

# Discrimination between trend and difference stationary processes based on adaptive lasso

Okyoung Na<sup>a,1</sup>

<sup>a</sup>Department of Applied Statistics, Kyonggi University

(Received October 14, 2020; Revised October 17, 2020; Accepted October 17, 2020)

---

## Abstract

In this paper, we study a method to discriminate between trend stationary and difference stationary processes. Since a crucial ingredient of this discrimination is to determine the existence of unit root, we can use a unit root testing strategy. So, we introduce a discrimination based on unit root testing and propose the method using the adaptive lasso. Our Monte Carlo simulation experiments show that the adaptive lasso improves the discrimination accuracy when the process is trend stationary, but has lower accuracy than unit root strategy where the process is difference stationary.

Keywords: trend stationary process, difference stationary process, unit root test, adaptive lasso

---

## 1. 서문

국내총생산, 국민소득, 물가, 환율, 이자율과 같은 거시경제 시계열 자료를 분석할 때는 (Elder와 Kennedy, 2001)의 설명에서 보듯이 결정적 추세와 정상시계열의 합으로 이루어진 시계열 모형과 확률보행과정처럼 정상시계열의 누적합으로 만들어지는 시계열 모형을 모두 고려할 수 있다. 전자에 해당하는 시계열 모형은 추세-정상시계열(trend stationary process)이라고 부르며, 다음 식을 만족한다.

$$y_t = g(t) + z_t, \quad g: \mathbb{R} \rightarrow \mathbb{R}, \{z_t\} \text{는 정상시계열이다.} \quad (1.1)$$

여기서  $g(t)$ 는 결정적 추세를 나타내는 부분으로 상수함수, 선형함수, 이차함수 등 다양한 형태를 고려할 수 있다. 그리고 후자는

$$(1 - B)^d y_t = z_t, \quad d \in \mathbb{N}, B \text{는 후향연산자, } \{z_t\} \text{는 정상시계열} \quad (1.2)$$

을 만족하는 경우로 차분-정상시계열(difference stationary process)이라고 한다.

기존 연구 결과들을 보면, 1970년대까지는 거시경제 시계열 자료를 분석할 때 주로 추세-정상시계열을 사용하였다. 그러나 Nelson과 Plosser (1982)가 국내총생산, 소비자 물가 지수 등의 거시경제 시계열 자료를 이용하여 단위근 검정을 실시하였고, 그 결과 상당수의 거시경제 시계열이 추세-정상성이 아닌 차분-정상성을 만족한다고 주장하였으며, 이후 차분-정상시계열에 대한 관심이 높아졌다. Nelson과 Plosser (1982)은 시계열 자료를 분석할 때 Dickey와 Fuller (1979)가 제안한 단위근 검정을 사용하였으나, 이 검정방법은 검

---

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

정력이 낮다는 단점을 가지고 있다. 즉 추세-정상시계열을 단위근이 존재하는 차분-정상시계열로 잘못 판단할 확률이 높다. 그리고 정책의 변화나 경제 위기처럼 시계열의 특성에 영향을 주는 사건이 존재하면, 단위근 검정 결과 차분-정상시계열로 판단할 가능성이 높아진다. 그러므로 Perron (1988), DeJong과 Whiteman (1991), Li (1995), Charles와 Darné (2012) 등은 Nelson과 Plosser (1982)에서 분석했던 시계열 자료를 새로운 기법으로 재분석하였으며, 최근까지도 여러 연구자들이 추세-정상성과 차분-정상성을 비교 검토하는 연구를 수행하고 있다.

실제로 Brockwell와 Davis (2006), Enders (2010), Elder와 Kennedy (2001) 등이나 그 안에 포함된 다양한 연구 결과들을 살펴 보면, 추세-정상시계열과 차분-정상시계열은 분산과 자기상관함수, 충격의 지속성 여부, 정상화 방법 등이 서로 다르다. 그러므로 시계열의 선택에 따라 자료를 분석하는 방법, 추정한 모형식을 해석하는 방법, 그리고 미래의 값을 예측하는 방법까지 모두 달라진다. 따라서 주어진 시계열 자료에 대한 추세-정상성과 차분-정상성을 판단하는 것은 매우 중요한 문제라고 할 수 있으며, 본 논문에서도 이 문제에 대해 연구하고자 한다.

식 (1.1)과 (1.2)에서 보듯이 추세-정상시계열과 차분-정상시계열의 가장 큰 차이점은 단위근의 존재 여부이다. 그러므로 Nelson과 Plosser (1982), Perron (1988) 등의 연구처럼 단위근 검정 결과를 이용하여 두 시계열을 분류할 수 있다. Dickey와 Fuller (1979)가 DF 검정을 제안한 이후로 Said와 Dickey (1984), Phillips와 Perron (1988), Kwiatkowski 등 (1992), Elliott 등 (1996), Enders (2010), Elder와 Kennedy (2001) 등과 같은 많은 연구자들이 다양한 단위근 검정방법과 단위근 검정을 수행하는 전략들을 연구, 제안하였다. 본 논문에서는 이 중에서 Enders (2010)가 제안한 단위근 검정 전략과 Elder와 Kennedy (2001)의 전략을 시계열을 분류하는데 이용하였으며, 이에 대해서는 2.2절에 자세하게 소개할 것이다.

Chen과 Chan (2011), Kwon 등 (2017), Na (2019) 등의 최근 연구 결과를 보면, 자기회귀누적이동평균 (ARIMA) 모형을 이용하여 시계열 자료를 분석할 때 adaptive lasso와 같은 벌점화 기법을 사용하고 있다. 특히 Na (2019)는 ARIMA 모형의 차분차수를 결정할 때, 즉 단위근의 존재 여부를 결정할 때 adaptive lasso를 사용하였으며, 단위근 검정에 비해 판단 결과가 나쁘지 않음을 모의실험을 통해 규명하였다. 본 논문에서는 이와 같은 연구결과를 추세-정상성과 차분-정상성을 판단하는 문제로 확장, adaptive lasso를 활용하여 시계열을 분류하는 방법에 대해 연구하였으며, 이에 대해서는 2.3절과 3장에서 상세하게 설명할 것이다.

## 2. 시계열의 분류

### 2.1. 모형

본 논문에서는 시계열  $\{y_t\}$ 가 다음 모형식을 만족한다고 가정한다:

$$y_t = a + bt + z_t, \quad a, b \in \mathbb{R}, t \in \mathbb{N}. \quad (2.1)$$

여기서  $\{(1-B)^d z_t\}$ 는 평균이 0이고, 인과성, 정상성, 가역성을 모두 만족하는 ARMA( $p, q$ ) 과정이다. 차분 차수  $d$ 는 이론적으로 0 이상의 정수이지만, 현실적으로 2 이상인 경우가 많지 않다. 그러므로 여기서는  $d$ 의 값으로 0과 1만 고려한다. 그리고 결정적 추세는  $a + bt$ 와 같은 선형 함수뿐만 아니라 2차나 3차 다항식, 삼각함수, 지수함수와 같은 형태일 수도 있으나, 실제 시계열 자료를 분석할 때 선형 추세를 가장 많이 사용하므로 본 논문에서는 식 (2.1)처럼 선형 함수만을 고려한다.

식 (2.1)에 주어진 모형은 기울기  $b$ 와 차분차수  $d$ 의 값에 따라 (1)  $b = 0, d = 0$ , (2)  $b \neq 0, d = 0$ , (3)  $d = 1$ 을 만족하는 세 가지 경우로 나누어지며, 각각의 경우 다른 정상화 방법을 가진다.

첫째,  $b = d = 0$ 일 때 식 (2.1)에 주어진  $\{y_t\}$ 는 평균이  $a$ 인 정상시계열이 되며, 정상화 과정이 필요없다. 그리고 시계열 그림이 수평적 락 형태이므로, 본 논문에서는 이 경우를 “level stationary (LS)”로 표기한다.

다음은  $b \neq 0, d = 0$ 인 두 번째 경우다. 이 경우 식 (2.1)은 결정적 선형 추세와 평균이 0인 정상시계열의 합으로 이루어진 추세-정상시계열이 되므로, “trend stationary (TS)”라고 표기한다. 시계열의 평균을 계산하면,  $E(y_t) = a + bt$  ( $b \neq 0$ )으로 시점에 따라 그 값이 달라진다. 그러므로 LS 경우와 다르게  $\{y_t\}$ 는 비정상시계열이며, 분석을 할 때 정상화 과정이 필요하다. TS-시계열을 정상시계열로 만들 때 주로 사용하는 방법은 추세제거(detrending), 즉 추세를 추정하여 원 시계열에서 빼는 것이다. 차분을 통해서도  $\{y_t\}$ 를 정상시계열로 만들 수 있으나, 이 때에는 과대차분의 문제가 발생한다. 그러므로 정상화 방법으로 추세제거를 선호한다.

마지막으로  $d = 1$ 인 경우는 앞에서 고려한 두 가지 경우와는 다르게  $\{z_t\}$ 가 ARIMA( $p, 1, q$ ) 과정, 즉 비정상시계열이 된다. 그러므로 기울기  $b$ 의 값에 관계없이  $\{y_t\}$ 는 항상 비정상시계열이 되고, 추세제거로는  $\{y_t\}$ 를 정상시계열로 만들 수 없다. 대신 1차 차분한 시계열  $\{(1 - B)y_t\}$ 은 평균이  $b$ 인 정상시계열이 되므로, 본 논문에서는 이 경우를 “difference stationary (DS)”라고 표시한다. DS-시계열의 경우  $b$ 의 값에 따라 경향(drift)이 존재하는 경우( $b \neq 0$ )와 그렇지 않은 경우( $b = 0$ )로 세분할 수 있으나, 정상화 방법에 차이가 없으므로 본 논문에서는  $b$ 의 값에 따른 분류까지는 고려하지 않겠다.

## 2.2. 단위근 검정을 이용한 모형 분류

이제 시계열 자료  $y_1, y_2, \dots, y_n$ 이 주어졌을 때, 앞서 정의한 LS, TS, DS 세 가지 중 하나로 시계열을 분류하는 문제를 생각해 보자. 기본적으로 이 문제는 식 (2.1)에서 단위근의 존재 여부를 판단하는 것과 같다. 그러므로 시계열 분류 문제에 단위근 검정을 활용할 수 있으며, adaptive lasso를 이용한 분류 방법에 대해 설명하기 전에 단위근 검정을 이용한 방법부터 소개한다.

**2.2.1. ADF 회귀모형** 단위근 검정에는 DF 검정 (Dickey와 Fuller, 1979), ADF 검정 (Said와 Dickey, 1984), PP 검정 (Phillips와 Perron, 1988), KPSS 검정 (Kwiatkowski 등, 1992), DF-GLS 검정 (Elliott 등, 1996) 등 다양한 검정 방법이 존재하며, 이 중에서 ADF 검정이 가장 대표적이다.

ADF 검정에서는 다음의 세가지 회귀모형을 사용한다:

$$\Delta y_t = \gamma y_{t-1} + \sum_{i=1}^k \delta_i \Delta y_{t-i} + \epsilon_t; \quad (2.2)$$

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

$$\Delta y_t = \gamma y_{t-1} + \alpha + \beta t + \sum_{i=1}^k \delta_i \Delta y_{t-i} + \epsilon_t. \quad (2.4)$$

위 식에서  $\Delta$ 은 1차 차분  $(1 - B)$ 을 의미하며,  $k$ 는 자료의 개수  $n$ 에 대한 함수로

$$\lim_{n \rightarrow \infty} n^{-\frac{1}{3}}k = 0, \quad k > r_0 n^{r_1} \quad (\text{단, } r_0, r_1 > 0) \quad (2.5)$$

을 만족하는 자연수이다. 식 (2.1)은 식 (2.2)–(2.4)으로 변환할 수 있다. 다만 식 (2.1)의  $\{z_t\}$ 에 대한 가정으로 인해 모형식 (2.4)의 오차항  $\{\epsilon_t\}$ 는 Dickey와 Fuller (1979)의 경우나 일반적인 회귀모형처럼 백색잡음이 아닐 수 있으며, 이 문제를 해결하기 위해 식 (2.5)에 주어진  $k$ 를 사용한다. 자세한 사항은 Said와 Dickey (1984)을 참조하기 바란다.

식 (2.1)과 위에 주어진 ADF 회귀모형식과의 관계를 좀 더 구체적으로 살펴보면, 식 (2.1)이 LS, TS, DS 중 무엇이나에 따라 변환 결과가 달라진다.

- 먼저 LS의 경우는 식 (2.2) 또는 식 (2.3)의 경우로 변환되며  $\gamma$ 가 음수가 된다. 즉,  $\gamma < 0, \beta = 0$ 인 식 (2.4)로 변환된다.
- TS는  $\gamma < 0$ 과  $\beta \neq 0$ 를 만족하는 식 (2.4)로 변환된다.
- DS 형태의 시계열은 식 (2.2) 또는 식 (2.3)으로 변환되며  $\gamma = 0$ 이다. 다시 말해서  $\gamma = 0, \beta = 0$ 인 식 (2.4)로 변환된다.

따라서 우리는 주어진 시계열 자료를 LS, TS, DS로 분류할 때 식 (2.1) 대신 식 (2.4)를 활용할 수 있다.

*Remark 2.1:* 식 (2.4)에서 이론적으로  $\gamma > 0$ 인 경우나  $\gamma = 0, \beta \neq 0$ 인 경우도 생각할 수 있다. 그러나, 이 경우는 생성된 자료를 가지고 시계열 그림을 그렸을 때 폭발적으로 증가하거나 감소하는 형태로 주어져서, 현실적으로 그 빈도가 높지 않다.

**2.2.2. Enders의 전략** ADF 검정은 세 가지 회귀모형식 (2.2)–(2.4) 중에서 적절한 하나의 모형식을 선택하여 회귀분석을 시행하고 구한 검정통계량

$$\tau = \frac{\tilde{\gamma}}{SE(\tilde{\gamma})} \quad (\text{단, } \tilde{\gamma} \text{와 } SE(\tilde{\gamma}) \text{는 } \gamma \text{에 대한 최소제곱추정량과 표준오차를 의미}) \quad (2.6)$$

의 값을 바탕으로  $H_0: \gamma = 0$ 와  $H_1: \gamma < 0$ 을 검정한다. 비록 식 (2.4)가 식 (2.2)과 식 (2.3)을 포함하는 가장 일반적인 모형식이지만, Dickey와 Fuller (1979)의 연구 결과에서 보듯이 검정통계량  $\tau$ 의 값과 임계값이 식 (2.2)–(2.4) 중 어느 모형식을 선택하느냐에 따라 달라진다. 만약 추세에 대한 사전정보가 있다면 Elder와 Kennedy (2001)처럼 ADF 검정에 사용할 회귀모형식을 선택할 수 있으나, 실제 시계열 자료를 분석할 때는 사전정보가 없는 경우도 많다. 활용할 수 있는 사전정보가 없을 때에는 Enders (2010)의 전략처럼 가장 일반적인 경우인, 추세를 고려한 식 (2.4)부터 시작하여 추세와 상수항이 모두 없는 식 (2.2)까지의 모형을 순차적으로 이용, ADF 검정을 반복적으로 실시하여 단위근의 존재 여부를 파악한다.

시계열 자료가 주어지면, Hacker와 Hatemi-J (2010)처럼 Enders의 전략을 이용하여 LS, TS, DS 중 하나로 시계열을 분류할 수 있으며, 구체적인 분류절차는 다음과 같다.

(A) 주어진 시계열 자료  $y_1, y_2, \dots, y_n$ 를 가지고 식 (2.4)의 회귀모형식을 적합시킨다. 그리고 그 결과를 바탕으로 다음에 주어진 (A1)–(A5) 단계를 수행한다.

- (A1) 가설  $H_0: \gamma = 0$ 와  $H_1: \gamma < 0$ 에 대한 ADF 검정을 실시한다. 검정통계량은 식 (2.6)에 주어진  $\tau$ 이며, 검정 결과 귀무가설이 기각되면 (A2)로 이동하고, 그렇지 않으면 (A3)로 이동한다.
- (A2) 식 (2.4)의 모형식에서 기울기  $\beta$ 의 유의성에 대한  $t$ -검정을 실시한다. 검정 결과  $H_0: \beta = 0$ 이 채택되면, 시계열을 LS로 분류한다. 그리고  $H_1: \beta \neq 0$ 이 채택되면 시계열을 TS로 분류하고, 분류 절차를 멈춘다.
- (A3) 식 (2.4)에서  $H_0: \gamma = \beta = 0$ 에 대한 우도비(likelihood ratio; LR) 검정을 실시한다. 검정통계량으로 Dickey와 Fuller (1981)의  $\Phi_3$ 를 사용한다. 검정 결과 귀무가설이 기각되면 (A4)로 이동하고,

그렇지 않으면 (B)로 이동한다.

- (A4) 1차 차분한 자료를 이용하여 회귀모형식  $\Delta y_t = \alpha + \beta t + \sum_{i=1}^k \delta_i \Delta y_{t-i} + \epsilon_t$ 을 적합시키고,  $H_0: \beta = 0$ 에 대한  $t$ -검정을 실시한다. 만약  $t$ -검정 결과  $H_0: \beta = 0$ 가 기각되면 (A5)로 이동하고, 그렇지 않으면 (B)로 이동한다.
- (A5) (A1)에서 구한 검정통계량  $\tau$ 를 가지고  $H_0: \gamma = 0$ 와  $H_1: \gamma < 0$ 에 대한  $t$ -검정을 실시한다. 검정 결과 귀무가설을 기각하면 시계열을 TS로 분류하고, 그렇지 않으면 DS로 분류한 후 분류 절차를 멈춘다.

(B) 회귀모형식 (2.3)을 적합시키고, 아래의 (B1)–(B4) 단계를 실시한다.

- (B1)  $H_0: \gamma = 0$ 와  $H_1: \gamma < 0$ 에 대한 ADF 검정을 실시한다. 검정 결과 귀무가설이 기각되면, 시계열을 LS로 분류하고 분류 절차를 멈춘다. 만약 귀무가설이 채택되면, (B2)로 이동한다.
- (B2) Dickey와 Fuller (1981)의  $\Phi_1$ 을 이용하여  $H_0: \gamma = \alpha = 0$ 에 대한 LR-검정을 실시한다. 그 결과 귀무가설을 기각하면 (B3)로 이동하고, 그렇지 않으면 (C)로 이동한다.
- (B3) 1차 차분한 자료를 이용하여 회귀모형식  $\Delta y_t = \alpha + \sum_{i=1}^k \delta_i \Delta y_{t-i} + \epsilon_t$ 을 적합시키고,  $H_0: \alpha = 0$ 에 대한  $t$ -검정을 실시한다. 만약  $t$ -검정 결과  $H_0: \alpha = 0$ 가 기각되면 (B4)로 이동하고, 그렇지 않으면 (C)로 이동한다.
- (B4) (B1)에서 구한 ADF 검정통계량  $\tau$ 를 이용하여  $H_0: \gamma = 0$ 에 대한 단측  $t$ -검정을 실시한다. 검정 결과 귀무가설을 기각하면 시계열을 LS로 분류하고, 그렇지 않으면 DS로 분류한 후 분류 절차를 멈춘다.

(C) 식 (2.2)에 주어진 회귀모형식을 적합시키고, 그 결과를 바탕으로 ADF 검정을 실시한다. 그 결과  $H_0: \gamma = 0$ 를 기각하면 시계열을 LS로 분류하고, 그렇지 않으면 DS로 분류한다.

**2.2.3. Elder-Kennedy의 전략** 앞서 Enders의 전략을 이용한 분류절차를 소개하였다. Enders의 단위근 검정 전략은 여러 단계에 걸쳐서 검정을 반복적으로 수행하기 때문에 그 절차가 복잡하다. Elder와 Kennedy (2001)는 이 점을 단점으로 지적하면서 현실성을 고려, 추세에 대한 사전정보가 없는 경우라도 식 (2.4)에 대한 ADF 검정을 한번만 실시한 것을 제안하였다. Elder와 Kennedy (2001)가 제안한 분류 절차는 다음과 같다.

- (D1) 시계열 자료  $y_1, y_2, \dots, y_n$ 을 가지고 (2.4)의 모형식을 적합시키고, ADF 검정을 실시한다. 검정 결과  $H_0: \gamma = 0$ 를 기각하면 (D2) 단계로 이동하고, 그렇지 않으면 시계열을 DS로 분류, (D3)로 이동한다.
- (D2) 식 (2.4)의 모형에서  $H_0: \beta = 0$ 에 대한  $t$ -검정을 실시한다. 검정 결과 귀무가설을 기각하면 시계열을 TS로 분류하고, 그렇지 않으면 LS로 분류한다.
- (D3) 1차 차분한 시계열의 평균이 0인지 검정하여, (D1)에서 DS로 구분된 시계열이 경향(drift)을 가지는지 여부를 추가로 판단한다.

### 2.3. Adaptive lasso를 이용한 모형 분류

2.2절에서 식 (2.1)에 주어진 시계열을 LS, TS, DS 중 하나로 구분하여 분류할 때 ADF 회귀모형식 (2.4)를 활용할 수 있음을 설명하였다. 그리고 Enders의 전략을 이용한 분류절차와 Elder와 Kennedy의 분류절차에 대해 알아보았다. 이들 두 가지 분류절차는 모두 단위근 검정과 회귀계수에 대한 검정을 기초로 만들어진

것이며, 순차적으로 시행한 검정 결과들을 바탕으로  $\gamma = 0$ 와  $\beta = 0$  여부를 판단하였다.

본 논문에서는 이와 다르게 모형식 (2.4)의 계수  $\theta = (\alpha, \beta, \gamma, \delta_1, \dots, \delta_k)^T$ 를 adaptive lasso 방법으로 추정,  $\gamma$ 와  $\beta$ 에 대한 추정값을 바탕으로 검정을 거치지 않고 직접 0인지 여부를 판단하려 한다.

조절모수  $\lambda \in \mathbb{R}_+$ 와 가중치 벡터  $w = (w_\beta, w_\gamma, w_1, \dots, w_k)^T \in \mathbb{R}_+^{k+2}$ 가 주어지면, 모수  $\theta$ 에 대한 adaptive lasso 추정량을 다음처럼 구할 수 있다:

$$\hat{\theta}(\lambda, w) = \arg \min_{\theta \in \mathbb{R}^{k+3}} \left\{ \frac{SSR(\theta)}{2N} + \lambda \left( w_\beta |\beta| + w_\gamma |\gamma| + \sum_{i=1}^k w_i |\delta_i| \right) \right\}, \quad (2.7)$$

여기서  $\mathbb{R}_+$ 는 양의 실수 집합을 의미하고,  $N$ 과  $SSR(\theta)$ 는 각각

$$N = n - k - 1, \quad SSR(\theta) = \sum_{t=k+2}^n \left\{ \Delta y_t - \left( \alpha + \beta t + \gamma y_{t-1} + \sum_{i=1}^k \delta_i \Delta y_{t-i} \right) \right\}^2$$

이다. 위 식에서 보듯이 조절모수와 가중치 벡터의 값은 adaptive lasso 추정값에 영향을 주므로 선택이 중요하다. Zou (2006), Chen과 Chan (2011), Kwon 등 (2017), Na (2019) 등에서 제안한 것처럼 보통 가중치 벡터는  $\theta$ 에 대한 최소제곱추정량이나 능선회귀추정량과 같은 일치추정량의 역수를 활용하여 결정하며, 조절모수는 정보함수나 교차검증을 이용하여 선택할 수 있다.

본 논문에서는  $\theta$ 에 대한 최소제곱추정량  $\tilde{\theta} = (\tilde{\alpha}, \tilde{\beta}, \tilde{\gamma}, \tilde{\delta}_1, \dots, \tilde{\delta}_k)^T$ 의 역수의 절대값, 즉

$$\tilde{w} = \left( \frac{1}{|\tilde{\beta}|}, \frac{1}{|\tilde{\gamma}|}, \frac{1}{|\tilde{\delta}_1|}, \dots, \frac{1}{|\tilde{\delta}_k|} \right)^T \quad (2.8)$$

을 가중치 벡터로 사용하였다. 그리고 조절모수  $\lambda$ 의 값은 정보함수

$$IC(\lambda) = \log \left\{ \frac{SSR(\hat{\theta}(\lambda, \tilde{w}))}{N} \right\} + \frac{q}{N} \|\hat{\theta}(\lambda, \tilde{w})\|_0 \quad (2.9)$$

가 최소가 되도록 선택하였다. 여기서  $\|\hat{\theta}(\lambda, \tilde{w})\|_0$ 는  $\hat{\theta}(\lambda, \tilde{w})$ 의 원소 중 0이 아닌 것의 개수, 즉 유의한 변수의 개수를 의미한다. 그리고  $q$ 는 변수의 개수에 대한 벌점상수로 본 논문에서는  $q$ 의 값으로 세 가지 값  $q = 2$ ,  $q = 2\log\log(N)$ ,  $q = \log(N)$ 을 고려하였으며, 이는 각각 AIC, HQC, BIC에 해당한다.

식 (2.8)에 주어진 가중치 벡터  $\tilde{w}$ 와 식 (2.9)의 정보함수를 이용하여 선택한 조절모수  $\hat{\lambda} = \operatorname{argmin}_\lambda IC(\lambda)$ 에 대응하는 adaptive lasso 추정량, 즉  $\hat{\theta}(\hat{\lambda}, \tilde{w})$ 을 간단히  $\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\gamma}, \hat{\delta}_1, \dots, \hat{\delta}_k)^T$ 라고 표기하자. 우리는  $\theta$ 의 원소 중  $\hat{\gamma}$ 와  $\hat{\beta}$ 의 값에 따라 시계열을 다음처럼 분류할 수 있다:

- 만약  $\hat{\gamma} \neq 0$ ,  $\hat{\beta} = 0$ 이면, 분석에 사용한 시계열을 LS로 분류한다;
- 만약  $\hat{\gamma} \neq 0$ ,  $\hat{\beta} \neq 0$ 이면, 시계열을 TS로 분류한다;
- 만약  $\hat{\gamma} = 0$ 이면, 시계열을 DS로 분류한다.

*Remark 2.2:* 앞서 설명한 것처럼 식 (2.1)을 만족하는 시계열은 식 (2.4)의 형태로 변환했을 때 결코  $\gamma = 0$ ,  $\beta \neq 0$  일 수 없다. 그러나 실제로 식 (2.1)에 의해 생성된 시계열 자료를 가지고 계산한 추정값이  $\hat{\gamma} = 0$ ,  $\hat{\beta} \neq 0$ 인 경우가 발생할 수 있으며, 이는 LS, TS, DS 중 어느 것에도 해당하지 않는다. Elder와 Kennedy

(2001)도 이와 같은 문제점을 인식하였으며, 이에 대한 해결책으로 단위근 검정을 시행,  $\gamma = 0$ 인지 여부를 판단하여 시계열을 DS와 나머지 경우로 1차 분류할 것을 제안하였다. 본 논문에서도 Elder와 Kennedy (2001)의 제안을 기초로 하여 추정값이  $\hat{\gamma} = 0, \hat{\beta} = 0$ 인 경우 뿐만 아니라  $\hat{\gamma} = 0, \hat{\beta} \neq 0$ 인 경우까지, 즉  $\hat{\gamma} = 0$ 인 경우를 DS로 분류하였다.

### 3. 모의실험

#### 3.1. Adaptive lasso 추정량에 기초한 분류의 정확도 실험

우선 본 논문에서 제안한 것처럼 adaptive lasso 추정값을 기초로 시계열의 정상성, 즉 LS, TS, DS를 구분하였을 때의 정확도를 모의실험을 통해 살펴보고자 한다. 이를 위해 다음과 같은 모형을 고려하였다:

$$y_t = a + bt + z_t, \quad t = 1, 2, \dots, n, \quad (3.1)$$

$$z_t = \phi z_{t-1} + \epsilon_t, \quad \epsilon_t \stackrel{iid}{\sim} N(0, 1). \quad (3.2)$$

이는 Dickey와 Fuller (1979)부터 현재까지 단위근 검정 연구에서 고려하는 시계열 모형 중 가장 기본적인 모형으로,  $b$ 와  $\phi$ 의 값에 따라 LS, TS, DS로 나누어진다. 위 모형식을 식 (2.4)의 형태로 변환하면,

$$\Delta y_t = (\phi - 1)y_{t-1} + (1 - \phi)a + \phi b + (1 - \phi)bt + \epsilon_t$$

이므로  $\gamma = \phi - 1, \beta = (1 - \phi)b$ 가 성립한다. 따라서  $a$ 의 값보다는  $b$ 와  $\phi$ 의 값이 분류 결과에 영향을 더 끼치며, 특히  $\phi$ 의 값이 1에 가까울수록 분류의 정확도가 떨어질 것임을 예상할 수 있다. 이와 같은 이유로 본 모의실험에서는  $a = 0.0, b = 0.0, 0.1, 0.5, 1.0, \phi = -0.9, -0.5, 0.0, 0.5, 0.8, 0.9 \sim 1.0$ 의 경우를 고려하였고,  $\phi$ 의 값 중 0.9부터 1.0의 구간은 0.01 간격으로 세분하여 살펴보았다.

각각의 경우, 위에 주어진 모형 (3.1)–(3.2)을 이용하여  $n = 50, 100, \dots, 5000$ 개의 자료를 생성하였다. 다만 식 (3.2)에 주어진  $z_t$ 를 생성할 때 초기값으로 0을 사용하였으며, 초기치 영향을 줄이기 위해 처음 100개의 값은 사용하지 않았다. 이렇게 생성한 자료를 가지고 2.3절에 설명한 방식대로 adaptive lasso 추정값을 계산하였고, 이 때  $k$ 는 Schwert (1989)의 제안처럼  $12(n/100)^{1/4}$ 을 넘지 않는 최대 정수를 사용하였다. 조절 모수  $\lambda$ 는 Friedman 등 (2010)의 방식대로 선택,  $\lambda$ 의 최소값을  $\lambda_{min} = 10^{-4} \min(1, \lambda_{max})$ 처럼 변경하여 사용하였다. 각각의 경우에 1,000번씩 반복 실험을 수행하였고, adaptive lasso 추정값을 기초로 시계열을 분류한 결과를 Figure 3.1–3.3처럼 정리하였다.

먼저 Figure 3.1은  $b = 0$ 일 때 1,000번 중 LS로 분류한 비율(단위, %)을 나타낸다.  $\phi$ 와 LS로 분류한 비율에 대한 선형 그래프를 정보함수 종류와 자료의 개수  $n$ 에 따라 구분하여 그린 그림으로, 위쪽 세 개는  $-1 < \phi \leq 1$ 의 구간에 대한 그림이고, 아래쪽 세 개는  $0.9 \leq \phi \leq 1$ 의 구간만을 확대한 그림이다. 식 (3.1)에서  $a = b = 0$ 이므로  $y_t = z_t$ 가 성립한다. 따라서  $\phi < 1$ 일 때는 LS가 참이고,  $\phi = 1$ 일 때는 DS가 참이다. 그러므로 그림에서  $\phi < 1$ 인 구간에서는 그래프의 높이가 높을수록, 그리고  $\phi = 1$ 일 때는 높이가 낮을수록 분류의 정확도가 높다고 생각할 수 있다. 그리고 Figure 3.3의 1행에 있는 그림은  $b = 0, \phi = 0.98, 0.99, 1.0$ 일 때 1,000번의 실험 결과 중 TS로 분류한 비율(단위, %)과 DS로 분류한 비율(단위, %)에 대한 선형 그래프를 겹쳐서 그린 것이다. 이 그림들로부터  $b = 0$ 일 때 다음과 같은 특징을 가짐을 알 수 있다.

- 자료의 개수가 커질수록 분류 정확도가 높아진다.
- $\phi \leq 0.9$ 인 경우에는 자료의 개수와 정보함수에 관계없이 분류 정확도가 80% 이상으로 높은 편이다. 특히 HQC와 BIC의 경우 자료의 개수가 100 이상일 때 100%에 가까운 정확도를 가진다.

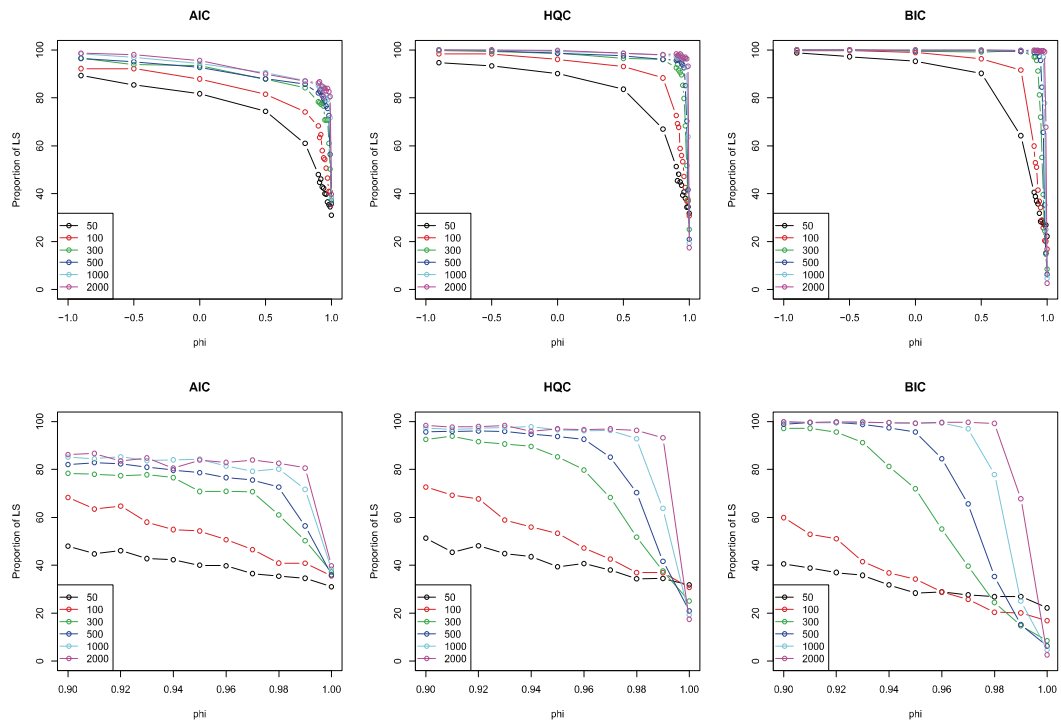


Figure 3.1. Proportion of LS when  $a = b = 0$ .

- $\phi \leq 0.9$ 이거나  $\phi = 1$ 일 때, 자료의 수와  $\phi$ 의 값이 같으면, AIC의 정확도가 가장 낮고 BIC의 정확도가 가장 높다.
- 자료의 수와 정보함수가 같을 때,  $\phi$ 의 값이 1에 가까워질수록 LS로 올바르게 분류하는 비율이 낮아지고, DS로 잘못 분류하는 비율이 높아진다. 이는  $\phi$ 의 값이 1에 가까울수록  $\gamma = \phi - 1$ 이 0에 가까워지기 때문이다.
- $0.9 < \phi < 1$ 이고 자료의 개수가 1,000 이하처럼 적은 경우에는 HQC의 분류 정확도가 BIC보다 높다. 이는 Figure 3.3의 1행 그림에서 보듯이 HQC보다 BIC를 사용했을 때  $\hat{\gamma} = 0$ 인 경우, 즉 DS로 판단하는 비율이 높기 때문이며, 선형회귀분석에서 BIC가 HQC보다 더 적은 변수를 선택하는 것과 비슷한 현상이라고 할 수 있다.
- 자료의 수가 증가할수록 HQC와 BIC가 LS 또는 DS에 속하는 시계열을 TS로 잘못 분류하는 비율이 0%로 감소하는 반면에, AIC는 그렇지 않다.

Figure 3.2와 Figure 3.3의 2-4행은  $b > 0$ 인 경우에 대한 분류 결과이며, 1,000번 중 TS로 분류한 비율(단위, %)과 DS로 분류한 비율(단위, %)을 각각  $\phi$ 와  $n$ 에 대하여 그린 선형 그래프이다. 모의실험 결과  $\phi < 0.5$ 인 경우 TS로 옳게 분류한 비율이  $\phi = 0.5$ 인 경우와 거의 차이가 없고 100%에 가깝게 나타났다. 그러므로 Figure 3.2에서는  $-1 < \phi \leq 1$ 의 전체 구간 대신  $0.5 \leq \phi \leq 1$ 의 구간에 대해서만 그림을 그렸다.

Figure 3.2와 Figure 3.3으로부터  $b > 0$ 인 경우에도  $b = 0$ 인 경우와 유사하게 자료의 개수  $n$ 이 증가할수록 adaptive lasso 추정량을 이용한 분류의 정확도가 높아지며,  $\phi$ 의 값이 1에 가까울수록 TS로 옳게 분류하는



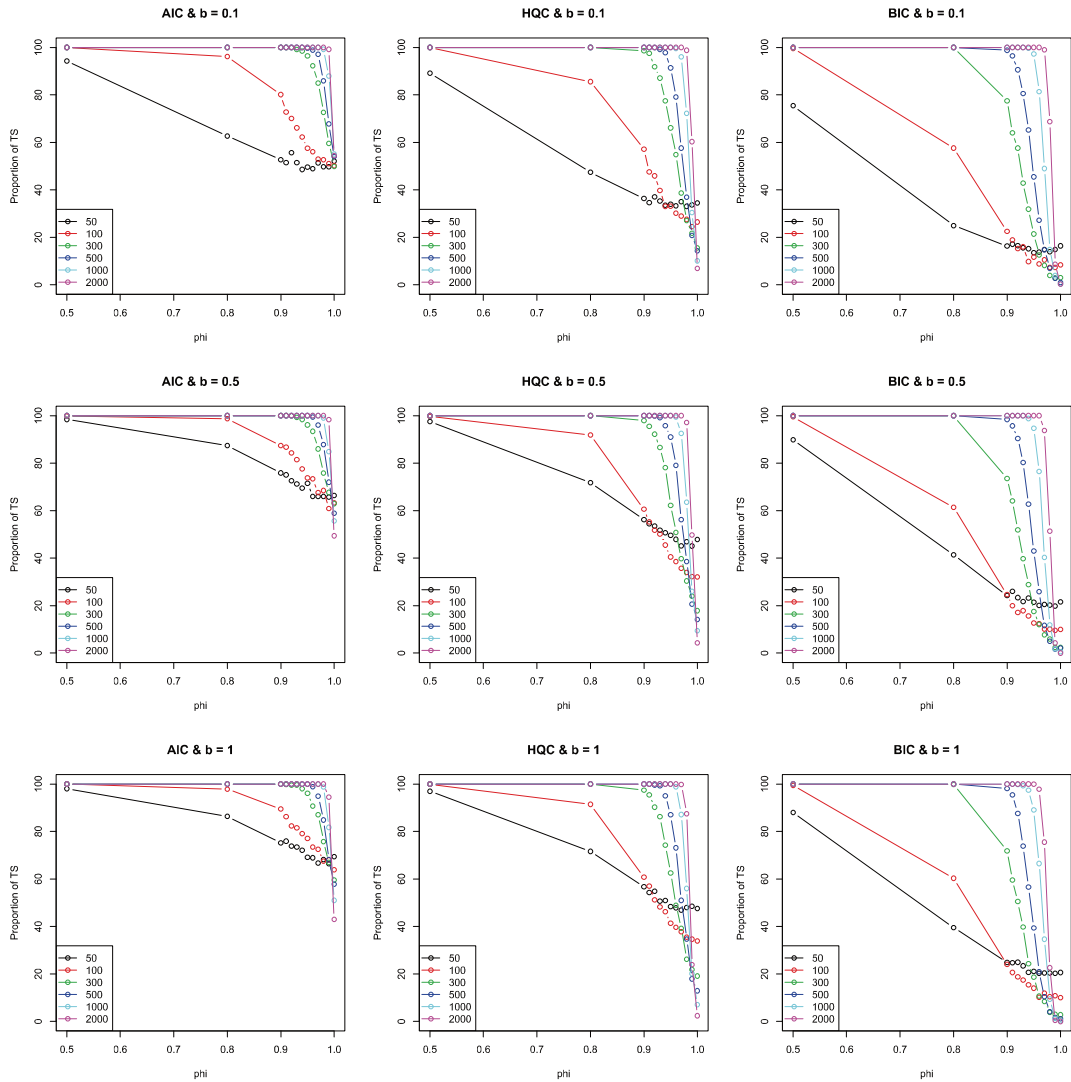


Figure 3.2. Proportion of TS when  $a = 0$  and  $b \neq 0$ .

비율이 낮아지고 DS로 잘못 분류하는 비율이 높아짐을 확인할 수 있었다. 그리고 추가적으로 다음과 같은 특징도 파악할 수 있었다.

- $\phi \leq 0.9$ 이고  $n \geq 300$ 인 경우, 정보함수의 선택에 관계없이 TS로 올바르게 분류하는 비율이 거의 100%에 가깝다. 그러나  $0.9 \leq \phi < 1$ 인 경우에는  $n$ 의 값이 같을 때 AIC의 정확도가 가장 높게 나타나고, BIC의 정확도가 가장 낮게 나타난다.
- $\phi$ 의 값이 크지 않은 경우, 예를 들면  $\phi \leq 0.95$ 인 경우에는 TS로 분류한 비율이  $b$ 의 값에 따라 크게 차이나지 않는다.

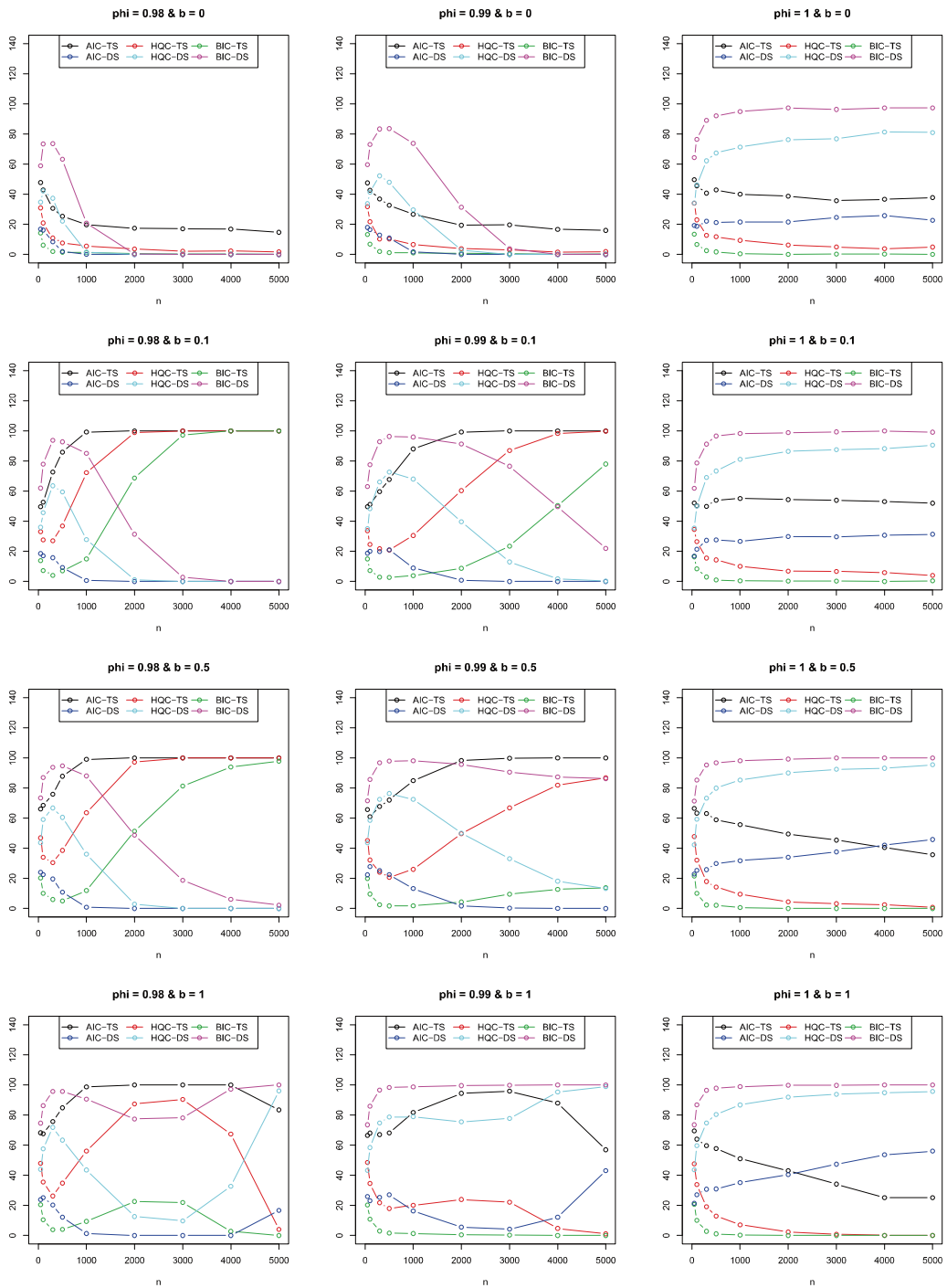


Figure 3.3. Proportions of TS and DS when  $\phi \geq 0.98$ .

- 그러나  $\phi$ 의 값이 극단적인 경우,  $\phi = 0.98$  또는  $\phi = 0.99$ 인 경우에는 Figure 3.3에서 보듯이  $b$ 의 값이 클수록 분류의 정확도가 떨어지며, TS와 DS로 분류하는 비율의 양상이 정보함수에 따라 다르다.  $b$ 가 크지 않을 때, AIC와 HQC는 자료의 수가 커질수록 TS로 옳게 분류하는 비율이 비교적 빠르게 증가한다. 그러나 BIC는 그 비율이 매우 천천히 증가하며,  $b = 0.5$ 일 때에는  $n = 5000$ 인 경우에도 옳게 분류하는 비율이 20%를 넘지 못 한다.
- $0.98 \leq \phi < 1$ 이고  $b = 1$ 로 기울기 큰 경우에는 자료의 수가 증가해도 분류의 정확도가 높아지지 않음을 확인할 수 있다. 오히려 자료의 수가 큰 경우에 TS로 옳게 분류하는 비율이 떨어지고, DS로 잘못 분류하는 비율이 높아진다.
- $\phi = 1$ 인 경우에는 BIC의 정확도가 가장 높으며, HQC의 정확도는 BIC보다는 낮지만 꽤 높은 값을 가진다. 이에 비해 AIC는 자료의 수가 큰 경우에도 정확도가 높지 않다.

위 실험 결과로부터 adaptive lasso 추정량을 이용하여 시계열의 정상성을 분류하는 것이 가능함을 알 수 있다. 그러나 조절모수를 선택하는 방법에 따라 분류의 정확도와 분류의 양상이 달라지므로 이에 대한 심도 깊은 추가 연구가 필요하다.

### 3.2. 비교 실험

다음으로 2장에서 설명했던 분류 방법들, 즉 ADF 단위근 검정을 기초로 시계열을 분류하는 방법과 adaptive lasso 추정량을 이용하여 분류하는 방법에 대한 비교 실험을 수행하고자 한다. 이를 위해

$$y_t = a + bt + z_t, \quad a = 0, 1, b = 0, 0.1, 0.5 \quad (3.3)$$

이고,  $z_t$ 가 다음의 다양한 모형을 따르는 경우를 고려하였다:

- $WN$  :  $\Delta^d z_t = \epsilon_t, \quad d = 0, 1, \epsilon_t \stackrel{iid}{\sim} N(0, 1)$ ;
- $AR(1)$  :  $\Delta^d z_t = 0.5 \Delta^d z_{t-1} + \epsilon_t, \quad d = 0, 1, \epsilon_t \stackrel{iid}{\sim} N(0, 1)$ ;
- $AR(2)$  :  $\Delta^d z_t = 0.5 \Delta^d z_{t-1} + 0.3 \Delta^d z_{t-2} + \epsilon_t, \quad d = 0, 1, \epsilon_t \stackrel{iid}{\sim} N(0, 1)$ ;
- $MA(1)$  :  $\Delta^d z_t = \epsilon_t + 0.5\epsilon_{t-1}, \quad d = 0, 1, \epsilon_t \stackrel{iid}{\sim} N(0, 1)$ ;
- $ARMA(1, 1)$  :  $\Delta^d z_t = 0.3 \Delta^d z_{t-1} + \epsilon_t + 0.7\epsilon_{t-1}, \quad d = 0, 1, \epsilon_t \stackrel{iid}{\sim} N(0, 1)$ .

여기서  $\Delta^d z_t$ 는  $z_t$ 를  $d$ 번 차분한 것을 나타내며,  $\Delta^0 z_t = z_t$ 이고  $\Delta^1 z_t = z_t - z_{t-1}$ 이다.  $d = 0$ 일 때  $\{z_t\}$ 는 정상성을 만족하고,  $b$ 의 값에 따라  $\{y_t\}$ 는 LS와 TS로 나누어진다. 그리고  $d = 1$ 일 때,  $\{z_t\}$ 는 비정상시계열이며  $\{y_t\}$ 는 DS가 된다.

각각의 모형에서 자료를  $n$ 개 생성하고, 2.2절과 2.3절에서 설명한 방법대로 시계열을 분류하였다. 일별이나 월별 자료의 경우에는 자료의 수가 많지 않은 경우가 많으므로, 비교 실험에서는 자료의 수로  $n = 50, 100, 300, 500, 1000$ 의 경우만을 고려하였다. 모의실험은 각각의 경우에 1,000번씩 반복하였으며, 1,000번의 분류 결과 중 시계열을 올바르게 참인 경우로 분류한 개수를 Table 3.1과 3.2에 정리하였다. Table 3.1은  $d = 0$ 일 때의 분류 정확도, Table 3.2는  $d = 1$ 일 때의 정확도를 나타낸다. 그리고 각 표에서 ENS는 Enders의 전략으로 구한 분류 결과이고, EKS는 Elder-Kennedy의 전략대로 분류를 수행한 결과를 나타낸다. AIC, HQC, BIC의 세 가지는 adaptive lasso 추정량을 기초로 시계열을 분류한 경우로, adaptive lasso 추



Table 3.2. Discrimination accuracy when  $d = 1$

$\Delta z_t$	$n$	$y_t = 0+0t+zt$					$y_t = 0+0.1t+zt$					$y_t = 0+0.5t+zt$				
		ENS	EKS	AIC	HQC	BIC	ENS	EKS	AIC	HQC	BIC	ENS	EKS	AIC	HQC	BIC
WN	50	774	882	208	366	638	791	888	192	353	637	787	872	222	425	719
	100	840	918	193	479	751	829	897	202	481	774	834	911	238	575	868
	300	863	946	220	625	898	855	921	239	674	920	885	940	273	730	950
	500	875	939	206	673	921	869	940	278	733	956	899	950	283	786	978
	1000	861	933	220	717	942	866	939	303	810	979	900	949	295	851	982
AR(1)	50	746	863	150	229	393	723	844	158	246	408	765	881	156	235	437
	100	816	904	144	310	557	819	918	136	304	549	844	915	175	346	551
	300	864	945	210	584	848	873	948	244	602	840	882	937	249	575	776
	500	860	946	216	653	907	869	947	234	659	895	868	931	252	627	836
	1000	881	957	245	738	954	894	942	308	790	965	904	953	270	727	909
AR(2)	50	718	860	188	258	389	717	854	205	256	373	715	864	165	228	375
	100	812	914	157	291	507	802	904	158	285	493	821	898	172	309	476
	300	853	945	183	535	799	853	950	172	536	790	860	939	245	550	782
	500	848	933	224	622	870	869	942	228	644	872	886	943	253	652	861
	1000	869	940	262	724	928	871	942	273	744	947	891	953	300	788	951
MA(1)	50	743	843	122	200	415	738	845	125	195	430	752	851	125	223	467
	100	801	895	119	255	491	825	903	113	268	517	822	896	150	292	521
	300	842	932	128	381	717	878	952	144	402	698	874	938	135	370	594
	500	867	937	156	486	766	853	924	144	424	724	885	952	136	374	576
	1000	862	941	141	518	826	886	939	167	503	802	876	928	150	407	587
ARMA(1,1)	50	739	865	127	178	314	750	844	123	186	305	727	833	102	155	291
	100	779	893	117	262	507	829	909	112	269	522	832	919	132	285	528
	300	854	936	102	338	683	850	926	111	359	690	859	934	117	371	625
	500	848	935	117	420	752	871	934	121	430	730	895	937	124	374	619
	1000	890	966	127	453	801	874	933	138	485	765	887	945	131	415	627

$\Delta z_t$	$n$	$y_t = 1 + 0t + z_t$					$y_t = 1 + 0.1t + z_t$					$y_t = 1 + 0.5t + z_t$				
		ENS	EKS	AIC	HQC	BIC	ENS	EKS	AIC	HQC	BIC	ENS	EKS	AIC	HQC	BIC
WN	50	774	883	207	367	621	777	865	192	345	624	783	877	214	414	716
	100	828	911	186	456	768	845	928	198	517	796	833	921	243	570	841
	300	860	941	212	620	898	873	948	246	646	931	887	942	279	746	960
	500	875	952	245	688	927	886	949	265	749	954	880	938	296	772	969
	1000	863	946	247	747	948	882	944	298	820	980	903	942	306	848	987
AR(1)	50	744	860	166	230	411	773	868	138	236	433	750	849	130	225	422
	100	825	913	151	341	561	821	915	166	363	578	801	891	166	334	528
	300	862	941	225	604	845	854	927	220	571	849	868	950	257	588	808
	500	869	937	246	668	903	907	956	265	704	920	864	934	260	620	832
	1000	860	943	249	744	945	894	951	311	794	959	896	951	297	724	901
AR(2)	50	702	863	163	223	393	683	846	188	246	382	687	845	171	225	369
	100	799	914	163	320	521	801	907	175	327	517	793	912	148	280	476
	300	840	935	208	515	785	846	935	192	537	777	887	943	200	529	772
	500	859	942	219	633	874	877	958	244	641	890	878	953	254	667	857
	1000	864	950	261	750	942	865	937	278	749	942	875	950	312	768	922
MA(1)	50	744	841	147	224	425	745	866	145	220	438	764	857	130	238	476
	100	837	918	130	276	526	811	907	138	290	535	825	911	135	272	532
	300	858	938	142	405	700	858	929	135	388	690	888	951	137	370	595
	500	878	948	161	465	770	874	938	157	462	744	876	933	152	395	603
	1000	860	949	142	529	821	882	943	150	500	796	895	947	178	436	618
ARMA(1,1)	50	747	861	125	174	307	742	839	120	164	315	740	840	128	180	298
	100	803	899	130	264	518	796	900	124	268	507	819	908	136	272	501
	300	865	939	122	358	671	857	929	116	354	672	876	945	118	333	598
	500	860	942	120	409	732	860	934	109	405	719	880	943	116	400	647
	1000	878	956	127	475	805	895	952	142	458	795	882	947	153	398	624

정량을 계산할 때 조절모수를 AIC, HQC, BIC로 선택하였음을 의미한다.

Table 3.1과 Table 3.2로부터 다음과 같은 특징을 파악할 수 있다.

- 자료의 수가 증가할수록 분류 정확도가 높아진다.
- $\alpha$ 의 값은 분류 결과에 큰 영향을 주지 못 한다.
- 대부분  $\{z_t\}$ 의 모형이 복잡해질수록 분류의 정확도가 떨어진다.
- LS가 참인 경우, AIC를 제외한 나머지 방법들은 모두  $n \geq 300$ 일 때 90% 이상의 높은 정확도를 가진다. 그러나  $n \leq 100$ 일 때는 EKS가 AIC를 포함한 다른 방법에 비해 정확도가 조금 낮게 나타난다. ENS와 HQC는 경우에 따라 조금씩 다르긴 하지만 대부분 정확도가 비슷하며, BIC가 전반적으로 가장 우세하다.
- TS가 참인 경우에는  $n \geq 300$ 일 때 모든 방법들이 거의 100%에 가깝게 시계열을 TS로 잘 분류한다. 그러나 자료의 수가 작은 경우, 즉  $n = 50$ 인 경우에는 ENS와 EKS가 adaptive lasso 추정량을 이용할 때에 비해 정확도가 많이 떨어진다.
- DS가 참인 경우를 살펴보면, EKS와 ENS가 adaptive lasso 추정량을 사용할 때보다 정확도가 높다.  $\{\Delta z_t\}$ 가 백색잡음(WN)인 경우에는 HQC와 BIC도 정확도가 높으며, 특히 BIC는 ENS나 EKS에 비해서도 정확도가 떨어지지 않는다. 그러나  $\{\Delta z_t\}$ 의 모형이 복잡한 경우, 특히 MA 항이 포함된 경우에는 adaptive lasso 추정량을 사용할 때 DS로 정확하게 분류하는 비율이 다소 떨어진다.

ADF 단위근 검정이 검정력이 낮다는 것은 잘 알려진 단점이다. 이로 인해 (2.4) 모형을 기초로 단위근 검정을 한번만 수행하는 Elder-Kennedy 전략(EKS)은 DS가 참일 때 분류 정확도가 높은 반면 LS나 TS가 참인 경우, 특히 자료의 수가 적은 경우 분류 정확도가 다소 떨어진다. Enders의 전략(ENS)도 ADF 단위근 검정을 사용하지만, (2.4) 모형부터 시작하여 (2.2) 모형까지 단위근 검정을 반복적으로 수행하면서 첫 검정에서 DS로 판단한 경우를 LS나 TS로 바꿔나간다. 그러므로 DS 분류 비율이 줄고 LS나 TS로 분류하는 비율이 증가, Table 3.1과 Table 3.2에 나타난 것처럼 Elder-Kennedy의 전략보다 DS일 때의 정확도가 조금 줄고 LS와 TS일 때의 정확도가 조금 늘어난다.

마지막으로 adaptive lasso 추정량, 특히 HQC나 BIC를 사용하여 조절모수를 선택한 경우에는 LS와 TS가 참일 때 분류 정확도가 다른 방법에 비해 높다. 그러나 DS가 참일 때는 분류 정확도가 떨어지며, 이는 adaptive lasso 추정량을 이용한 분류 절차의 단점이라 할 수 있다.

#### 4. 결론

본 논문에서는 Na (2019)의 연구 결과를 선형추세가 있는 경우로 확장, adaptive lasso 추정량을 추세-정상시계열과 차분-정상시계열을 분류하는 데 사용하였다. 모의실험 결과를 보면, 비록 차분-정상시계열의 경우는 단위근 검정 결과를 사용한 것에 비해 정확도가 다소 떨어지지만, 추세-정상시계열이 참인 경우에는 정확도가 단위근 검정보다 높은 것을 알 수 있다. 그리고 대부분의 경우 자료의 수가 증가함에 따라 분류 정확도가 증가하며,  $\phi$ 의 값이 0.95처럼 1에 비교적 가까운 경우에도 정확도가 매우 높다. 따라서 단위근 검정 이외에 adaptive lasso 추정량도 추세-정상성과 차분-정상성을 판단하는데 사용 가능하다. 다만 조절모수를 선택하는 방법에 따라 분류 정확도에 차이가 발생하므로, 이에 대한 심도 깊은 추가 연구가 필요하다.

## References

- Brockwell, P. J. and Davis, R. A. (2006). *Time Series: Theory and Methods* (2nd ed), Springer.
- Chen, K. and Chan, K. S. (2011). Subset ARMA selection via the adaptive lasso, *Statistics and Its Interface*, **4**, 197–205.
- Charles, A. and Darné, O. (2012). Trends and random walks in macroeconomic time series: A reappraisal, *Journal of Macroeconomics*, **34**, 167–180.
- DeJong, D. N. and Whiteman, C. H. (1991). Reconsidering ‘Trends and random walks in macroeconomic time series’, *Journal of Monetary Economics*, **28**, 221–254.
- 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.
- Dickey, D. and Fuller, W. (1981). Likelihood ratio statistics for autoregressive time series with a unit root, *Econometrica*, **49**, 1057–1072.
- Elder, J. and Kennedy, P. E. (2001). Testing for unit roots: What should students be taught?, *Journal of Economic Education*, **32**, 137–146.
- 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* (3rd ed), John Wiley & Sons.
- Friedman, J., Hastie, T., and Tibshirani, R. (2010). Regularization paths for generalized linear models via coordinate descent, *Journal of Statistical Software*, **33**, 1–22.
- Hacker, S. and Hatemi-J, A. (2010). The properties of procedures dealing with uncertainty about intercept and deterministic trend in unit Root Testing, *Working Paper Series in Economics and Institutions of Innovation 214*, Royal Institute of Technology, CESIS - Centre of Excellence for Science and Innovation Studies.
- 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.
- Kwon, S., Lee, S., and Na, O. (2017). Tuning parameter selection for the adaptive LASSO in the autoregressive model, *Journal of the Korean Statistical Society*, **46**, 285–297.
- Li, H. (1995). A reexamination of the Nelson-Plosser data set using recursive and sequential tests, *Empirical Economics*, **20**, 501–518.
- Na, O. (2019). Model selection for unstable AR process via the adaptive LASSO, *The Korean Journal of Applied Statistics*, **32**, 909–922.
- Nelson, C. R. and Plosser, C. I. (1982). Trends and random walks in macroeconomic time series, *Journal of Monetary Economics*, **10**, 139–162.
- Perron, P. (1988). Trends and random walks in macroeconomic time series: Further evidence from a new approach, *Journal of Economic Dynamics and Control*, **12**, 297–332.
- 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.
- Schwert, G. W. (1989). Tests for unit roots: A Monte Carlo investigation, *Journal of Business and Economic Statistics*, **7**, 147–160.
- Zou, H. (2006). The adaptive lasso and its oracle properties, *Journal of the American Statistical Association*, **101**, 1418–1429.

# Adaptive lasso를 이용하여 추세-정상시계열과 차분-정상시계열을 판별하는 방법에 대한 연구

나옥경 <sup>a,1</sup>

<sup>a</sup>경기대학교 응용통계학과

(2020년 10월 14일 접수, 2020년 10월 17일 수정, 2020년 10월 17일 채택)

---

## 요약

본 논문에서는 추세-정상시계열과 차분-정상시계열을 판별하는 방법에 대해 연구한다. 두 시계열 모형은 시계열적 특징, 충격의 지속성 여부, 시계열을 정상화시키는 방법 등이 모두 다르므로, 어떤 모형을 선택하나에 따라 분석 방법이나 해석에 차이가 발생한다. 따라서 시계열 자료를 분석할 때 추세-정상성과 차분-정상성을 판별하는 것은 매우 중요한 일이다. 두 시계열을 구분하는 중요한 기준은 단위근의 존재 여부이므로, 단위근 검정 결과를 활용할 수 있다. 최근 연구 결과들을 살펴보면, 다양한 시계열 모형을 적합시킬 때 뿐만 아니라 비정상 자기회귀모형의 차분 차수를 결정할 때도 adaptive lasso와 같은 벌점화 추정방법을 도입, 사용하고 있다. 본 논문에서도 adaptive lasso를 이용하여 추세-정상시계열과 차분-정상시계열을 판별하는 방법을 제안, 연구를 진행하였다. 단위근 검정을 이용한 분류 방법과 adaptive lasso 추정량을 기초로 한 분류 방법에 대한 비교 모의 실험을 수행하였고, 그 결과 추세-정상시계열이 참인 경우는 adaptive lasso 방법의 분류 정확도가 단위근 검정 방법보다 좀 더 우수하며, 차분-정상시계열의 경우에는 반대로 정확도가 떨어지는 것을 확인할 수 있었다.

**주요용어:** 추세-정상시계열, 차분-정상시계열, 단위근 검정, adaptive lasso 추정

---

---

<sup>1</sup>(16227) 경기도 수원시 영통구 광교산로 154-42, 경기대학교 응용통계학과. E-mail: okna@kyonggi.ac.kr