형 근사를 사용하기 때문인 것으로 판단된다. 마지막으로 n 이 작을 때는 $\{u_t\}$ 를 가지고 구한 γ 에 대한 추정값 중 0인 비율이 높게 나타난 반면 $\{y_t^d\}$ 를 가지고 γ 를 추정한 경우에는 0의 비율이 다소 낮게 나타났다. 시계열의 길이가 짧은 경우, 즉 n 이 작은 경우에는 α 의 값이 1에 가까울 때 $\beta_1 = 0$ 일지라도 시계열 그림을 그렸을 때 증가나 감소의 경향이 나타나는 경우가 종종 발생한다. 이런 경우 β_1 에 대한 추정값은 참값 0과 다소 차이가 발생하지만, $\{y_t^d\}$ 처럼 추세 제거를 통해 증가하거나 감소하는 경향을 줄이거나 없앨 수 있다. 그러므로 $\{u_t\}$ 를 사용할 때는 경향 때문에 단위근이 존재한다고 판단할 수 있지만, 오히려 $\{y_t^d\}$ 는 경향이 제거되었으므로 단위근이 존재하지 않는다고 판단할 수 있기 때문에 이와 같은 특이한 현상이 발생한 것으로 생각한다.

3.2. Adaptive lasso를 이용한 단위근 식별 : 원 시계열 vs. 추세를 제거한 시계열

본 절에서는 adaptive lasso로 ADF-회귀모형식을 추정하여 단위근을 식별할 때 원 시계열보다 추세를 제거한 시계열을 사용하는 것이 검정력을 향상시키는지 모의실험을 통해 알아보려 한다.

먼저 2.2절에서 설명한 방식, 즉 원 시계열 $\{y_t\}$ 을 가지고 adaptive lasso 방법으로 단위근의 존재 여부를 판단하는 방법에서는 두 가지 ADF-회귀모형식

$$(MC) \Delta y_t = c_0 + \gamma y_{t-1} + \sum_{i=1}^k \delta_i \Delta y_{t-i} + e_t;$$

$$(MT) \Delta y_t = c_0 + c_1 t + \gamma y_{t-1} + \sum_{i=1}^k \delta_i \Delta y_{t-i} + e_t$$

을 고려하였으며, k 는 $12(n/100)^{1/4}$ 을 넘지 않는 최대의 자연수, 가중치는 모형식에 주어진 계수에 대한 최소 제곱추정값의 역수의 절대값을 사용하였다. 그리고 Na (2020)의 모의실험 결과를 반영하여 조절모수는 BIC를 이용하여 선택하였다. 이에 대한 결과는 Table 3과 Figures 4-6의 “ALASSO” 부분에 정리하였다.

다음으로 추세를 제거한 시계열 $\{y_t^d\}$ 을 가지고 adaptive lasso를 이용하여 단위근의 존재를 판단하는 방법에서는 3.1절에서 고려한 다양한 방법 중에서 OLS를 통한 추세 제거와 BIC를 이용한 조절모수 선택 두 가지를 조합한 방법을 사용하였다. OLS, ERS, GLS 중 OLS 추정방법으로 추세를 추정한 것은 OLS 추정량이 가장 기본적인 추정량일 뿐만 아니라 3.1절의 모의실험 결과에서 보듯이 다른 추정방법에 비해 추정 결과와 단위근 식별 정도가 나쁘지 않으며, 추세 함수로 상수함수를 사용하는 경우 결과가 단위근 검정의 결과와 유사하기 때문이다. ADF 검정이나 PP 검정과 같은 대부분의 단위근 검정에서는 귀무가설로 “단위근이 존재한다”를 설정하는데, 이는 단위근이 존재할 때 단위근을 제대로 파악하는 것이 좀 더 중요하고, 검정력 조정은 2차적인 문제임을 뜻한다고 생각할 수 있다. 이와 같은 맥락에서 adaptive lasso의 조절모수를 선택할 때 AIC, HQC, BIC 세 가지 정보함수 중 $\alpha = 1$ 일 때 단위근이 존재한다고 판단한 비율이 높게 나타나는 BIC를 사용하였다. OLS로 추세를 제거한 뒤 단위근 존재 여부를 판단한 결과는 Table 3과 Figures 4-6의 “DT-OLS” 부분에 정리하였으며, “C”와 “LT”는 각각 추세를 제거할 때 사용한 추세함수의 형태가 상수함수와 t 에 대한 1차 함수인 경우를 의미한다.

Table 3은 $n = 500$ 일 때 1000번 중 단위근이 존재한다고 판단한 경우의 수를 모의실험에서 고려한 자료생

Table 3: Number of unit root non-rejections

α	ϕ	θ	$\beta_1 = 0$								$\beta_1 = 1$							
			ADF		DF-GLS		ALASSO		DT-OLS		ADF		DF-GLS		ALASSO		DT-OLS	
			MC	MT	C	LT	MC	MT	C	LT	MC	MT	C	LT	MC	MT	C	LT
0	0	0	0	0	147	40	0	0	0	0	1000	0	1000	41	1000	0	1000	0
	-0.5	0	0	0	387	196	0	0	0	0	1000	0	1000	190	1000	0	1000	0
	0.5	0	0	0	21	2	0	0	0	0	1000	0	1000	1	1000	0	1000	0
	0	-0.8	0	0	759	712	0	0	0	0	1000	0	1000	688	1000	0	1000	0
	0	-0.5	0	0	487	323	0	0	0	0	1000	0	1000	316	1000	0	1000	0
	0	0.5	0	0	63	6	0	0	0	0	1000	0	1000	13	1000	0	1000	0
	0	0.8	0	0	56	7	0	0	0	0	1000	0	1000	14	1000	0	1000	0
0.4	0	0	0	0	24	3	0	0	0	0	1000	0	1000	1	1000	0	1000	0
	-0.5	0	0	0	168	48	0	0	0	0	1000	0	1000	41	1000	0	1000	0
	0.5	0	0	0	1	0	0	0	0	0	1000	0	1000	0	1000	0	1000	0
	0	-0.8	0	0	605	452	0	0	0	0	1000	0	1000	482	1000	0	1000	0
	0	-0.5	0	0	255	104	0	0	0	0	1000	0	1000	89	999	0	1000	0
	0	0.5	0	0	4	1	0	0	0	0	1000	0	1000	0	1000	0	1000	0
	0	0.8	0	0	9	0	0	0	0	0	1000	0	1000	0	1000	0	1000	0
0.7	0	0	0	0	0	0	0	0	0	0	1000	0	1000	0	1000	0	1000	0
	-0.5	0	0	0	12	1	0	0	0	0	1000	0	1000	1	1000	0	1000	0
	0.5	0	0	0	0	0	0	0	0	0	1000	0	1000	0	1000	0	1000	0
	0	-0.8	0	0	355	171	0	0	0	0	1000	0	1000	174	1000	0	1000	0
	0	-0.5	0	0	44	8	0	0	0	0	1000	0	1000	3	1000	0	1000	0
	0	0.5	0	0	0	0	0	0	0	0	1000	0	1000	0	1000	0	1000	0
	0	0.8	0	0	0	0	0	0	0	0	1000	0	1000	0	1000	0	1000	0
0.9	0	0	0	0	0	0	0	0	0	0	1000	0	1000	0	1000	30	1000	0
	-0.5	0	0	0	0	0	0	0	0	0	1000	0	1000	0	1000	6	1000	0
	0.5	0	0	0	0	0	0	0	0	0	1000	0	1000	0	1000	9	1000	0
	0	-0.8	0	0	143	52	0	0	0	0	1000	0	1000	52	1000	0	1000	0
	0	-0.5	0	0	6	1	0	0	0	0	1000	0	1000	0	1000	0	1000	0
	0	0.5	0	0	0	0	0	0	0	0	1000	0	1000	0	1000	3	1000	0
	0	0.8	1	14	0	0	0	0	0	0	1000	7	1000	4	1000	30	1000	2
0.95	0	0	13	96	0	12	64	49	61	30	1000	74	1000	7	1000	624	1000	23
	-0.5	0	14	77	2	16	60	40	60	34	1000	81	1000	14	1000	197	1000	29
	0.5	0	22	115	0	13	89	61	89	51	1000	117	1000	19	1000	279	1000	40
	0	-0.8	0	0	116	48	0	0	0	0	1000	3	1000	44	1000	2	1000	0
	0	-0.5	2	23	3	9	1	0	1	0	1000	26	1000	13	1000	39	1000	2
	0	0.5	27	162	1	26	50	27	49	26	1000	146	1000	28	1000	150	1000	38
	0	0.8	47	182	0	36	94	60	96	54	1000	202	1000	51	1000	172	1000	57
0.96	0	0	63	240	0	50	164	137	163	102	1000	216	1000	44	1000	789	1000	76
	-0.5	0	58	234	4	65	157	114	153	93	1000	228	1000	59	1000	303	1000	92
	0.5	0	78	282	0	64	192	130	191	111	1000	288	1000	69	1000	423	1000	136
	0	-0.8	1	4	132	53	1	0	1	0	1000	4	1000	65	1000	6	1000	0
	0	-0.5	22	93	11	35	25	16	26	12	1000	77	1000	23	1000	102	1000	12
	0	0.5	110	317	1	96	148	90	150	89	1000	319	1000	95	1000	232	1000	70
	0	0.8	138	340	5	128	181	121	185	109	1000	361	1000	146	1000	250	1000	111
0.97	0	0	219	450	8	205	387	342	388	250	1000	444	1000	231	1000	888	1000	248
	-0.5	0	216	433	25	229	371	288	368	222	1000	438	1000	213	1000	396	1000	228
	0.5	0	250	490	3	236	424	353	429	282	1000	506	1000	231	1000	527	1000	305
	0	-0.8	9	28	133	84	8	5	9	6	1000	15	1000	72	1000	38	1000	3
	0	-0.5	87	203	15	99	84	44	84	37	1000	206	1000	97	1000	143	1000	38
	0	0.5	293	522	15	288	297	202	295	183	1000	517	1000	280	1000	335	1000	205
	0	0.8	327	522	19	290	336	248	342	222	1000	500	1000	278	998	298	1000	198
0.98	0	0	548	711	105	527	694	656	695	542	1000	681	1000	500	1000	955	1000	483
	-0.5	0	505	687	115	505	639	569	640	476	1000	699	1000	509	1000	534	1000	468
	0.5	0	499	650	75	494	654	597	655	487	1000	694	1000	521	1000	639	1000	515
	0	-0.8	46	76	153	104	53	21	55	20	1000	80	1000	141	999	116	1000	23
	0	-0.5	268	413	73	282	264	164	262	131	1000	440	1000	293	1000	270	1000	146
	0	0.5	541	695	110	514	527	406	530	348	1000	707	1000	533	1000	428	1000	341
	0	0.8	556	713	116	538	527	417	525	375	1000	695	1000	520	999	372	1000	354
0.99	0	0	784	854	464	779	850	822	858	726	1000	850	1000	778	1000	979	1000	729
	-0.5	0	803	852	488	806	877	825	880	723	999	843	1000	785	999	590	1000	703
	0.5	0	786	830	458	778	847	808	848	704	1000	857	1000	794	1000	715	1000	725
	0	-0.8	200	226	218	276	242	107	234	99	1000	237	1000	283	999	233	1000	95
	0	-0.5	626	670	335	607	577	429	574	363	1000	691	1000	627	1000	401	1000	361
	0	0.5	793	859	480	787	772	664	775	574	999	851	1000	774	999	466	1000	557
	0	0.8	775	828	468	774	731	624	730	550	1000	844	1000	782	997	495	1000	576
1	0	0	917	892	878	888	943	929	948	832	992	904	1000	898	991	975	1000	837
	-0.5	0	899	893	889	894	942	910	950	802	994	887	1000	875	992	533	1000	790
	0.5	0	902	891	880	887	936	896	948	819	990	904	1000	894	992	752	1000	812
	0	-0.8	635	400	656	465	603	262	611	216	999	418	1000	481	993	338	1000	233
	0	-0.5	839	758	834	763	826	607	834	520	996	777	1000	764	994	413	1000	513
	0	0.5	894	891	901	894	882	752	888	692	993	911	1000	907	982	533	1000	706
	0	0.8	907	901	907	893	877	715	889	681	987	884	1000	892	962	487	1000	677

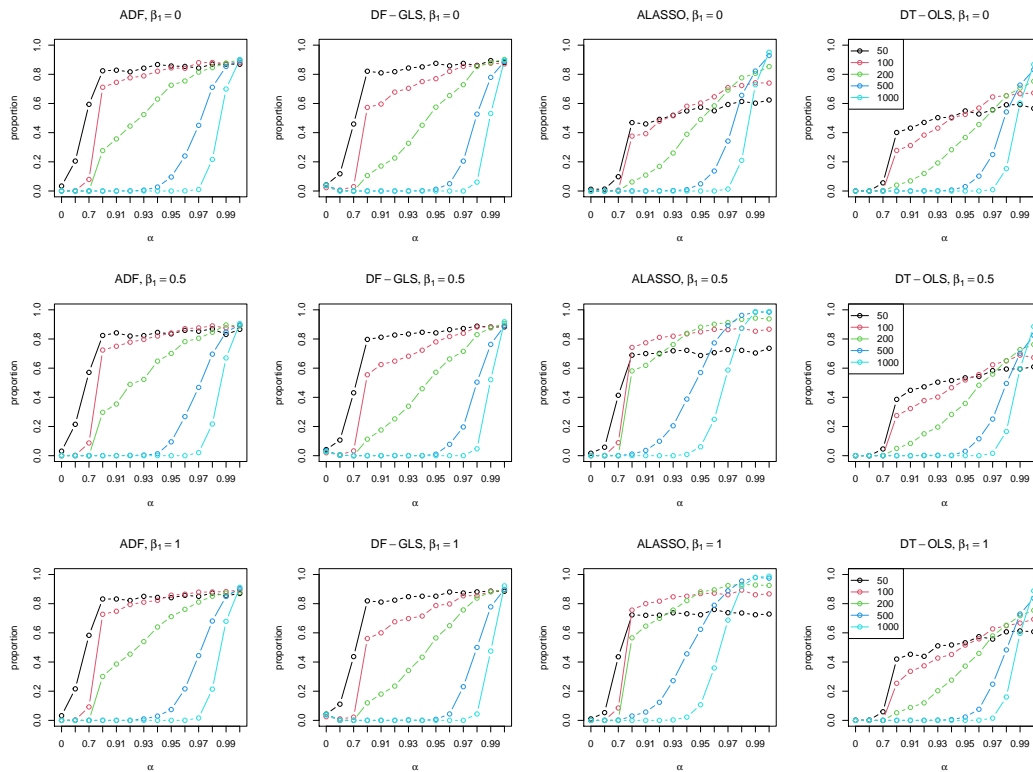


Figure 4: Proportion of unit root non-rejections when $\{v_t\} \sim \text{WN}$.

성과정별로 정리한 것이며, 이로부터 다음과 같은 사실을 파악할 수 있다.

- 원 시계열 $\{y_t\}$ 를 가지고 상수항만 포함된 MC 모형식을 adaptive lasso로 추정하여 단위근의 존재를 판단한 결과와 OLS로 상수항수 C를 이용하여 추세를 제거한 후 adaptive lasso로 단위근을 판단한 결과는 유사하다. 참 추세가 0인 경우에는 $\alpha \leq 0.95$ 또는 $\alpha = 1$ 일 때 단위근을 정확하게 판단하는 비율이 약 90% 이상으로 높다. 이에 비해 참 추세의 기울기가 0이 아닌 경우에는 보통의 단위근 검정과 마찬가지로 α 의 값에 관계없이 모두 단위근이 존재한다고 판단한다.
- 참 추세의 기울기가 0일 때는 $\{y_t\}$ 를 가지고 선형 추세가 포함된 MT 모형식을 추정할 경우에 비해 선형 추세 LT를 추정하여 제거한 $\{y_t^d\}$ 로 (2.3)을 추정한 경우가 단위근이 존재한다고 판단한 비율이 0~10% 정도 낮았다. 그러나 그 차이는 크지 않으며, MC와 C의 경우에 비해 검정력이 나쁘지 않았다.
- 참 추세의 기울기가 0이 아닐 때는 $\{y_t\}$ 로 MT를 적합시킨 경우에 비해 LT 추세를 제거한 $\{y_t^d\}$ 를 사용한 경우에 $\alpha < 1$ 일 때 단위근이 존재하지 않는다고 올바르게 판단한 비율이 상승하였다. 특히 $0.9 \leq \alpha \leq 0.97$ 이고 $\{v_t\}$ 가 WN의 경우에는 약 60% 정도 검정력이 향상되었다.

Figures 4-6는 α 와 단위근이 존재한다고 판단한 비율에 대한 선형 그래프를 n 의 크기별로 구분해서 그린 것으로 “ALASSO”는 $\{y_t\}$ 를 가지고 추세가 포함된 MT 모형식을 적합시켜서 단위근을 식별한 경우이고, “DT-OLS”는 선형 추세 LT를 OLS 방법으로 추정하여 제거한 $\{y_t^d\}$ 를 이용하여 식 (2.3)을 적합시켜서 단위근을 판단한 경우이다. 이 그림들은 다음과 같은 특징을 가진다.

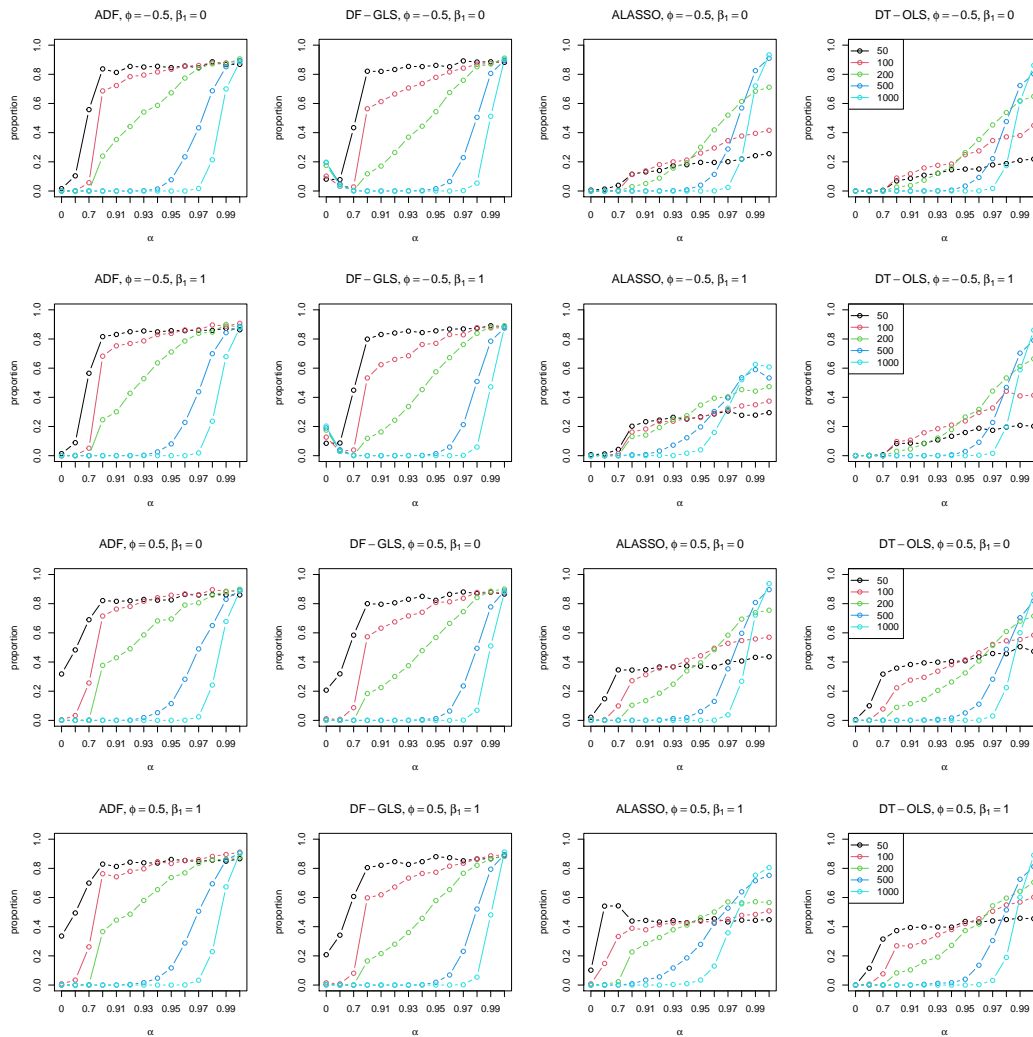


Figure 5: Proportion of unit root non-rejections when $\{v_t\} \sim \text{AR}(1)$.

- 자료의 개수 n 이 커질수록 단위근의 존재를 판단하는 정확도가 높아진다.
- 원 시계열 $\{y_t\}$ 를 사용한 ALASSO보다 추세가 제거된 시계열 $\{y_t^d\}$ 를 사용한 DT-OLS가 높은 검정력을 가진다. 특히 $\{v_t\}$ 가 WN이고 α 가 0.9 초반일 때 검정력이 대폭 상승한다.

3.3. 단위근 검정과의 비교

Adaptive lasso를 이용한 두 가지 방법 외에 추가로 ADF 검정과 DF-GLS 검정과 비교 실험도 실시하였다.

ADF 검정에서는 세 가지의 회귀모형식 중 상수항만 포함된 MC 모형식과 추세까지 포함된 MT 모형식 두 가지를 모두 사용하였으며, 모형식 안에 포함된 차수 k 는 $12(n/100)^{1/4}$ 를 넘지 않는 최대의 자연수를 최대값으로 하여 BIC가 최소가 되도록 선택하였다. 그리고 Elliott 등 (1996)이 제안한 DF-GLS 검정에서는 상수항수 C

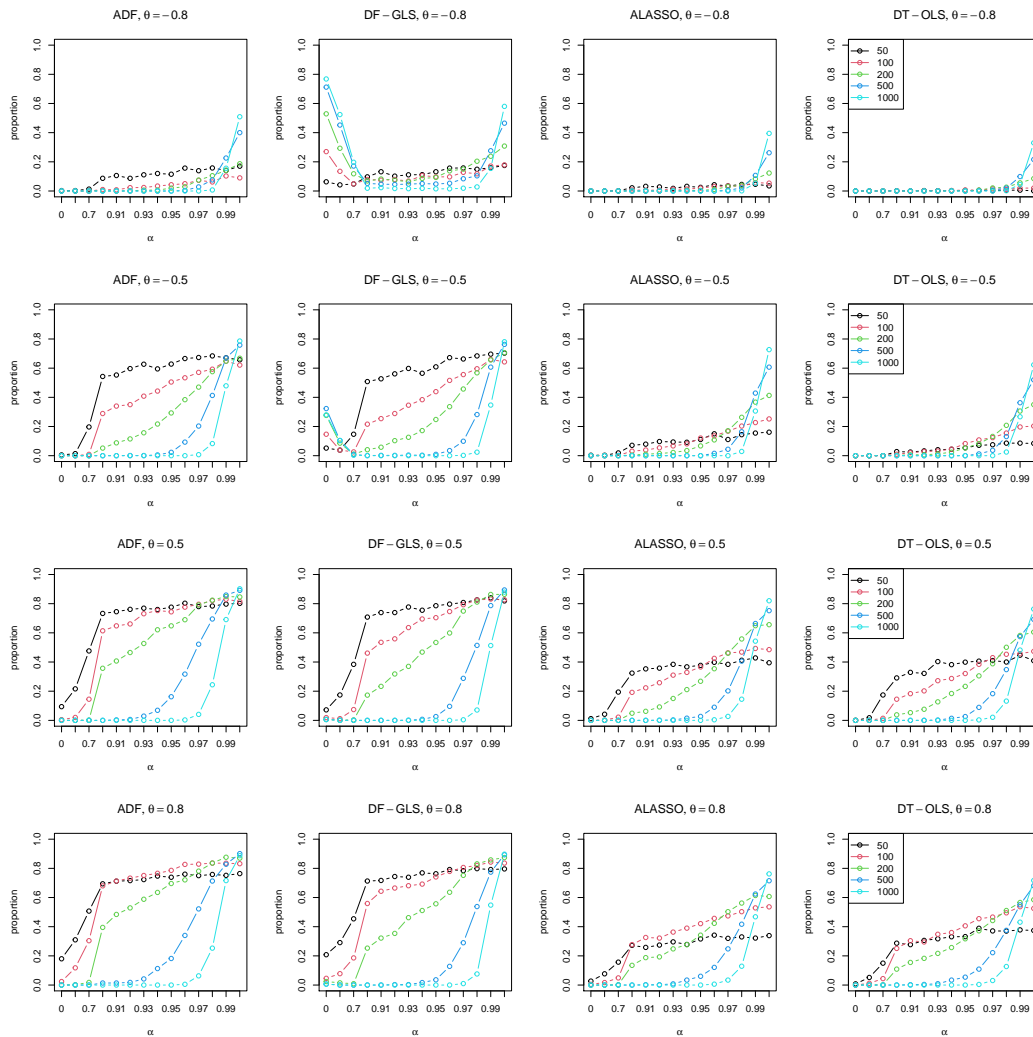


Figure 6: Proportion of unit root non-rejections when $\{v_t\} \sim MA(1)$.

와 1차 함수 LT를 두 가지 방식으로 추세를 제거하였으며, (2.3)의 차수는 ADF 검정에서와 마찬가지로 BIC를 이용하여 선택하였다. 두 가지 단위근 검정에서 유의수준을 모두 10%로 하여 검정을 실시하였으며, 검정 결과 귀무가설을 기각하지 못한, 즉 단위근이 존재한다고 판단한 비율을 Table 3과 Figures 4-6의 “ADF” 부분과 “DF-GLS” 부분에 정리하였다.

Table 3에서 보듯이 원 시계열 $\{y_t\}$ 과 상수항만 포함된 MC 모형식을 가지고 ADF 검정을 실시한 결과와 상수항수 C로 추세를 제거한 시계열로 DF-GLS 검정을 실시한 결과는 ALASSO & MC, DT-OLS & C의 결과와 큰 차이가 없다. 이와 같은 이유로 Figures 4-6은 추세가 포함된 MT 모형식에 기초한 ADF 검정 결과와 1차 함수 LT를 이용하여 선형 추세를 제거한 시계열을 이용한 DF-GLS 검정 결과, ALASSO & MT, DT-OLS & LT에 대한 것만 작성하여 비교하였으며, 다음과 같은 특징을 발견하였다.

- 대부분의 경우 ADF 검정과 DF-GLS 검정에서 단위근이 존재한다고 판단한 비율이 본 논문에서 제한한

Table 4: Nelson & Plosser original data

	n	ALASSO			DT-OLS			ADF	DF-GLS
		AIC	HQC	BIC	AIC	HQC	BIC		
Real GNP	62	0	0	0	0	0	0	-0.075 (-1.751)	-0.036 (-1.080)
Nominal GNP	62	0.048	0.064	0.064	0	0	0	0.029 (1.369)	-0.001 (-0.080)
Real Per Capita GNP	62	-0.095	-0.095	-0.095	-0.108	-0.108	-0.108	-0.184 (-2.984)	-0.109 (-2.164)
Industrial Production Index	111	0.067	0.053	0.053	0	0	0	0.008 (0.344)	0.001 (0.039)
Total Employment	81	-0.140	-0.064	-0.064	-0.156	-0.120	-0.120	-0.155 (-3.044)	-0.147 (-3.013)*
Total Unemployment Rate	81	-0.259	-0.259	-0.144	-0.261	-0.261	-0.141	-0.164 (-2.874)	-0.162 (-2.889)*
GNP deflator	82	0	0	0	0	0	0	-0.021 (-0.727)	-0.012 (-0.520)
Consumer Price Index	111	0	0	0	0	0	0	-0.012 (-0.682)	-0.008 (-0.548)
Nominal Wages	71	0.010	0.012	0.012	0	0	0	0.008 (0.444)	-0.000 (-0.008)
Real Wages	71	-0.022	0	0	-0.028	0	0	-0.089 (-2.246)	-0.039 (-1.345)
Money Stock	82	0.025	0.007	0	0	0	0	-0.001 (-0.047)	-0.007 (-0.934)
Velocity of Money	102	-0.089	-0.089	-0.089	0	0	0	-0.117 (-3.468)**	-0.009 (-0.505)
Bond Yield	71	0	0	0	-0.045	0	0	0.090 (1.872)	-0.024 (-0.614)
Stock Prices	100	-0.067	-0.048	0	-0.071	-0.071	0	-0.023 (-0.878)	-0.019 (-0.830)

DT-OLS & LT 방법으로 단위근이 존재한다고 판단한 비율보다 높게 나타났다.

- Figures 4-5을 보면, $n = 50$ 이고 $\alpha \geq 0.9$ 일 때 ADF 검정과 DF-GLS 검정은 80%보다 조금 높은 비율로 단위근이 존재한다고 판단하였다. 이에 비해 본 논문에서 제안한 DT-OLS 방법은 40~50% 정도의 비율로 단위근이 존재한다고 판단하였다. 비록 비율의 수치에는 차이가 있지만, 네 가지 방법 모두 n 이 작을 때는 $0.9 \leq \alpha \leq 0.99$ 와 $\alpha = 1$ 를 구분하지 못 한다.
- 네 가지 방법 모두 n 이 커질수록 단위근을 식별하는 정확도가 높아진다. 특히 $n \geq 500$ 인 경우에는 네 가지 방법 모두 매우 높은 정확도를 가지며, α 에 따른 그래프의 패턴도 유사하다.
- $n \geq 500$ 일 때, $\theta \leq -0.5$ 인 경우를 제외한 대부분의 경우 DF-GLS 검정과 DT-OLS는 비슷한 결과를 가진다. 두 방법은 ADF 검정과 ALASSO보다 검정력, 즉 $\alpha < 1$ 일 때 단위근이 존재하지 않는다고 판단한 비율이 높다.
- $n = 200$ 일 때는 DT-OLS가 DF-GLS 검정보다 검정력이 조금 높지만, $\alpha = 1$ 일 때 단위근이 존재한다고 올바르게 판단한 비율은 조금 낮다. 그러므로 두 방법 중 어느 것이 더 우수하다고 말하기는 어렵다.
- Figure 5을 보면, $\{v_t\}$ 가 AR인 경우는 단위근 식별 결과가 WN 경우와 좀 다르기는 하지만 그 차이가 크지 않다. 그러나 Figure 6를 보면, $\{v_t\}$ 가 MA인 경우에는 단위근 식별 결과가 WN 경우와 확연하게 다르며, 특히 $\theta = -0.8$ 처럼 작은 경우에는 그 차이가 더 크다. $\theta \leq -0.5$ 인 경우에는 단위근이 존재한다고 판단하는 비율이 다른 경우에 비해 낮으며, $\theta = -0.8$ 인 경우에는 α 의 값에 관계없이 대부분 단위근이 존재하지 않는다고 판단한다.
- $\theta = -0.8$ 일 때 한가지 특이한 현상은 DF-GLS 검정의 경우 $\alpha = 0$ 이고 n 이 클 때 단위근이 존재한다고 판단한 비율이 높게 나타나는 것이다. 다른 세 가지 방법에서는 이런 현상이 발견되지 않는데, 이는 DF-GLS가 추세를 제거할 때 ERS 방법을 사용하기 때문에 생기는 문제로 생각한다.

종합적으로 볼 때, 검정을 이용한 방법과 adaptive lasso를 이용한 방법 중 어떤 것이 더 우수하다고 말하기 어렵다. 기울기가 0이 아닌 선형 추세가 존재하는 경우에는 추세를 제거한 시계열을 사용하는 DF-GLS 검정과 DT-OLS 방법이 원 시계열을 사용하는 경우에 비해 검정력이 좀 더 좋은 것으로 나타나는데, 이는 Elliott 등 (1996)과 같은 결과다. 그리고 α 의 값이 0에 가까운 경우에는 DF-GLS 검정보다는 adaptive lasso를 사용한 방법과 ADF 검정의 검정력이 높게 나타났다.

3.4. 자료 분석

Nelson과 Plosser (1982)가 분석한 자료는 국민총생산, 소비자 물가 지수, 총 실업률 등의 거시경제 시계열 자료로 단위근 검정에 관한 연구에서 자주 사용된다. 본 논문에서도 Nelson과 Plosser (1982)의 자료에 3.2절과 3.3절에서 살펴본 네 가지 방법을 적용하여 단위근의 존재 여부를 판단하였으며, Table 4에 분석 결과를 정리하였다. 14개의 시계열 중 상당수가 시계열 그림을 그렸을 때 증가하거나 감소하는 경향을 가지고 있으며, 추세에 대한 사전 정보가 없으므로 본 연구에서는 선형 추세 $d_t = \beta_0 + \beta_1 t$ 를 이용하여 분석하였다. 표에 주어진 ALASSO 부분의 숫자는 원시계열에 대한 모형식 (2.2)을 adaptive lasso 방법으로 적합시켜서 구한 γ 의 추정값이다. 그리고 DT-OLS 부분의 숫자는 OLS 방법으로 선형 추세를 제거한 뒤 모형식 (2.3)과 adaptive lasso를 이용하여 구한 추정값이다. ADF 부분은 원 시계열 자료를 가지고 ADF 검정을 실시한 결과로 () 앞에 있는 숫자는 모형식 (2.2)와 최소제곱법을 이용하여 구한 γ 에 대한 추정값이고, () 안의 숫자는 ADF 검정통계량의 값이다. 마지막으로 DF-GLS 부분은 ERS 방법으로 선형 추세를 제거한 후 최소제곱법을 이용하여 구한 γ 의 추정값과 DF-GLS 검정통계량의 값이다. ADF 검정과 DF-GLS 검정의 결과에서 *와 **은 각각 유의수준 10%와 5%에서 귀무가설을 기각한 경우, 즉 단위근이 존재하지 않는다고 판단한 경우를 의미하며, adaptive lasso를 이용한 ALASSO와 DT-OLS에서는 추정값이 0일 때 단위근이 존재한다고 판단한다.

분석 결과 ADF 검정은 1개의 시계열을 제외한 나머지 13개의 시계열이 단위근이 존재한다고 판단하였고, ALASSO에서 BIC의 경우는 7개의 시계열이 단위근이 존재한다고 판단하였다. ADF 검정과 ALASSO 두 가지 방법은 모두 원 시계열을 가지고 모형식 (2.2)을 추정함에도 불구하고 6개의 시계열에 대해 단위근의 존재를 판단한 결과가 달랐다. 이는 두 방법에서 모형식을 적합할 때 사용한 추정방법이 최소제곱법과 adaptive lasso로 서로 다르며, 자료의 수가 크지 않고, 모형식에 포함된 두 개의 독립변수 t 와 y_{t-1} 이 상관관계가 강하게 존재하기 때문인 것으로 보인다. 분석 결과 ALASSO에서 t 의 계수 c_1 과 y_{t-1} 의 계수 γ 에 대한 추정값이 동시에 0이거나 또는 동시에 0이 아닌 경우는 드물었는데, 이 또한 t 와 y_{t-1} 의 강한 상관관계가 원인인 것으로 추측되며, 이에 대한 추가 연구가 필요하다.

추세를 제거한 시계열 (y_t^d)를 가지고 단위근의 존재를 판단하는 두 가지 방법인 DT-OLS와 DF-GLS 검정은 Table 4에서 보듯이 결과가 비슷하다. DF-GLS 검정은 14개 중 2개의 시계열이 단위근을 갖지 않는다고 판단하였고, BIC를 이용한 DT-OLS는 3개의 시계열이 단위근을 갖지 않는다고 판단하였다. 비록 1인당 실질 GNP의 경우 DT-OLS와 DF-GLS 검정에 의한 단위근 판단 결과가 달랐지만, γ 에 대한 추정값을 비교해보면 -0.108 과 -0.109 로 거의 비슷하였다.

4. 결론

결정적 추세에 대한 사전 정보가 없는 상태에서는 ADF 검정이나 adaptive lasso로 ADF-회귀모형식을 적합시킬 때 식 (2.2)처럼 선형 추세가 포함된 가장 일반적인 모형식을 사용할 수 밖에 없다. 그러나 이 경우 기존의 연구 결과 및 본 모의실험 결과에서 보듯이 단위근의 존재를 판단하는 검정력이 높지 않았다. 특히 adaptive lasso를 사용한 방법에서는 참 추세의 기울기가 0이 아닌 경우 검정력이 더 낮았다. 본 논문에서는 adaptive lasso를 사용할 때의 검정력을 향상시키기 위한 수정 방안을 연구하였으며, Elliott 등 (1996)이 추세를 제거한 뒤 ADF 검정을 실시한 것처럼 adaptive lasso 방법에서도 추세가 제거된 시계열을 사용하기로 하였다.

모의실험 결과, 기존의 방법처럼 원 시계열 (y_t)에 (2.2)의 모형식을 adaptive lasso로 적합시킬 때보다 추세가 제거된 시계열 (y_t^d)로 (2.3)의 모형식을 적합시킬 때 단위근의 존재를 판단하는 검정력이 높게 나타났다. 비록 자료의 수가 $n \leq 100$ 으로 작은 경우에는 단위근이 존재할 때의 정확도가 ADF 검정과 DF-GLS 검정에 비해 낮지만, n 이 커질수록 그 차이가 줄어들고 $n \geq 500$ 으로 자료의 수가 큰 경우에는 차이가 거의 나지 않

았다. 그리고 $n \geq 200$ 인 경우에는 단위근 검정과 원 시계열을 사용하는 adaptive lasso 방법에 비해 검정력이 높거나 비슷하게 나타났다.

그러나 추세가 제거된 시계열을 사용하더라도 adaptive lasso 방법에서는 자료의 수 n 이 작거나 $\{v_i\}$ 가 백색잡음이 아닌 이동평균 모형을 따르는 경우 단위근이 존재할 때 정확도가 낮다는 문제가 여전히 존재한다. 그리고 adaptive lasso에 사용되는 가중치나 조절모수 선택에 대한 연구도 미흡한 상태이므로 이들에 대한 추가 연구가 필요하다. 마지막으로 본 논문에서는 (1.1)처럼 선형 추세만을 고려하였으나, Elliott 등 (1996)처럼 다양한 추세를 갖는 경우로 확장 가능할 것으로 예상하며, 이에 대한 후속 연구를 계속 진행할 예정이다.

References

- Akaike, H. (1974). A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, **19**, 716–723.
- Dickey, D. and Fuller, W. (1979). Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, **74**, 427–431.
- Elliott, G., Rothenberg, T., and Stock, J. (1996). Efficient tests for an autoregressive Unit Root, *Econometrica*, **64**, 813–836.
- Enders, W. (2010). Applied econometric time series, third edition, John Wiley & Sons.
- Hannan, E. and Quinn, B. (1979). The determination of the order of an autoregression, *Journal of the Royal Statistical Society, Series B*, **41**, 190–195.
- Kwiatkowski, D., Phillips, P. C., Schmidt, P., and Shin, Y. (1992). Testing the null hypothesis of stationarity against the alternative of a unit root, *Journal of Econometrics*, **54**, 159–178.
- Na, O. (2019). Model selection for unstable AR process via the adaptive LASSO, *The Korean Journal of Applied Statistics*, **32**, 909–922.
- Na, O. (2020). Discrimination between trend and difference stationary processes based on adaptive lasso, *The Korean Journal of Applied Statistics*, **33**, 723–738.
- Ng, S. and Perron, P. (2001). Lag length selection and the construction of unit root tests with good size and power, *Econometrica*, **69**, 1519–1554.
- Nelson, C. R. and Plosser, C. I. (1982). Trends and random walks in macroeconomic time series, *Journal of Monetary Economics*, **10**, 139–162.
- Phillips, P. C. and Perron, P. (1988). Testing for a unit root in time series regression, *Biometrika*, **75**, 335–346.
- Said, S. E. and Dickey, D. A. (1984). Testing for unit roots in autoregressive-moving average models of unknown order, *Biometrika*, **71**, 599–607.
- Schwarz, G. (1978). Estimating the dimension of a model, *Annals of Statistics*, **6**, 461–464.
- Schwert, G. W. (1989). Tests for unit roots: A Monte Carlo investigation, *Journal of Business and Economic Statistics*, **7**, 147–160.
- Vougas, D. (2007). GLS detrending and unit root testing, *Economics Letters*, **97**, 222–229.
- Zou, H. (2006). The adaptive lasso and its oracle properties, *Journal of the American Statistical Association*, **101**, 1418–1429.

추세 제거된 시계열을 이용한 단위근 식별

나옥경^{1,a}

^a경기대학교 응용통계학과

요 약

본 논문에서는 adaptive lasso 방법을 이용하여 단위근의 존재 여부를 판단하는 방법에 대해 연구하였다. 최근 원 시계열에 상수항과 선형 추세가 포함된 ADF-회귀모형식을 adaptive lasso로 추정하여 단위근을 식별하는 방법이 제안되었으나, 미지의 선형 추세가 존재할 때 검정력이 떨어지는 것으로 나타났다. 이 문제를 해결하기 위해 본 논문에서는 ADF-회귀모형식을 적합시킬 때 원 시계열 대신 선형 추세가 제거된 시계열을 사용하는 수정안을 제안하였다. 그리고 수정안에서는 일차적으로 선형 추세를 제거한 후 모형식을 적합시키기 때문에 ADF-회귀모형식 중 상수항과 선형 추세를 모두 포함하지 않는 모형식을 사용하였다. 기존의 방법보다 수정안을 사용할 때 단위근의 존재를 판단하는 검정력이 향상되는지 모의실험을 통해 검토하였으며, ADF 검정과 DF-GLS 검정과의 비교 실험도 진행하였다. 모의실험 결과 adaptive lasso를 이용하여 단위근의 존재를 판단할 때 원 시계열보다 추세가 제거된 시계열을 사용하는 경우가 높은 정확도를 가지며, 자료의 개수가 충분히 많을 때 단위근을 잘 판단함을 확인할 수 있었다.

주요용어: 단위근 검정, adaptive lasso 추정, 추세 제거

¹(16227) 경기도 수원시 영통구 광고산로 154-42, 경기대학교 응용통계학과. E-mail: okna@kyonggi.ac.kr