2022.03.28.
김은영

제6장 비정상시계열

6.1 시계열의 비정상성

6.2 ARIMA모형

■ **차분**: 시계열에 **추세**가 있는 경우, '**차분**'을 통해 정상적 시계열로 변환

- 1차 차분: $\triangle Z_t = Z_t - Z_{t-1} = Z_t - BZ_t = (1-B)Z_t$
- 2차 차분: $\triangle^2 Z_t = \triangle(\triangle Z_t) = \triangle(Z_t - Z_{t-1}) = \triangle Z_t - \triangle Z_{t-1}$

$$= (Z_t - Z_{t-1}) - (Z_{t-1} - Z_{t-2}) = Z_t - 2Z_{t-1} + Z_{t-2}$$

$$= (1-2B+B^2)Z_t = (1-B)^2 Z_t$$

- d차 차분: $\triangle^d Z_t = (1-B)^d Z_t, \quad d = 1,\ 2,\ \cdots$

[정의 6.1] d차 누적시계열 (Integrated Process of Order d)

> d차 차분 후 시계열이 처음으로 정상적이 될 때, 원 시계열을 d차 누적시계열이라 하며, $I(d)$로 표현한다.

[정의 6.2] ARIMA모형

> d차 차분 후 시계열이 정상적 ARMA(p,q)모형을 따를 때, 원 시계열이 ARIMA(p,d,q)모형을 따른다고 한다.

- ARIMA(p,d,q)모형: $\Phi_p(B)(1-B)^d Z_t = \Theta_q(B)a_t$

[예] ARIMA(1,1,1)모형

$$(1-\phi_1 B)(1-B)Z_t = (1-\theta_1 B)a_t, \quad -1 < \phi_1 < 1; -1 < \theta_1 < 1 \text{ 또는}$$

$$Z_t = (1+\phi_1)Z_{t-1} - \phi_1 Z_{t-2} + a_t - \theta_1 a_{t-1}$$

6.3 IMA모형 [~지수가중이동평균(EWMA)]

[정의 6.3] IMA모형

> ARIMA(p,d,q)모형에서 $p=0$인 경우를 IMA(d,q)모형이라 한다. 즉,
> $$(1-B)^d Z_t = \Theta_q(B)a_t$$

· IMA(1,1)모형: $(1-B)Z_t = (1-\theta B)a_t$, $\quad -1 < \theta < 1$

→ AR형태로 표현하면 $Z_t = \sum_{j=1}^{\infty} \pi_j Z_{t-j} + a_t$

여기서, $\pi_j = (1-\theta)\theta^{j-1}$, $\quad j = 1,\ 2,\ \cdots$

$\lambda = 1-\theta,\ \ 0 < \lambda < 1 \ \rightarrow \ \pi_j = \lambda(1-\lambda)^{j-1}$ [=EWMA에서의 가중치]

⇒ 시점 T에서의 시계열 수준: $L_T = \sum_{j=1}^{\infty} \pi_j Z_{T+1-j} = \lambda \sum_{j=1}^{\infty} (1-\lambda)^{j-1} Z_{T+1-j}$

시점 $T+1$에서의 시계열 수준: $\underline{L_{T+1} = \lambda Z_{T+1} + (1-\lambda)L_T}$

※ 시점 $T+1$의 시계열 값이 관측된 후의 EWMA의 갱신식(updating formula)와 동일

6.4 계절성 ARIMA모형

▪ 추세가 없는 시계열에서 주기(period) s의 계절성(seasonality)을 갖는다면

$$E[Z_t] = E[Z_{t+s}]$$

▪ 추세가 없는 계절성 시계열을 정상적인 것으로 변환하기 위해서는 계절성 차분(seasonal differencing)이 필요하다.

▪ 주기(period) s의 계절성(seasonality) 차분

$$\nabla_s Z_t = (1-B^s)Z_t = Z_t - Z_{t-s}$$

[예_1] 주기 12를 가지며, 추세가 없는 월별 시계열이 있다.

이때 1월의 데이터들이 MA(1)모형을 따른다면

$$Z_t = (1-\Theta B^{12})a_t$$

여기서, $a_t, a_{t-12}, a_{t-24}, \cdots$: 월이 고정되면 해당 오차항들은 서로 상관관계가 없다.

→ 그러나 인접한 월간에는 상관관계가 있을 수 있으므로 오차항들에 대한 새로운 모형이 필요하다.

시계열이 MA(1)모형을 따른다면: $a_t = (1-\theta B)a_t$

해당 시계열모형: $\underline{Z_t = (1-\theta B)(1-\Theta B^{12})a_t}$

↳ 비계절성 MA(1)모형과 계절성 MA(1)모형이 결합된 형태

[예_2] 주기 12를 가지며, 추세가 없는 월별 시계열이 있다.

이때 1월의 데이터들이 IMA(1,1)모형을 따른다면

$$(1-B^{12})Z_t = (1-\Theta B^{12})a_t$$

여기서, $a_t, a_{t-12}, a_{t-24}, \cdots$: 월이 고정되면 해당 오차항들은 서로 상관관계가 없다.

→ 그러나 인접한 월간에는 상관관계가 있을 수 있으므로 오차항들에 대한 새로운 모형이 필요하다.

시계열이 IMA(1,1)모형을 따른다면: $(1-B)a_t = (1-\theta B)a_t$

해당 시계열모형: $\underline{(1-B)(1-B^{12})Z_t = (1-\theta B)(1-\Theta B^{12})a_t}$

↳ 비계절성 IMA(1,1)모형과 계절성 IMA(1,1)모형이 결합된 형태

⇒ 계절성 ARIMA$(0,1,1)\times(0,1,1)_{12}$ 모형

[일반화]

- 주기 s를 갖는 시계열에 대하여 특정 계절별 시계열이 ARIMA(P,D,Q)를 따르고

$$\Phi_P(B^s)(1-B^s)^D Z_t = \Theta_Q(B^s)a_t$$

- 오차항들이 ARIMA(p,d,q)를 따르면

$$\phi_p(B)(1-B)^d a_t = \theta_q(B)a_t$$

- 원 시계열: $\underline{\phi_p(B)\Phi_P(B^s)(1-B)^d(1-B^s)^D Z_t = \theta_q(B)\Theta_Q(B^s)a_t}$

→ 비계절성 ARIMA(p,d,q)모형과 계절성 ARIMA$(P,D,Q)_s$모형이 결합된 형태

⇒ 계절성 ARIMA$(p,d,q)\times(P,D,Q)_s$ 모형

■ 계절성 ARIMA모형 식별을 위한 절차

1. 시계열도를 그려보고 추세 및 계절성의 존재 여부를 판단한다.
2. 아래 사항을 고려하여 적절히 차분한다.

(1) 추세는 없고 계절성이 있는 경우

: 해당 주기에 대한 계절성 차분을 실시한다.

(2) 추세는 있고 뚜렷한 계절성이 없는 경우

① 선형추세가 있는 경우 1차 차분을 한다.

② 곡선추세가 있는 경우 차분 전에 변환을 시도한다.

(3) 추세와 계절성이 있는 경우

: 우선 계절성 차분을 실시하고 추세를 다시 조사한다.

추세가 여전히 남아있는 경우 1차 차분을 추가로 실시한다.

3. 차분 시계열에 대한 ACF 및 PACF를 바탕으로 p, q, P, Q를 결정한다.
(1) 비계절성 계수인 p, q는 ARMA모형의 경우와 동일한 요령으로 결정한다.
(2) 계절성 계수인 P, Q는 주기의 배수에서 나타나는 ACF 및 PACF의 패턴을 보고 결정한다.

4. 모형을 추정한다.
5. 잔차검정을 실시한다.

6.5 ARIMA모형의 예측

■ 시점 n에서 k단계 이후 예측값과 예측오차분산

$$f_{n,k} = E[Z_{n+k} | Z_n,\ Z_{n-1},\ \cdots],\ k = 1,\ 2,\ \cdots$$

$$v_{n,k} = Var[Z_{n+k} | Z_n,\ Z_{n-1},\ \cdots],\ k = 1,\ 2,\ \cdots$$

6.6 분산안정화

■ 시계열의 분산이 시간에 따라 변하는 비정상적 시계열의 경우 종종 로그함수 등과 같은 변환을 통해 시계열의 분산이 안정화될 수 있어 모형화에 도움이 된다.

■ 시계열 $\{Z_t,\ t = 1,2,\cdots\}$의 분산이 상수가 아니라 다음과 같이 기댓값의 함수 f에 비례한다고 하자.

$$Var(Z_t) = cf(\mu_t)$$

여기서, c: 비례상수, $\mu_t = E(Z_t)$

■ 시계열 Z_t에 어떤 변환을 취해야 변환된 시계열의 분산이 일정한 상수가 되는지 알아보자.

· g라는 함수를 이용하여 시계열을 변환시키고 μ_t를 중심으로 1차 테일러 전개를 하면

$$g(Z_t) = g(\mu_t) + g'(\mu_t)(Z_t - \mu_t)$$

· 위의 변환 시계열의 분산은

$$Var[g(Z_t)] = [g'(\mu_t)]^2 Var(Z_t) = c[g'(\mu_t)]^2 f(\mu_t)$$

· 분산이 상수 c가 되기 위해서는 변환 함수 g가 다음 성질을 만족해야 한다.

$$g'(\mu_t) = \frac{1}{\sqrt{f(\mu_t)}} \quad \Rightarrow \quad \text{변환함수: } g(\mu_t) = \int \frac{1}{\sqrt{f(\mu_t)}} d\mu_t$$

- Box and Cox(1964): 여러 형태를 갖는 파워 변환(power transformation)을 제안

$$g(Z_t)=\begin{cases}\dfrac{Z_t^{\lambda}-1}{\lambda}, & \lambda \neq 0 \\ \ln(Z_t)\ , & \lambda = 0\end{cases}$$

→ λ값에 따라 다양한 변환함수를 갖게 된다.

표 6.1 Box-Cox 변환 함수

λ 값	$g(Z_t)$
-2.0	$g(Z_t)=1/Z_t^2$
-1.0	$g(Z_t)=1/Z_t$
-0.5	$g(Z_t)=1/\sqrt{Z_t}$
0.0	$g(Z_t)=\ln(Z_t)$
0.5	$g(Z_t)=\sqrt{Z_t}$
2.0	$g(Z_t)=Z_t^2$

6.8 단위근검정

■ 시계열의 정상성 여부를 판단하는 통계적 검정법

- DF(Dickey-Fuller) 검정: Dickey and Fuller(1979)
- ADF(Augmented Dickey-Fuller) 검정: Said and Dickey(1984)

■ ADF 검정

- 가정: 모든 정상적 시계열은 고차원의 AR모형으로 근사될 수 있다.

  → 시계열이 다음과 같은 AR(p)모형을 따른다.

$$\Phi_p(B)Z_t = a_t$$

- 다항식 $\Phi_p(B)$에 단위근이 포함된다고 가정하면

$$\Phi_p(B)Z_t = (1-B)\Psi_{p-1}(B)$$

여기서, $\Psi_{p-1}(B)$: $(p-1)$차 다항식

$$\rightarrow\ (1-B)(1-\psi_1 B- \cdots - \psi_{p-1}B^{p-1})Z_t = a_t \quad \text{또는}$$

$$(1-B)Z_t = (1-B)(\psi_1 B+ \cdots + \psi_{p-1}B^{p-1})Z_t + a_t \ \text{또는}$$

$$Z_t - Z_{t-1} = \sum_{j=1}^{p-1}\psi_j(Z_{t-j}-Z_{t-j-1}) + a_t$$

■ 다음 모형을 고려하여 가설 $H_0:\ \phi=1$ 여부를 검정하는 것이 단위근검정과 동일

$$Z_t=\phi_1 Z_{t-1}+\sum_{j=1}^{p-1}\psi_j \triangle Z_{t-j}+a_t \text{ ------ ①}$$

$$\text{여기서, } \triangle Z_{t-j}=Z_{t-j}-Z_{t-j-1}$$

· 참고로 다음 관계가 성립한다.

$$\psi_j=\phi_1+\ \cdots\ +\phi_j-1,\quad j=1,2,\cdots,p-1$$

■ 시계열 관측값 $\{Z_1, Z_2, \cdots, Z_n\}$이 주어질 때 식 ①의 모형 파라미터를 추정하기 위해 최소제곱법(OLS)을 사용한다.

$$y=X\beta+\epsilon$$

여기서,

$$y=\begin{pmatrix} Z_{p+1} \\ Z_{p+2} \\ \cdots \\ Z_n \end{pmatrix},\ X=\begin{bmatrix} Z_p & \Delta Z_p & \Delta Z_{p-1} \cdots & \Delta Z_2 \\ Z_{p+1} & \Delta Z_{p+1} & \Delta Z_p \ \cdots & \Delta Z_3 \\ \vdots & \vdots & \vdots & \vdots \\ Z_{n-1} & \Delta Z_{n-1} & \Delta Z_{n-2} \cdots & \Delta Z_{n-p+1} \end{bmatrix},\ \beta=\begin{bmatrix} \phi \\ \varphi_1 \\ \vdots \\ \varphi_{p-1} \end{bmatrix},\ \epsilon=\begin{pmatrix} a_{p+1} \\ a_{p+2} \\ \cdots \\ a_n \end{pmatrix}$$

· 회귀계수 추정치

$$\hat{\beta}=\left(X^T X\right)^{-1} X^T y$$

■ 단위근 검정을 위한 검정통계량

$$T=\frac{\hat{\phi}-1}{se(\hat{\phi})}$$

■ 단위근검정 결과에 대한 해석

· 귀무가설($H_0:\ \phi=1$) 기각 ○ → 단위근이 없다 → 정상적 시계열이다.

· 귀무가설($H_0:\ \phi=1$) 기각 × → 단위근이 있다 → 비정상적 시계열이다.

⇒ 차분을 취해 정상적 시계열로 변환 후 추가분석 시행