# 시계열분석

2021-15115 권이태

September 18, 2024

# Contents

# Chapter 1

# 시계열자료

시간에 따라 관측된 자료를 **시계열자료**라 한다. 시계열은 연속적으로 생성되는 **연속시계열**과 이산적 시점에서 생성되는 **이산시계열**로 나뉘나, 대개 관측의 한계로 인해 이산시계열을 분석한다. 시계열자료는 일반적으로 시간 $t$를 이용하여

$$\{Z_t, t = 1, 2, \cdots, \}$$

와 같이 표현한다. 이 시계열자료로부터 시계열의 생성 특성을 파악하여 미래의 관측값을 예측하는 것이 주요 목표가 된다.

시계열자료 분석 시에는 먼저 탐색적 자료분석을 위하여 **시계열그림**을 그린다. 이는 시간 $t$를 가로축으로, 시계열의 관측값 $Z_t$를 세로축으로 하여 그린다. 시계열그림을 그리면 $Z_t$의 평균 수준이 일정한지, 그 분산은 일정한지, 어떠한 추세가 있는지 등을 간단히 확인해볼 수 있다는 점에서 유용하다.

시계열분석의 과정은 아래의 단계를 따른다.

1. 모형의 식별(identification)
2. 모형의 추정(model estimation)
3. 모형의 진단(model diagnostics)
4. 예측(forecasting)
5. 안정성 조사(stability checking)
6. 예측 갱신(forecast updating)

한편 모형의 적합과 예측 방법은 다양한 선택 기준을 가지고 있다. 이는 모형의 정확성, 소요되는 시간, 복잡도, 데이터에의 접근성, 예측 기간의 길이 등을 총체적으로 고려하여 결정한다.

**Definition 1.** *관측된 자료에서 시점 $t$에서의 $l$-시차 이후의 예측값을 $\hat{Z}_t(l)$처럼 쓸 수 있으며,* ***예측오차**는*

$$\hat{e}_t(l) = Z_{t+l} - \hat{Z}_t(l)$$

*으로 정의한다.*

**Definition 2.** *만약 시계열 $\{\epsilon_t\}$가 아래 세 조건을 만족하면,* ***백색잡음**이라 한다.*

- $\mathbb{E}[\epsilon_t] = 0$
- $\mathrm{Var}(\epsilon_t) = \sigma^2$
- $Cov(\epsilon_t, \epsilon_s) = 0 \text{ if} t \neq s$

# Chapter 2

# 추세분석

전통적으로 시계열자료의 분석 과정에서 많이 사용되어 온 방법은 $Z_t$를 시간의 함수로 표현하는 것이다. 특히 회귀분석을 적용할 수 있는 **선형다항추세모형**

$$Z_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \cdots + \beta_p t^p + \epsilon_t$$

를 많이 사용한다. 만약 시계열이 일정한 수준을 중심으로 위아래로 흔들리면 $p=0$인 **상수평균모형**을, 선형적인 증가추세를 가지면 $p=1$인 **선형추세모형**을 주로 적용한다. 만약 성장이 지수적이라면 로그를 씌운 뒤 분석하는 것을 고려해볼 수 있다. 이는

$$Z_t = e^{\beta_0 + \beta_1 t} \epsilon_t$$

와 같은 **비선형추세모형**으로 써질 수 있다.

먼저 상수평균모형

$$Z_t = \beta_0 + \epsilon_t$$

에서는 $\hat{\beta}_0 = \frac{1}{n}\sum_{t=1}^{n} Z_t = \bar{Z}$으로 추정하며, $n$시점에서 $l$시차 이후의 추정된 예측값은

$$\hat{Z}_n(l) = \hat{\beta}_0 = \bar{Z}$$

이며, 예측오차는

$$\hat{e}_n(l) = Z_{n+l} - \hat{Z}_n(l) = \beta_0 + \epsilon_{n+l} - \bar{Z}$$

이다. 예측값의 기대값은 $\beta_0$, 그 분산은 $\left(1+\frac{1}{n}\right)\sigma^2$으로 나타난다. $\sigma^2$의 불편추정량으로는

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{t=1}^{n}(Z_t - \bar{Z})^2$$

을 사용할 수 있다. 만약 $\epsilon_t$가 정규분포를 따른다고 가정하면, $Z_{n+l}$의 95퍼센트 예측구간은

$$\bar{Z} \pm t_{\alpha/2}(n-1) \times \hat{\sigma}\sqrt{1+\frac{1}{n}}$$

으로 주어진다. 여기에서 만약 $Z_{n+1}$을 새롭게 얻었다면, $n+1$ 시점에서 $l-1$시차 이후의 추정된 예측값은

$$\begin{aligned}
\hat{Z}_{n+1}(l-1) &= \frac{1}{n+1}(Z_1 + \cdots + Z_n + Z_{n+1}) \\
&= \frac{1}{n+1}(n\bar{Z} + Z_{n+1}) \\
&= \hat{Z}_n(l) + \frac{1}{n+1}(Z_{n+1} - \bar{Z})
\end{aligned}$$

$$= \hat{Z}_n(l) + \frac{1}{n+1}\hat{e}_n(1)$$

이 된다. 즉 예측값은 새로운 정보가 들어올 때마다 전 시기의 예측값과 예측오차를 이용하여 쉽게 갱신할 수 있다. 만약 실제 $\beta_0$에 비하여 $\hat{\beta}_0$이 작았다면, $\hat{e}_n(1)$이 양수일 가능성이 크고, 이러한 갱신법은 우리의 $\hat{\beta}_0$을 올바른 방향으로 지도하여 줄 수 있다.

만약 다항추세를 가지고 있는 경우, 선형회귀모형을 적합하여 $\beta_0, \beta_1, \cdots, \beta_p$를 추정한다. 다만 설명변수로 $t, t^2, t^3$ 등을 곧이곧대로 넣는 경우 다중공선성의 문제가 발생할 가능성이 크므로, 각각을 센터링하여 넣는 방법 또한 자주 이용된다. 각각의 신뢰구간과 $Z_{t+l}$의 예측은 일반적인 회귀분석에서와 동일하게 수행할 수 있다.

## 2.1 계절성분이 존재하는 경우의 추세분석

계절성이 존재하는 경우, 주기를 표현하기 위해 삼각함수를 이용하거나, 각 시기에 대한 가변수를 넣어 회귀분석을 수행하는 방법이 자주 이용된다. 삼각함수를 이용하는 경우에는 계절에 따른 연속성을 고려할 수 있다는 장점이 있지만, 특이한 형태의 계절성을 모델링하지 못한다는 단점 역시 있다. 일반적으로 주기가 $s$인 계절성이 있다고 생각되는 경우,

$$Z_t = (\beta_0 + \beta_1 t + \cdots + \beta_p t^p) + \gamma_1 \sin\left(\frac{2\pi t}{s}\right) + \gamma_2 \cos\left(\frac{2\pi t}{s}\right) + \epsilon_t$$

로 모형화할 수 있다. 계절성분이 여러 개라 생각된다면 삼각함수 항 두 개를 다른 주기 $s'$을 통해 적용함으로써 추가할 수 있을 것이다. 지시변수를 이용하는 경우에는

$$Z_t = (\beta_0 + \beta_1 t + \cdots + \beta_p t^p) + \sum_{i=1}^{s} \gamma_i I(t \equiv i \pmod{s}) + \epsilon_t$$

를 사용할 수 있다. 단 식별조건을 위하여 $i = 0$일 때는 기저수준으로 보고 추가하지 않는다. 이때의 추론은 회귀분석에서 이미 많이 해보았을 것이므로 생략한다.

혹은 계절성분만이 아니라 비선형추세 등 다양한 형태의 모형이 있을 수 있다. 이 경우에는 로그변환 등의 방법을 통하여 선형성을 만족하도록 모형을 바꾼 뒤 회귀분석을 적용하면 된다.

## 2.2 자기회귀오차모형

위와 같은 회귀모형을 적합해 분석하는 경우 $\epsilon_t$의 독립성이 중요하다. 이 경우에만 OLS 추정량이 BLUE가 되기 때문이다. 그러나 시계열자료 상에서는 오차항 역시 시간에 따라 일정한 상관성을 가지는 경우가 많다. 특히 오차항이 자기회귀과정을 따르는 **자기회귀오차모형**이 사용될 수 있다. 이는 회귀분석에서의 GLSAR과도 유사하다.

외생변수 $X_{t1}, X_{t2}, \cdots, X_{tp}$에 대하여

$$Z_t = \beta_0 + \beta_1 X_{t1} + \cdots + \beta_p X_{tp} + \epsilon_t$$

로 모형화하자. 이때 $\epsilon_t$가 백색잡음이 아니고, 특정한 차수 $k$의 자기회귀구조를 따른다고 가정한다. 즉

$$\epsilon_t = \phi_1 \epsilon_{t-1} + \cdots + \phi_k \epsilon_{t-k} + a_t$$

에 의해 순차적으로 $\epsilon_t$가 결정되고, $a_t$가 백색잡음이라는 것이다. 이를 적합하기 위해서는 먼저 $k$를 결정한다. 그 방법은 아래와 같다.

1. OLS를 통해 오차항 $\epsilon_t$의 추정치인 잔차 $e_t$를 구한다.

2. $j=1$에 대하여 DW 검정을 진행한다. 즉

$$D_j = \frac{\sum_{t=j+1}^{n}(e_t - e_{t-j})^2}{\sum_{t=1}^{n} e_t^2}$$

을 구한 뒤 이 값이 $d_L(j)$보다 작으면 양의 상관성이, $d_U(j)$보다 크면 음의 상관성이, 그 사이에 있으면 미결로 판단한다.

3. 만약 자기상관이 있다고 나왔다면, $j=2$에 대하여 DW 검정을 진행하여 결과를 본다.

4. 이를 $j$를 하나씩 늘려가면서 진행하고, 처음으로 자기상관성이 있다고 판단되지 않는 $j$를 $k$로 선택한다.

그 다음으로는 오차항에

$$\epsilon_t = \phi_1\epsilon_{t-1} + \cdots + \phi_k\epsilon_{t-k} + a_t$$

구조를 준 뒤 그에 맞게 최적의 추정량 $\beta_i$들을 찾는다. 다만 그 과정은 컴퓨터로 해야 하는 만큼 복잡하여 생략한다.

# Chapter 3

# 평활법

시계열자료는 자료들이 시간에 따라 장기간에 걸쳐 관측된다. 그러나 시간이 경과함에 따라 시계열이 생성되는 시스템 자체에 변화가 있을 수 있고, 하나의 모형을 통해 이를 파악하기 어렵다. 이러한 경우 최근의 자료에 더 많은 가중치를 두고 시계열을 평활화하여 예측에 사용하는 **지수평활법**을 사용할 수 있다.

## 3.1 단순지수평활법

시계열자료가 아래와 같은 모형을 따른다고 생각하자.

$$Z_t = \beta_0(t) + \epsilon_t$$

이때 앞선 경우와는 달리 $\beta_0(t)$는 시간에 따라 변화할 수 있는 모수이며, 국지적으로는 동일한 수준을 가지지만 전체적으로는 시간대별로 평균수준이 변할 수 있다고 가정한다. 이는 앞선 경우처럼 상수평균모형에 따라 추정할 경우 현 시점의 $\beta_0$가 아니라 전체적으로 평균의 $\beta_0$과 유사한 값을 얻게 되어 예측에 문제가 있다. 여기에서는 최근 값에 더 큰 가중치를 두게 함으로써 이를 해결할 수 있다.
$Z_1, \cdots, Z_n$까지의 자료를 바탕으로 만든 $\hat{\beta}_0$의 예측값으로써

$$\hat{Z}_n(1) = \hat{\beta}_{0,n} = \frac{1}{n}\sum_{t=1}^{n} Z_t = \bar{Z}$$

를 사용한다면, 예측오차는

$$\hat{e}_n(1) = Z_{n+1} - \hat{Z}_n(1) = Z_{n+1} - \hat{\beta}_{0,n}$$

이다. 이제 $0 < \omega < 1$에 대하여 $Z_{n+1}$이 관측되었을 때의 $\hat{\beta}_{0,n+1}$을 아래처럼 갱신할 수 있다.

$$\begin{aligned}\hat{\beta}_{0,n+1} &= \hat{\beta}_{0,n} + \omega\hat{e}_n(1) \\ &= \hat{\beta}_{0,n} + \omega(Z_{n+1} - \hat{\beta}_{0,n}) \\ &= (1-\omega)\hat{\beta}_{0,n} + \omega Z_{n+1}\end{aligned}$$

지수평활법에서는 $\hat{\beta}_{0,n}$을 $S_n^{(1)}$처럼 쓴다. 따라서

$$\begin{aligned}S_{n+1}^{(1)} &= \omega Z_{n+1} + (1-\omega)S_n^{(1)} \\ &= \omega\sum_{j=0}^{n}(1-\omega)^j Z_{n+1-j} + (1-\omega)^{n+1}S_0^{(1)}\end{aligned}$$

으로 전개될 수 있다. 여기에서 $\omega$는 **평활상수**, $S_0^{(1)}$은 **초기 평활값**, $S_n^{(1)}$은 **단순지수평활** 통계량이다. 그러면 $n \to \infty$임에 따라

$$S_n^{(1)} = \omega \sum_{j=0}^{n-1} (1-\omega)^j Z_{n-j}$$

으로, 지수적으로 감소하는 가중치가 있는 가중평균이 된다. 예측 역시도 $\hat{Z}_n(l) = S_n^{(1)}$으로 쉬우며, 그 갱신 역시도 간단하다. 초기 평활값으로는 표본평균, 최초 관측값 등 다양한 값을 사용할 수 있으며, $\omega$는

$$\mathrm{SSE}(\omega) = \sum_{t=2}^{n} \hat{e}_{t-1}^2(1)$$

을 최소로 하는 $\omega$를 사용하는 것이 바람직하다. 한편 이러한 지수평활법은 가중치를 $\omega_t = (1-\omega)^{n-t}$로 하는 가중최소제곱법과도 유사하다.

## 3.2 이중지수평활법

시계열이 선형추세에 따라 증가하는 경우에도 지수평활법의 적용이 가능하다.

$$Z_t = \beta_0(t) + \beta_1(t)t + \epsilon_t$$

에서, 앞에서 구한 $S_n^{(1)}$의 기대값은

$$\mathbb{E}[S_n^{(1)}] = \beta_0 + \beta_1(n+1) - \frac{1}{\omega}\beta_1$$

으로 $Z_{n+1}$의 예측값으로 사용할 시 $\beta_1/\omega$만큼의 편향이 생기게 된다. 이를 위해서는 $S_n^{(1)}$에 한 번 더 단순지수평활을 적용해 얻어지는 **이중지수평활** 통계량 $S_n^{(2)}$를 도입하여야 한다.

$$\begin{aligned} S_n^{(1)} &= \omega Z_n + (1-\omega)S_{n-1}^{(1)} \\ S_n^{(2)} &= \omega S_n^{(1)} + (1-\omega)S_{n-1}^{(2)} \end{aligned}$$

그 다음, 이로부터 시점 $n$에서의 절편 $\beta_0$과 기울기 $\beta_1$의 추정량을 각각

$$\begin{aligned} \hat{\beta}_{0,n} &= 2S_n^{(1)} - S_n^{(2)} - \hat{\beta}_{1,n}n \\ \hat{\beta}_{1,n} &= \frac{\omega}{1-\omega}(S_n^{(1)} - S_n^{(2)}) \end{aligned}$$

으로 하여 얻을 수 있으며,

$$\hat{Z}_n(l) = \left(2 + \frac{\omega}{1-\omega}l\right) S_n^{(1)} - \left(1 + \frac{\omega}{1-\omega}l\right) S_n^{(2)}$$

으로 주어진다. 초기 평활값이나 $\omega$의 선택 등은 생략한다. 이를 확장하면, $Z_t = \beta_0 + \beta_1 t + \beta_2 \frac{t^2}{2} + \epsilon_t$와 같은 상황에서도 삼중지수평활법을 수행할 수 있다.

## 3.3 계절지수평활법

시계열의 평균수준이 시간의 흐름에 따라 변화하지만, 계절에 따른 변동폭이 시간의 흐름에 관계없이 일정한 경우에는 **가법계절모형**으로 모형화할 수 있다. Winters의 가법계절모형에서 $Z_{n+l}$은

$$Z_{n+l} = T_{n+l} + S_{n+l} + I_{n+l}$$

으로 분해된다. $T_{n+l}$은 시간의 흐름에 따라 지속적으로 감소하거나 증가할 수 있는 **추세성분**, $S_{n+l}$은 계절에 따라 주기적으로 운동하는 **계절성분**, $I_{n+l}$은 오차항인 **불규칙성분**이다. 추세가 선형추세를 따르고, 계절성분의 주기가 $s$라고 하면 $n$에서 $Z_{n+l}$의 예측값은

$$\hat{Z}_n(l) = \hat{T}_n + \hat{\beta}_{1,n}l + \hat{S}_{n+l-([(l-1)/s]+1)s}$$

으로 $\hat{T}_n, \hat{\beta}_{1,n}$과 계절성분 $\hat{S}_{n+1-s}, \cdots, \hat{S}_n$만 이용하여 추후의 예측을 쉽게 수행해줄 수 있다. 이들은 각각 아래처럼 얻는다.

$$\begin{aligned}
\hat{T}_{n+1} &= \omega_1(Z_{n+1} - \hat{S}_{n+1-s}) + (1-\omega_1)(\hat{T}_n + \hat{\beta}_{1,n}) \\
\hat{\beta}_{1,n+1} &= \omega_2(\hat{T}_{n+1} - \hat{T}_n) + (1-\omega_2)\hat{\beta}_{1,n} \\
\hat{S}_{n+1} &= \omega_3(Z_{n+1} - \hat{T}_{n+1}) + (1-\omega_3)\hat{S}_{n+1-s}
\end{aligned}$$

즉 추세성분은 관측값에서 계절성분을 조정한 값과 이전 추세성분의 지수평활로, 기울기는 추세성분 간의 차이로 얻은 값과 이전 기울기의 지수평활로, 계절성분은 관측값에서 추세성분을 조정한 값과 이전 계절성분의 지수평활로 새롭게 업데이트된다. 초기 평활값은 회귀식 등을 이용하여 적절히 결정한다.

**승법계절지수평활법**에서는 승법적으로

$$Z_{n+l} = T_{n+l}S_{n+l} + I_{n+l}$$

으로 계절에 의한 변동이 수준에 따라 달라질 수 있다고 모형화한다. 그렇다면 예측값은

$$\hat{Z}_n(l) = (\hat{T}_n + \hat{\beta}_{1,n}l)\hat{S}_{n+l-([(l-1)/s]+1)s}$$

처럼 나타나고, 각 성분은 아래처럼 업데이트한다.

$$\begin{aligned}
\hat{T}_{n+1} &= \omega_1\left(\frac{Z_{n+1}}{\hat{S}_{n+1-s}}\right) + (1-\omega_1)(\hat{T}_n + \hat{\beta}_{1,n}) \\
\hat{\beta}_{1,n+1} &= \omega_2(\hat{T}_{n+1} - \hat{T}_n) + (1-\omega_2)\hat{\beta}_{1,n} \\
\hat{S}_{n+1} &= \omega_3(\frac{Z_{n+1}}{\hat{T}_{n+1}}) + (1-\omega_3)\hat{S}_{n+1-s}
\end{aligned}$$

마찬가지로 초기 평활값이나 가중치는 적절한 회귀식이나 가중평균 등으로 결정한다.

# Chapter 4

# 계절조정

평활법에서는 시계열을 구성하는 계절성분 등을 구분하지 않고 불규칙성분을 제거하여 미래의 값을 예측하는 것이 목적인 반면에, 분해법은 각 성분들을 구분한 뒤 그들로써 미래의 값을 예측하려 한다. 또한 이는 계절조정을 가능하게 하여 장기적인 추세의 분석을 가능하게 한다.

## 4.1 추세모형에 의한 분해

추세모형에 의한 분해를 하는 경우, 아래처럼 분해한다.

$$T_t = \beta_0 + \sum_{i=1}^{k} \beta_i t^i$$

$$S_t = \sum_{i=1}^{s} \delta_i I(t \equiv i \ (\mathrm{mod} s))$$

이를 바탕으로 아래처럼 각 성분을 추정한다.

1. 원시계열 $Z_t$에 추세모형 $Z_t = \beta_0 + \sum_{i=1}^{k} \beta_i t^i + \epsilon_t$를 얻어 얻은 추정식으로부터 추세성분의 추정계열 $\hat{T}_t = \hat{\beta}_0 + \sum_{i=1}^{k} \hat{\beta}_i t^i$를 얻는다.

2. 원시계열 $Z_t$에서 추정계열 $\hat{T}_t$를 뺀 잔차계열에 계절모형을 적합시키고, 그 추정식으로부터 계절성분의 추정계열 $\hat{S}_t = \sum_{t=1}^{s} \hat{\delta}_i I(t \equiv i \ (\mathrm{mod} s))$를 얻는다.

3. 불규칙성분의 추정계열 $\hat{I}_t = Z_t - \hat{T}_t - \hat{S}_t$에 대한 분석을 통해 추가적인 정보가 남아 있는지 진단한다.

삼각함수를 이용한 계절모형을 적합하는 경우 2단계의 적합 방식만 바꾸면 되고, 승법모형을 이용하는 경우 로그를 취하면 가법모형이 되므로 이 분석법을 그대로 이용할 수 있다. 만약 경기변동 등을 설명하는 추가적인 순환성분 $C_t$가 있다면, 그 주기를 적절히 찾고, 그에 맞게 $\hat{I}_t$에서 추가적으로 조정해주는 것이 필요하다.

## 4.2 이동평균법에 의한 분해

가장 간단한 $Z_t = \beta_0 + \epsilon_t$ 모형을 생각할 때, 지수평활법을 사용하는 이유는 $\beta_0$가 변할 수 있기 때문이었다. 이를 반영하기 위해서는 관측값 모두를 이용하지 않고 최근 관측값만을 이용하는 **이동평균법**이 고려될 수 있다. 예측의 목적으로 이동평균법을 이용하는 경우, $m$개의 관측값만을 이용하여

$$M_n^{(1)} = \sum_{t=n-m+1}^{n} \frac{Z_t}{m}$$

을 $Z_n(l)$으로 사용할 수 있다. 이는 기대값이 $\beta_0$인 불편추정량이다. $m$이 클수록 평활의 수준이 커 국지적인 변화에 둔감하게 반응하나 안정적이고, $m$이 작으면 그 변화를 빨리 반영하나 불안정하다. $Z_{n+1}$이 관측되면

$$M_{n+1}^{(1)} = M_n^{(1)} + \frac{Z_{n+1} - Z_{n-m+1}}{m}$$

으로 쉽게 갱신할 수 있다.

만약 평활의 목적으로 이동평균법을 사용한다면, 중심이동평균을 통해 $m = 2l + 1$일 때

$$M_n^{(1)} = \frac{Z_{n-l} + \cdots + Z_{n+l}}{2l+1}$$

을 사용할 수 있다. 특히 계절성분이 있는 $Z_t = \beta_0 + S_t + \epsilon_t$ 모형에서, 주기가 $s$이므로 $s$항 이동평균을 수행하면 계절에 의한 변동이 제거되고 불규칙성분 역시 평활되는 특성을 가지고 있다. 이를 종합하면,

$$Z_t = T_t + S_t + C_t + I_t$$

라는 가법모형에서 이동평균법을 이용한 분해법은 아래와 같은 과정을 따른다.

1. 원시계열 $Z_t$의 계절성분 주기 $s$에 적절한 이동평균을 적용하여 계절성분과 불규칙성분이 제거된 추세+순환성분 $\widehat{T_t + C_t}$를 얻는다.

2. 원시계열 $Z_t$에서 $\widehat{T_t + C_t}$를 빼 계절+불규칙성분 $\widehat{S_t + I_t}$를 얻는다.

3. $\widehat{S_t + I_t}$에 $s$의 배수가 아닌 항의 이동평균을 적용하여 불규칙성분만 제거하고 계절성분의 추정계열 $\hat{S}_t$를 얻는다.

4. 추세+순환성분을 종속변수로 하는 추세다항식을 적합시켜 $\hat{T}_t$를 얻는다.

5. $\widehat{T_t + C_t}$에서 $\hat{T}_t$를 빼어 $\hat{C}_t$를 얻는다.

6. 불규칙성분의 추정계열 $\hat{I}_t = Z_t - \hat{T}_t - \hat{C}_t - \hat{S}_t$를 구하고, 해당 시계열을 진단하여 정보가 남아 있는지 확인한다.

이를 이용하여 시점 $n$에서의 $l$시차 이후 예측값을

$$\hat{Z}_n(l) = \hat{T}_n(l) + \hat{S}_n(l) + \hat{C}_n(l)$$

으로 얻게 된다. $\hat{T}_n(l)$은 적합된 추세모형에 $t = n + l$을 대입하여 얻고, $\hat{S}_n(l)$은 $n + l$과 주기가 같은 계절성분을 통해 얻는다. $\hat{C}_n(l)$ 역시 $\hat{C}_t$에 적절한 주기모형을 적합한 뒤 그 적합을 통해 얻는다. 승법모형의 경우 뺄셈을 모두 나눗셈으로 대체하는 등의 방법을 통해 동일한 방식으로 분해가 가능하다.

## 4.3 계절조정

### 4.3.1 전년동기비

승법모형 $Z_t = T_t S_t C_t I_t$를 가정하는 경우, **계절조정**은

$$Z_t^{(sa)} = \frac{Z_t}{S_t}$$

를 얻는 과정을 의미한다. 승법모형 하에서 전년동기비는

$$r_t = \frac{Z_t}{Z_{t-12}} = \frac{T_t}{T_{t-12}} \times \frac{C_t}{C_{t-12}} \times \frac{S_t}{S_{t-12}} \times \frac{I_t}{I_{t-12}}$$

로 얻어지며, 만약 계절성분과 불규칙성분이 안정적이어 $S_t \approx S_{t-12}, I_t \approx I_{t-12}$라면 추세+순환성분의 변동만 잘 반영함을 알 수 있다. 즉 계절조정을 수행하는 것이다.

### 4.3.2 계절조정을 위한 이동평균법

계절조정을 위해서는 기본적으로 원시계열에서 계절성분의 크기를 추정하는 것이 중요하다. 이동평균법으로써 이를 수행했었으므로, 계절조정 역시 X-12 ARIMA와 같은 다양한 이동평균법을 통해 수행될 수 있다. 먼저, 다양한 이동평균법에 대해 알아보자.

- 대칭 $(2d+1)$항 이동평균:

$$\mathrm{MA}_t(2d+1) = \sum_{j=-d}^{d} \frac{Z_{t+j}}{2d+1}$$

- 비대칭 $(2d)$항 이동평균:

$$\mathrm{MA}_{t,a}(2d) = \sum_{j=-d+1}^{d} \frac{Z_{t+j}}{2d}$$

- 대칭 $(2d+1)$항 가중이동평균:

$$\mathrm{MA}_{t,w}(2d+1) = \sum_{j=-d}^{d} w_j Z_{t+j}, \quad \sum_{j=-d}^{d} w_j = 1,\ w_j = w_{-j}$$

- $(a \times (2d+1))$항 이동평균:

$$\mathrm{MA}_t(a \times (2d+1)) = \sum_{j=-d}^{d} \frac{\mathrm{MA}_{t+j}(a)}{2d+1}$$

- $(a \times 2d)$항 이동평균:

$$\mathrm{MA}_t(a \times (2d)) = \sum_{j=-d+1}^{d} \frac{\mathrm{MA}_{t+j}(a)}{2d}$$

- Henderson의 대칭 가중이동평균:

$$\mathrm{MA}_{t,H}(2d+1) = \sum_{j=-d}^{d} w_j Z_{t+j}$$

### 4.3.3 X-11 ARIMA와 X-12 ARIMA

X-11 ARIMA와 X-12 ARIMA는 계절조정을 위해 사용되는 프로시져로, 실무에서 자주 사용된다. X-11 ARIMA에서는 시계열자료에 적절한 모형을 적합시켜 자료 시작점 이전 자료를 후향예측하고 끝점 이후 자료를 예측하여 확장된 자료를 이동평균함으로써 계절조정을 수행한다. 이외에도 사전조정이나 그 통계적 평가 등이 포함되어 있다. X-12 ARIMA는 그 확장판으로, 아래의 승법모형을 가정한다.

$$Z_t = T_t C_t S_t D_t H_t O_t I_t$$

여기에서 $D_t$는 **거래일성분**, $H_t$는 **명절성분**, $O_t$는 **이상점**에 의한 효과를 나타낸다. 이를 분해하는 과정을 알아보자.

1. **RegARIMA 모형을 이용한 사전조정**: 거래일, 명절 효과 등을 제거하기 위하여, 이들을 설명변수로 하는 RegARIMA 모형을 $Z_t$에 적합시킨다. 이상점 역시 여기에서 제거한다. 이들 성분들을 제거함으로써 사전조정된 안정적인 원계열을 산출하며, 적합된 모형을 바탕으로 시계열의 양끝을 연장하여 결측치 문제를 제거한다.

2. **시계열 구성성분 분해과정**: 이 과정은 사전조정된 시계열을 추세, 순환, 계절, 불규칙성분으로 분해한다. 먼저 중심 12항 이동평균을 수행해 추세+순환성분을 산출한 뒤, Henderson의 가중이동평균을 통해 잠정 계절조정계열을 개선한다. 그 이후 계절+불규칙성분에 $(3 \times 3), (3 \times 5), (3 \times 7), (3 \times 9)$항 이동평균을 적용하여 계절성분을 얻는다. 여기에는 이상점 제거 과정 등이 포함된다.

3. **통계적 평가과정**: Q-통계량, M1-M11 품질관리 통계량, 스펙트럼 추정량, 슬라이딩 스팬, 리비젼 히스토리 진단 등이 수행되어 불규칙성분에 추가적인 자료가 없는지 살펴본다.

# Chapter 5

# 확률과정

시계열분석에서는 **이산형 확률과정**을 통해 모형을 세운다. 우리가 관측하는 것은 해당 확률과정의 표본경로 중 하나이다.

**Definition 3.** *시계열이* ***정상성****을 가진다는 것은 시계열의 확률적인 성질들이 시간의 흐름에 불변함을 의미한다.* ***강한 정상성****은 임의의 자연수* $t_1, t_2, \cdots, t_n$*과* $k$*에 대하여*

$$(Z_{t_1}, Z_{t_2}, \cdots, Z_{t_n}) \stackrel{d}{\equiv} (Z_{t_1+k}, Z_{t_2+k}, \cdots, Z_{t_n+k})$$

*임을 의미한다. 한편* ***약한 정상성****, 혹은* ***공분산정상성****은 모든 시점* $t$*에서*

$$\begin{aligned} \mu &= \mathbb{E}[Z_t] \\ \sigma^2 &= \mathrm{Var}(Z_t) \\ \gamma_h &= \mathrm{Cov}(Z_t, Z_{t+h}) \end{aligned}$$

*가 성립함을 의미한다. 특히* $\gamma_h$*는* ***자기공분산****이라 부른다. 일반적으로 정상성이라 하면, 약한 정상성을 의미하며, 정상 확률과정은 정상성을 만족하는 확률과정을 의미한다. 강한 정상성과 약한 정상성은 둘 중 하나가 서로를 포함하는 개념이 아니다.*

**Definition 4.** $\epsilon_t$*가* $W.N.(0, \sigma^2)$*을 따를 때,* ***확률보행과정****은*

$$Z_t = Z_{t-1} + \epsilon_t$$

*에 의하여 정의된다. 이는* $Var(Z_t) = t\sigma^2$*으로 분산이 일정하지 않기에,* ***비정상 확률과정****이다.* ***절편이 있는 확률보행과정****은*

$$Z_t = \delta + Z_{t-1} + \epsilon_t$$

*에 의해 정의되어 추세 역시 있다.*

**Definition 5.** $\epsilon_t$*가* $W.N.(0, \sigma^2)$*을 따를 때,* ***선형과정****은*

$$Z_t = \mu + \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$$

*처럼 정의되는 정상확률과정으로, 특히* $\sum_{j=0}^{\infty} \psi_j^2$*가 유한해야 한다.*

## 5.1 자기상관함수

**Definition 6.** *정상시계열에서, 자기공분산함수는*

$$\gamma_k = \mathrm{Cov}(Z_t, Z_{t+k})$$

*, **자기상관함수** ACF는*

$$\rho_k = Corr(Z_t, Z_{t+k}) = \frac{\gamma_k}{\gamma_0}$$

*으로 정의한다. 이는 대칭성을 가져* $\rho_k = \rho_{-k}$*를 만족한다.*
*표본에서는 이를 **표본자기공분산함수***

$$\hat{\gamma}_k = \frac{1}{n}\sum_{t=1}^{n-k}(Z_t - \bar{Z})(Z_{t+k} - \bar{Z})$$

*와 **표본자기상관함수** SACF*

$$\hat{\rho}_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0}$$

*로 추정할 수 있다.*

**Theorem 1.** $n$*이 충분히 클 때* $\hat{\rho}_k$*는 점근적으로 정규분포를 따르며,* $i = 1, 2, \cdots, q$*에 대해* $\rho_i \neq 0$*이고* $k > q$*에서* $\rho_k = 0$*이라면*

$$\mathrm{Var}(\hat{\rho}_k) \approx \frac{1}{n}\left(1 + 2\sum_{i=1}^{q}\rho_i^2\right)$$

*임이 알려져 있다. 따라서* $H_0 : \rho_k = 0$*의 검정은* $\hat{\rho}_k$*의 절대값이* $\frac{2}{\sqrt{n}}$*보다 큰지를 확인함으로써 수행된다.*

## 5.2 부분자기상관함수

**Definition 7.** ***부분자기상관함수** PACF는 아래의 Yule-Walker 방정식*

$$\begin{pmatrix} \rho_0 & \rho_1 & \rho_2 & \cdots & \rho_{k-1} \\ \rho_1 & \rho_0 & \rho_1 & \cdots & \rho_{k-2} \\ \rho_2 & \rho_1 & \rho_0 & \cdots & \rho_{k-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{k-1} & \rho_{k-2} & \rho_{k-3} & \cdots & \rho_0 \end{pmatrix} \begin{pmatrix} \phi_{k1} \\ \phi_{k2} \\ \phi_{k3} \\ \vdots \\ \phi_{kk} \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_k \end{pmatrix}$$

*를 풀어 얻는* $\phi_{kk}$*으로,*

$$\phi_{kk} = Corr(Z_t - \mathrm{proj}_{Z_{t+1},\cdots,Z_{t+k-1}}(Z_t), Z_{t+k} - \mathrm{proj}_{Z_{t+1},\cdots,Z_{t+k-1}}(Z_{t+k}))$$

*과도 같다. **표본부분자기상관함수** SPACF는 Yule-Walker 방정식에서* $\rho_j$*를* $\hat{\rho}_j$*로 대체하여 얻은* $\hat{\phi}_{kk}$*를 의미한다.*

**Theorem 2.** ***Durbin-Levinson 알고리즘**은 PACF를 쉽게 구하는 알고리즘은 아래와 같게 이루어진다.*

$$\hat{\phi}_{11} = \hat{\rho}_1$$

$$\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^{k}\hat{\phi}_{kj}\hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^{k}\hat{\phi}_{kj}\hat{\rho}_j}$$

$$\hat{\phi}_{k+1,j} = \hat{\phi}_{kj} - \hat{\phi}_{k+1,k+1}\hat{\phi}_{k,k+1-j}$$

# Chapter 6

# 자기회귀이동평균과정

## 6.1 자기회귀과정

**Definition 8.** *어떠한 확률과정 $Z_t$가*

$$Z_t = f(Z_{t-1}, Z_{t-2}, \cdots) + \epsilon_t, \quad \epsilon_t \sim W.N.(0, \sigma^2)$$

*의 관계를 만족하면,* ***자기회귀과정****이라 부른다.*

일반적으로 평균이 $\mu$인 정상 $p$차 자기회귀과정 AR($p$) 과정은

$$Z_t - \mu = \sum_{j=1}^{p} \phi_j (Z_{t-j} - \mu) + \epsilon_t$$

처럼 쓴다.

**Definition 9.** *아래와 같은* ***후진작용소****를 이용하면 모형을 쉽게 쓸 수 있다.*

$$BZ_t = Z_{t-1}, \quad B^j Z_t = Z_{t-j}$$

*$B$ 대신 $L$을 사용하기도 한다.*

따라서 p는 $\delta = (1 - \phi_1 - \cdots - \phi_p)\mu$를 이용하여

$$Z_t = \delta + \sum_{j=1}^{p} \phi_j Z_{t-j} + \epsilon_t = \delta + \Phi(B) Z_t + \epsilon_t$$

처럼 쓰기도 한다. 이때

$$\Phi(B) = \phi_1 B + \phi_2 B^2 + \cdots + \phi_p B^p$$

인 작용소 $B$의 다항식이다.

**Theorem 3.** *AR(1) 모형*

$$Z_t = \delta + \phi Z_{t-1} + \epsilon_t$$

*가 정상시계열일 조건은 $|\phi| < 1$이다. 또한*

$$\mu = \frac{\delta}{1 - \phi}$$

$$\gamma_0 = \frac{\sigma^2}{1 - \phi^2}$$

$$\gamma_k = \phi^k \frac{\sigma^2}{1-\phi^2}$$
$$\rho_k = \phi^k$$

*이다. 따라서 AR(1) 모형의 ACF는 지수적으로 감소하는 특성을 가지고 있다. 반면 PACF의 경우* $\phi_{11} = \rho_1 = \phi$*이지만,* $k \geq 2$*에서* $\phi_{kk} = 0$*으로 시차 2 이후로는 절단되는 형태를 가진다.*

**Theorem 4.** *AR(2) 모형*

$$Z_t = \delta + \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \epsilon_t$$

*가 정상시계열일 조건은* $\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 = 0$*의 근* $z$*가 모두 단위원 밖에 있는 것이다. 즉*

$$\phi_1 + \phi_2 < 1$$
$$\phi_2 - \phi_1 < 1$$
$$-1 < \phi_2 < 1$$

*인 것이다. 또한*

$$\mu = \frac{\delta}{1-\phi_1-\phi_2}$$
$$\gamma_0 = \frac{1-\phi_2}{1+\phi_2}\frac{\sigma^2}{(1-\phi_2)^2-\phi_1^2}$$
$$\rho_1 = \frac{\phi_1}{1-\phi_2}$$
$$\rho_k = \phi_1\rho_{k-1} + \phi_2\rho_{k-2} \quad (k \geq 2)$$
$$\phi_{11} = \rho_1$$
$$\phi_{22} = \phi_2$$
$$\phi_{kk} = 0 \quad (k \geq 3)$$

*으로, ACF는 지수적으로 감소하고 PACF는 시점 3 이후로 절단되는 형태를 보인다.*

**Theorem 5.** *일반적인* $AR(p)$ *모형*

$$\Phi(B)Z_t = \delta + \epsilon_t$$

*에서 정상성 조건은* ***특성방정식*** $\Phi(z) = 0$*의 모든 근* $z$*가 단위원 밖에 있는 것이다. 이는* $Z_t$*가* ***short-memory****를 가져 이전 관측값의 영향이 빠르게 사라짐과도 동치이다. ACF는 지수적으로 감소하며, 일반적으로* ***Yule-Walker 방정식***

$$\begin{pmatrix} \rho_0 & \rho_1 & \rho_2 & \cdots & \rho_{p-1} \\ \rho_1 & \rho_0 & \rho_1 & \cdots & \rho_{p-2} \\ \rho_2 & \rho_1 & \rho_0 & \cdots & \rho_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \cdots & \rho_0 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \\ \phi_p \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_p \end{pmatrix}$$

*에 의해 PACF를 구하면 시점* $p+1$ *이후로 절단되는 것을 알 수 있다.*

## 6.2 이동평균과정

**Definition 10.** ***이동평균과정****은 선형과정 중 유한한* $\epsilon_{t-j}$*들의 결합으로 이루어지는 과정을 말한다. 대표적으로* $MA(q)$ *모형은*

$$Z_t = \mu + \epsilon_t - \theta_1\epsilon_{t-1} - \cdots - \theta_q\epsilon_{t-q} = \mu + \Theta(B)\epsilon_t$$

*와 같은 형태를 가진다. 이동평균과정은 항상 정상성을 가진다.*

**Theorem 6.** *MA(1) 과정은*

$$Z_t = \mu + \epsilon_t - \theta\epsilon_{t-1}$$

*과 같은 형태를 가지며,*

$$\gamma_k = \begin{cases} (1+\theta^2)\sigma^2 & k=0 \\ -\theta\sigma^2 & k=1 \\ 0 & k \geq 2 \end{cases}$$

*이다. 따라서 ACF는*

$$\rho_k = \begin{cases} -\dfrac{\theta}{1+\theta^2} & k=1 \\ 0 & k=2 \end{cases}$$

*으로 시차 2 이후로 절단되는 형태가 되며, PACF는 Yule-Walker 방정식에 따라*

$$\phi_{kk} = \theta^k \times \frac{-(1-\theta^2)}{1-\theta^{2(k+1)}}$$

*으로 지수적으로 감소한다.*

**Definition 11.** *MA 과정에서* ***가역성****은 MA 모형의 백색잡음 $\epsilon_t$ $t$ 시점 이전의 $Z_t$들의 결합으로써 표현할 수 있는 성질을 의미한다. 즉*

$$\sum_{j=0}^{\infty} \pi_j X_{t-j} = \epsilon_t$$

*인 $\pi_j$들이 존재하면 가역적인 MA 과정이다. MA(1) 모형의 경우, $|\theta| < 1$이 가역성 조건이다.*

**Theorem 7.** *MA(2) 과정은*

$$Z_t = \mu + \epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2}$$

*와 같은 형태를 가지며,*

$$\gamma_k = \begin{cases} (1+\theta_1^2+\theta_2^2)\sigma^2 & k=0 \\ -\theta_1(1-\theta_2)\sigma^2 & k=1 \\ -\theta_2\sigma^2 & k=2 \\ 0 & k=3 \end{cases}$$

*이다. 따라서 ACF는*

$$\rho_k = \begin{cases} \dfrac{-\theta_1+\theta_1\theta_2}{1+\theta_1^2+\theta_2^2} & k=1 \\ -\dfrac{\theta_2}{1+\theta_1^2+\theta_2^2} & k=2 \\ 0 & k=3 \end{cases}$$

*으로 시차 3 이후로 절단되는 형태가 되며, PACF는 지수적으로 감소한다. 이 시계열이 가역일 필요충분조건은 $\Theta(z) = 1 - \theta_1 z - \theta_2 z^2 = 0$의 $z$가 모두 단위원 밖에 있는 것이다.*

**Definition 12.** *AR($p$) 과정의 ACF와 PACF는 $MA(p)$ 과정의 PACF와 ACF 형태와 유사하다. 이러한 성질을 AR과정과 MA과정의* ***쌍대성****이라 부른다.*

**Theorem 8.** *$MA(q)$ 과정*

$$Z_t = \mu - \theta_1 \epsilon_{t-1} - \cdots - \theta_q \epsilon_{t-q} = \mu + \Theta(B)\epsilon_t$$

*이 가역일 필요충분조건은 $\Theta(z) = 0$의 모든 근 $z$가 단위근 밖에 있는 것이다. 이 시계열에서*

$$\gamma_k = \begin{cases} (-\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q)\sigma^2 & k = 1, 2, \cdots, q \\ 0 & k \geq q+1 \end{cases}$$

*이며, ACF는*

$$\rho_k = \begin{cases} \dfrac{-\theta_k + \theta_1\theta_{k+1} + \cdots + \theta_{q-k}\theta_q}{1 + \theta_1^2 + \cdots + \theta_q^2} & k = 1, 2, \cdots, q \\ 0 & k \geq q+1 \end{cases}$$

*이다. 앞에서와 마찬가지로 ACF는 $q+1$ 이후로 절단되는 형태이며, PACF는 지수적으로 감소한다.*

## 6.3 자기회귀이동평균과정

**Definition 13.** *시계열을 묘사함에 있어* ***모수의 절약****을 하기 위하여, 자기회귀 부분과 이동평균 부분을 동시에 포함하는* ***자기회귀이동평균*** *과정 $ARMA(p, q)$를 세울 수 있다. 그 형태는*

$$Z_t = \delta + \phi_1 Z_{t-1} + \cdots + \phi_p Z_{t-p} + \epsilon_t - \theta_1\epsilon_{t-1} - \cdots - \theta_q\epsilon_{t-q}$$

*혹은*

$$\Phi(B)\dot{Z}_t = \Theta(B)\epsilon_t, \quad \dot{Z}_t = Z_t - \mu$$

*이다. 이때*

$$\mu = \frac{\delta}{1 - \phi_1 - \cdots - \phi_p}$$

*이다. $ARMA(p, q)$ 과정이 정상적일 조건은 $\Phi(z) = 0$의 모든 근 $z$가 단위원 밖에 있는 것이며, 가역적일 조건은 $\Theta(z)$의 모든 근 $z$가 단위원 밖에 있는 것이다.*

*한편 ACF와 PACF는 모형에 따라 다르나, 일반적으로 ACF는 시차 $q-p+1$ 이후로는 지수적으로 감소하는 형태이며, PACF는 시차 $p-q+1$ 이후로는 지수적으로 감소하는 형태이다.*

# Chapter 7

# 비정상시계열의 분석

만약 시간에 따라 시계열의 평균수준이 달라지거나 분산이 달라진다면, 그 시계열이 **비정상시계열**이라 한다. 추세를 갖거나 계절성을 갖는 시계열이 대표적인 비정상시계열이다. 이때 추세에는 두 가지 종류가 있다.

- **결정적 추세**: 결정적이고 영원히 지속되는 추세
- **확률적 추세**: 자료 간의 강한 양의 상관관계 때문에 관찰되는 추세

## 7.1 비정상시계열의 정상화

### 7.1.1 분산이 일정하지 않은 경우

시계열의 분산이 일정하지 않은 경우, 변동폭을 안정화하기 위하여 **분산안정화변환**을 수행할 수 있다. 예를 들어, 분산을 상수화하기 위해 로그변환이나 제곱근변환 등을 사용할 수 있다. 대표적으로 많이 사용되는 것은 $\lambda$에 대한 **Box-Cox** 변환으로,

$$y(\lambda) = \begin{cases} \dfrac{y^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \log y & \lambda = 0 \end{cases}$$

으로 정의된다.

### 7.1.2 수준이 일정하지 않은 경우

수준이 안정하지 않은 경우에는 적절한 추세를 밝혀낸 뒤 이를 제거하는 것이 필요하다. 먼저 결정적 추세를 갖는 경우,

$$Y_t = T_t + S_t + I_t$$

로 모형화하자. 그 다음 $I_t$는 불규칙성분,

$$T_t = \beta_0 + \beta_1 t + \cdots + \beta_k t^k$$
$$S_t = \sum_{i=1}^{s-1} \gamma_i I(t = i \ (\text{mod} s))$$

로 모형화한 뒤 이들을 추정하여 $\hat{I}_t = Y_t - \hat{T}_t - \hat{S}_t$ 를 얻는다. 이제 이는 평균이 0인 정상시계열의 형태를 보일 것으로 기대된다. 이처럼 적절한 설명변수를 통해 추세를 모형화하는 방식을 **자기상관오차를 갖는 회귀모형**이라 부른다.

한편 확률적 추세를 갖는 경우에는 추세모형으로 이를 해결할 수 없다. 이처럼 국지적으로 다른 시간대에서는 시계열의 움직임이 다르지만 본질적으로 시계열의 움직임이 같은 경우에는 **동질적 비정상성**을 가진다고

말하며, 이때는 차분을 통해 정상시계열로 바꾸어줄 수 있다. 차분연산자 $\Delta$는

$$\Delta Y_t = Y_t - Y_{t-1}$$

처럼 정의한다. 일반적으로 최고차항이 $k$인 결정적 추세를 가지면, $k$번 차분할 경우 정상시계열이 됨이 알려져 있다.

## 7.2 단위근 검정

확률보행과정을 포함한 많은 비정상과정들의 경우,

$$\Phi(B)Z_t = \epsilon_t$$

처럼 묘사되는 상황에서 $\Phi(z) = 0$의 근 중 $|z| = 1$인 근인 **단위근**이 존재한다는 것이 특징이다. 따라서 차분을 통해 확률적 추세를 보정하는 경우, 적절히 차분한 뒤 단위근 검정을 수행하여 정상시계열이 되었는지 확인하는 과정이 필요하다.

### 7.2.1 Diceky-Fuller 검정

다음과 같은 AR(1) 과정을 생각해 보자.

$$Z_t = \phi Z_{t-1} + \epsilon_t$$

이때 $\epsilon_t \sim W.N.(0, \sigma^2)$이다. 잘 알려져 있듯, $|\phi| < 1$일 때에 이 시계열은 정상시계열이 된다.
한편 단위근검정에서는 $\phi = 1$인지를 판정하려 한다. 일반적으로 그 검정에는 $\phi$의 최소제곱추정량 $\hat{\phi}$를 이용한다. 이때 아래의 세 가지 모형이 존재하며, 각 모형 하에서 $\hat{\phi}$의 분포가 다르기에 자료의 성질에 맞는 검정법을 택해야만 한다.

**Case 1**
실제 모형: $Z_t = Z_{t-1} + \epsilon_t$ (절편이 없는 확률보행과정)
적합 모형: $Z_t = \phi Z_{t-1} + \epsilon_t$ (절편이 없는 AR(1) 과정)
이때 귀무가설 $\phi = 1$ 하에서 $\phi$-statistic은 Wiener process $W$에 대하여

$$T(\hat{\phi} - 1) = \frac{T^{-1}\sum_{t=2}^{T} Z_{t-1}Z_t}{T^{-2}\sum_{t=2}^{T} Z_{t-1}^2} \xrightarrow{d} \frac{\int_0^1 W(r)dW}{\int_0^1 W(r)^2 dr}$$

, $t$-statistic은

$$t_T = \frac{\hat{\phi} - 1}{\hat{\sigma}} \xrightarrow{d} \frac{\int_0^1 W(r)dW}{\left(\int_0^1 W(r)^2 dr\right)^{1/2}}$$

으로 점근분포를 가진다. 따라서 이를 바탕으로 분포표에서 분위수를 확인하여 각 통계량의 값이 충분히 작으면 귀무가설 $H_0 : \phi = 1$을 기각할 수 있다.

**Case 2**
실제 모형: $Z_t = Z_{t-1} + \epsilon_t$ (절편이 없는 확률보행과정)
적합 모형: $Z_t = \delta + \phi Z_{t-1} + \epsilon_t$ (절편이 있는 AR(1) 과정)
실제 데이터가 절편이 없는 확률보행과정을 따르는데 이를 절편이 있는 과정으로 적합하는 경우, $\hat{\phi}$에 대한

두 통계량은 아래의 점근분포를 가진다.

$$T(\hat{\phi}-1) \xrightarrow{d} \frac{\int_0^1 W(r)dW - W(1)\int_0^1 W(r)dr}{\int_0^1 W(r)^2dr - \left(\int_0^1 W(r)^2dr\right)^2}$$
$$t_T \xrightarrow{d} \frac{\int_0^1 W(r)dW - W(1)\int_0^1 W(r)dr}{\left(\int_0^1 W(r)^2dr - \left(\int_0^1 W(r)^2dr\right)^2\right)^{1/2}}$$

한편 일반적으로는 귀무가설로 $H_0 : \phi = 1$보다는 $H_0 : \phi = 1, \delta = 0$을 사용한다. 왜냐하면 $H_0 : \phi = 1$은 절편이 있는 확률보행과정 역시 허용하는데, 이는 결정적 추세를 가지게 되어 실제 데이터와 멀어지기 때문이다. 따라서 이에 대해 OLS F검정통계량을 통해 동시검정을 하는 방법이 더욱 적절하다고 생각된다. 같은 이유로, 아래 역시 허용되지 않는다.
실제 모형: $Z_t = \delta + Z_{t-1} + \epsilon_t$ (절편이 있는 확률보행과정; 결정적 추세 **존재**)
적합 모형: $Z_t = \delta + \phi Z_{t-1} + \epsilon_t$ (절편이 있는 AR(1)과정; 결정적 추세 **부재**)

**Case 3**
실제 모형: $Z_t = \delta + Z_{t-1} + \epsilon_t$ (절편이 있는 확률보행과정; 결정적 추세 존재)
적합 모형: $Z_t = \delta + \phi Z_{t-1} + \beta t + \epsilon_t$ (절편과 **추세**가 포함된 모형)
이때에도 마찬가지로 $T(\hat{\phi}-1)$과 $t_T$는 적절한 수렴성을 갖는다. 여기에서도 마찬가지로, $\phi = 1$과 $\beta = 0$을 동시에 검정하는 상황이 더욱 선호된다. 이때 $F$검정통계량을 마찬가지로 사용한다. 주의할 것은 **Case 2**와 **Case 3** 모두에서 $F$통계량은 그 형태만 $F$통계량의 형태일 뿐 실제로 $F$분포를 따르지는 않으며, Dickey와 Fuller가 시뮬레이션을 통해 얻은 별도의 분포표를 따라 검정해야 한다. 이 경우의 검정은 시계열의 추세가 존재할 때 이 추세가 절편 $\delta$에 의한 것인지, 혹은 실제적인 추세 $\beta t$의 존재에 따른 것인지를 파악하는 효과도 있다.
따라서 일반적으로 데이터를 보고 나서, 평균수준이 0으로 생각되면 **Case 1**을, 평균수준이 0이 아니나 추세가 없다면 **Case 2**를, 추세가 존재하는 경우 **Case 3**을 선택하여 검정하는 것이 올바르다.

### 7.2.2 Augmented Dickey-Fuller 검정

만약 시계열자료가 AR($p$)를 따르는 경우, Dickey-Fuller 검정을 적용하기에 무리가 있다. 시계열자료가 AR($p$)를 따른다면,

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \cdots + \phi_p Z_{t-p} + \epsilon_t$$

로 쓸 수 있을 것이다. 이때 $\phi = \phi_1 + \phi_2 + \cdots + \phi_p$, $\xi_j = -[\phi_{j+1} + \phi_{j+2} + \cdots + \phi_p]$를 정의하면, **ADF representation**

$$Z_t = \phi Z_{t-1} + \xi_1 \Delta Z_{t-1} + \xi_2 \Delta Z_{t-2} + \cdots + \xi_{p-1}\Delta Z_{t-p+1} + \epsilon_t$$

를 얻게 된다.
한편 이 시계열이 단위근을 가질 경우, 특성방정식

$$\Phi(z) = 1 - \phi_1 z - \phi_2 z^2 - \cdots - \phi_p z^p$$

이 근 $z = 1$을 가지므로

$$\phi = \phi_1 + \phi_2 + \cdots + \phi_p = 1$$

을 얻는다. 따라서 ADF representation 하에서 $\phi = 1$을 검정하는 것이 단위근검정이 된다.
만약 귀무가설이 참이라면, 즉 $\phi = 1$이라면,

$$\Delta Z_t = Z_t - \phi Z_{t-1} = \xi \Delta Z_{t-1} + \cdots + \xi_{p-1}\Delta Z_{t-p+1} + \epsilon_t$$

으로 $\Delta Z_t$가 AR$(p-1)$ 과정을 따르게 된다. 한편 AR(1) 모형 하에서

$$Z_t = \phi Z_{t-1} + \epsilon_t$$

를 검정할 때, $\phi = 1$이라면 $\epsilon_t = \Delta Z_t$이다. 즉 $\epsilon_t$의 독립성을 가정하는 일반적인 DF 검정과 달리, ADF 검정에서는 오차항에 자기상관이 존재할 수 있다고 가정함으로써 이를 AR$(p)$ 모형에까지 확장하였다고 볼 수 있다.
한편 여기에서도 마찬가지로 세 가지의 케이스가 존재한다. DF 검정의 스피릿과 동일하게, 서로 비교가능한 실제 모형과 적합모형 사이에서 검정을 수행한다. 단지 다른 것은 오차항의 자기상관성을 허용함에 따라 검정통계량이 살짝 달라졌다는 것뿐이다.

**Case 1**
실제 모형: $Z_t = Z_{t-1} + \xi_1 \Delta Z_{t-1} + \cdots + \xi_{p-1}\Delta Z_{t-p+1} + \epsilon_t$
적합 모형: $Z_t = \phi Z_{t-1} + \xi_1 \Delta Z_{t-1} + \cdots + \xi_{p-1}\Delta Z_{t-p+1} + \epsilon_t$
이때 검정통계량은 아래와 같다.

$$Z_\phi = \frac{T(\hat{\phi}-1)}{1-\hat{\xi}_1 - \cdots - \hat{\xi}_{p-1}} \xrightarrow{d} \frac{\int_0^1 W(r)dr}{\int_0^1 W(r)^2 dr}$$
$$t_T = \frac{\hat{\phi}-1}{\hat{\sigma}_{\hat{\phi}}} \xrightarrow{d} \frac{\int_0^1 W(r)dr}{\left(\int_0^1 W(r)^2 dr\right)^{1/2}}$$

즉 동일한 검정통계량의 분포를 따르지만, $Z_\phi$에서 오차항의 자기상관을 고려하여 $1-\hat{\xi}_1-\cdots-\hat{\xi}_{p-1}$로 나눠준 것만 다르다.

**Case 2**
실제 모형: $Z_t = Z_{t-1} + \xi_1 \Delta Z_{t-1} + \cdots + \xi_{p-1}\Delta Z_{t-p+1} + \epsilon_t$
적합 모형: $Z_t = \delta + \phi Z_{t-1} + \xi_1 \Delta Z_{t-1} + \cdots + \xi_{p-1}\Delta Z_{t-p+1} + \epsilon_t$
동일한 방식으로 $Z_\phi$와 $t_T$를 구할 수 있으며, 동일한 조정을 취하면 DF 검정의 **Case 2**와 동일한 극한분포를 가진다. OLS F검정 역시 마찬가지로 DF 검정에서와 동일하다.

**Case 3**
실제 모형: $Z_t = \delta + Z_{t-1} + \xi_1 \Delta Z_{t-1} + \cdots + \xi_{p-1}\Delta Z_{t-p+1} + \epsilon_t$
적합 모형: $Z_t = \delta + \phi Z_{t-1} + \beta t + \xi_1 \Delta Z_{t-1} + \cdots + \xi_{p-1}\Delta Z_{t-p+1} + \epsilon_t$
검정통계량과 OLS F검정통계량의 분포는 역시 적절한 조정을 취한다면 DF 검정의 **Case 3**과 같다.

### 7.2.3 Phillips-Perron 검정

DF 검정에서 ADF 검정으로 발전하며 AR$(p)$를 가정하여 단위근을 검정하는 방법을 배웠다. 이는 사실상 오차항 $\epsilon_t$의 독립성 가정을 포기하고 자기상관을 허용함으로써 이루어진 것이다. PP 검정에서는 이를 더욱 확장하여 오차항 $\epsilon_t$의 자기상관성을 AR$(p-1)$ 모형 대신 MA$(\infty)$ 모형을 통해 처리한다.
먼저 **Beverage-Nelson 분해**에 대해 알아 보자.

$$Z_t = Z_{t-1} + u_t$$

라고 할 때, $u_t$가 $MA(\infty)$ 모형을 따라

$$u_t = \Psi(L)\epsilon_t = \sum_{j=0}^{\infty} \psi_j \epsilon_{t-j}$$

라 하자. $\epsilon_t \sim W.N.(0,\sigma^2)$이며 $\sum_{j=0}^{\infty} j|\psi_j| < \infty$를 만족한다고 하자. 그렇다면 초기값이 $Z_0$인 시계열 $Z_t$는

$$\begin{aligned} Z_t &= Z_{t-1} + u_t \\ &= Z_{t-2} + u_{t-1} + u_t \\ &= \cdots \\ &= Z_0 + u_1 + u_2 + \cdots + u_{t-1} + u_t \\ &= \psi(1)(\epsilon_1 + \epsilon_2 + \cdots + \epsilon_t) + \eta_t - \eta_0 + Z_0 \end{aligned}$$

처럼 쓸 수 있다. 이때 $\psi(1) = \sum_{j=0}^{\infty} \psi_j$, $\eta_t = \sum_{j=0}^{\infty} \alpha_j \epsilon_{t-j}$, $\alpha_j = -\sum_{k=j+1}^{\infty} \psi_k$이다. 한편 조건에 의하여 $\sum_{j=0}^{\infty} |\alpha_j| < \infty$이기에 $\eta_t$는 잘 정의되는 정상시계열이다. ADF 검정에서는 $u_t$가 AR$(p-1)$을 따름을 가정하는데, 이를 MA$(\infty)$로 변환하면 BN 분해를 위한 가정을 잘 만족함이 알려져 있다. PP 검정은 적절한 수정항을 추가함으로써 AR$(p)$로부터 비롯되는 MA$(\infty)$ 모형이 아닌 일반적인 MA$(\infty)$ 모형을 따르는 $u_t$에 대해서도 단위근 검정을 수행할 수 있게 한다.
BN 분해 하에서 $\eta_t$는 정상성을 갖는 과정이고 $Z_0, \eta_0$는 고정된 값이다. 따라서 $\sqrt{T}\bar{u}$ 따위의 점근분포를 구하려 할 때, $\sqrt{T}\psi(1)\bar{\epsilon}$이 큰 도움이 된다. 자기상관성을 가지는 $u_t$들에 비해 $\epsilon_t$는 독립성을 가져 계산이 쉬우므로, 점근분포의 계산에 큰 역할을 한다. Hamilton 책의 506p를 보면 그 과정이 꽤 자세하게 등장한다.
과정은 생략하고, 아래의 값들을 정의하여 보자.

$\hat{u}_j$ : 적절한 모형 하에서 구한 잔차
$\hat{\sigma}^2_{\hat{\phi}}$ : 적절한 모형 하에서 구한 추정량의 분산
$s_T^2 = \dfrac{1}{T-2}\displaystyle\sum_{t=2}^{T} \hat{u}_t^2$ ($u_t$ 분산의 OLS 추정량)
$\hat{\gamma}_j = \dfrac{1}{T}\displaystyle\sum_{t=j+1}^{T} \hat{u}_t \hat{u}_{t-j}$
$\hat{\lambda}^2 = \hat{\gamma}_0 + 2\displaystyle\sum_{j=1}^{q} [1 - j/(q+1)]\hat{\gamma}_j$ ($q$시점까지 자기공분산함수가 유의해 보이는 경우의 Newey-West 추정량)
$\Delta_1 = \dfrac{1}{2}\left(\dfrac{T^2\hat{\sigma}^2_{\hat{\phi}}}{s_T^2}\right)(\hat{\lambda}^2 - \hat{\gamma}_0)$
$\Delta_2 = \dfrac{1}{2}\left(\dfrac{T\hat{\sigma}_{\hat{\phi}}}{s_T}\right)\dfrac{\hat{\lambda}^2 - \hat{\gamma}_0}{\hat{\lambda}}$

그렇다면, 조정된 통계량

$$Z_{\phi}^* = T(\hat{\phi} - 1) - \Delta_1$$

는 각 케이스에서 DF 검정에서의 $T(\hat{\phi} - 1)$와 동일한 점근분포를 따른다.
만약 $t$통계량을 이용하고자 한다면, 아래의 조정을 이용할 수 있다.

$$t_T^* = \left(\frac{\hat{\gamma}_0}{\hat{\lambda}^2}\right)^{1/2} t_T - \Delta_2$$

이 역시 각 케이스에서 DF 검정에서의 $t_T$와 동일한 점근분포를 따른다.
한편 마지막으로 $\hat{u}_t$를 구하는 방법에 대해 의문이 있을 수 있다. 각 케이스와 해당 경우의 $\hat{u}_t$를 알아 보자.

**Case 1**
실제 모형: $Z_t = Z_{t-1} + u_t,\ u_t = \Psi(L)\epsilon_t$
적합 모형: $Z_t = \phi Z_{t-1} + u_t$
잔차: $\hat{u}_t = Z_t - \hat{\phi} Z_{t-1}$

**Case 2**
실제 모형: $Z_t = Z_{t-1} + u_t,\ u_t = \Psi(L)\epsilon_t$
적합 모형: $Z_t = \delta + \phi Z_{t-1} + u_t$
잔차: $\hat{u}_t = Z_t - \hat{\delta} - \hat{\phi} Z_{t-1}$

**Case 3**
실제 모형: $Z_t = \delta + Z_{t-1} + u_t,\ u_t = \Psi(L)\epsilon_t$
적합 모형: $Z_t = \delta + \phi Z_{t-1} + \beta t + u_t$
잔차: $\hat{u}_t = Z_t - \hat{\delta} - \hat{\phi} Z_{t-1} - \hat{\beta} t$
즉 종합하면 단위근검정에서는 크게 두 가지를 고려해야 한다. 첫째. 어떤 검정을 사용할 것인가? 둘째. 그리고 어떤 모형을 가정할 것인가? 분석하고자 하는 시계열에서 오차항의 자기상관구조, 시계열의 평균과 추세를 고려하여 9개 케이스 중 하나를 잘 택한 뒤 검정해야만 할 것이다.

## 7.3 자기회귀누적이동평균과정

**Definition 14.** *만약 $d$차 차분된 $(1-B)^d$가 $ARMA(p,q)$ 모형을 따른다면, $Z_t$는* ***자기회귀누적이동평균과정****을 따른다고 하고, $ARIMA(p,d,q)$처럼 쓸 수 있다. 혹은*

$$\Phi(B)(1-B)^d Z_t = \Theta(B)\epsilon_t$$

*처럼 쓸 수도 있다. 만약 $d=1$이라면 $Z_t$는 정상적인 시계열 $W_t$의 누적합이 된다.*

시계열에 자기회귀누적이동평균과정을 적합하고자 한다면 시계열그림을 그린 뒤 추세를 확인하고, 확률적 추세가 있다고 판단되면 차분을 순차적으로 하며 정상시계열이 되는 시점까지 차분하여 $d$를 정한 뒤 ARMA 모형을 적합한다. 시계열도로 보았을 때 확인이 어렵다면, SACF도를 그려보는 것 역시 유용하다. 단위근이 존재하는 경우 SACF가 매우 느리게 감소하므로, SACF가 지수적으로 감소하지 않고 매우 느리게 감소하는 경우 차분을 하여 적합하는 것이 추천된다.
만약 원래의 시계열이 비정상적인 양상을 보이면 차분을 수행하는데, 실제로는 그 시계열이 정상시계열임에도 차분을 해버릴 수 있다. 이러한 상황을 **과대차분**이라 한다. 과대차분을 하면 정상성에는 문제가 없으나, ACF를 복잡하게 만들고 시계열의 분산을 증가시킨다는 점에서 선호되지 않는다. 따라서 시계열을 차분했을 때 오히려 분산이 증가하는 상황이 나오면, 과대차분을 의심하고 차분을 수행하지 않는 것이 권고된다.

# Chapter 8

# 자기회귀누적이동평균과정의 적합

앞서 본 자기회귀누적이동평균과정에서 적절한 차분의 차수 $d$를 결정한다면, 이후에는 $W_t = (1-B)^d Z_t$에 ARMA$(p,q)$ 모형을 적합하기만 하면 된다. 따라서 여기에서는 **Box-Jenkins** 방법에 기반하여 일반적인 정상시계열에 대하여 자기회귀이동평균과정을 적합하는 방법에 대해 알아본다.

## 8.1 모형의 식별

시계열이 적절히 차분되어 정상성을 가지게 되었다면, AR차수 $p$와 MA차수 $q$를 결정해 주어야 한다. 여기에는 SACF, SPACF, SIACF를 이용하는 방법과 정보기준을 이용하는 방법의 두 가지가 대표적이다.

### 8.1.1 SACF, SPACF, SIACF를 이용하는 방법

먼저 AR$(p)$ 과정을 적합하는 경우에는 SPACF를 사용할 수 있다. 이 과정의 경우 $k \geq p+1$인 경우 $\phi_{kk} = 0$이 됨이 알려져 있으므로, $\hat{\phi}_{kk}$를 구하고 그것이 0인지 여부를 판단함으로써 가설검정을 수행할 수 있다. 특히 점근적으로 그 평균이 0, 분산이 $\mathrm{Var}(\hat{\phi}_{kk}) = \frac{1}{n}$이므로

$$|\hat{\phi}_{kk}| > \frac{2}{\sqrt{n}}$$

이라면 유의수준 $\alpha = 0.05$에서 $\phi_{kk} = 0$을 기각하게 된다. SPACF 그림을 그렸을 때 특정한 점 이후에서 이것이 0에 매우 가깝다면, 이러한 검정을 통해 최대 시차 $p$를 결정할 수 있다.

MA$(q)$ 과정을 적합하는 경우에는 SACF를 주로 사용한다. SACF는 $k \geq q+1$인 경우에 $\rho_k = 0$이 되는데, 특히 이때

$$\mathrm{Var}(\hat{\rho}_k) = \frac{1}{n}(1 + 2\rho_1^2 + \cdots + 2\rho_q^2)$$

이다. 따라서 검정하고자 하는 시차 이전의 모든 $\rho_j$들로써 이 분산을 추정한 뒤 $H_0 : \rho_k = 0$을 검정하여 해당 시차 이전까지만 MA 모형이 분포함을 확인할 수 있다.

**Definition 15.** *정상 $ARMA(p,q)$ 과정*

$$\Phi(B) Z_t = \Theta(B) \epsilon_t$$

*에서* ***역과정***은 *$ARMA(q,p)$ 과정*

$$\Theta(B) W_t = \Phi(B) \epsilon_t$$

*에 의해 생성되는 $W_t$를 의미한다. 이 $W_t$의 ACF를* ***역자기상관함수*** *IACF라 부르며, 이는 PACF와 유사하게 기능하여 그 절단점으로써 AR 차수를 식별할 수 있다.*

### 8.1.2 모형선택 기준을 이용하는 방법

ARMA$(p,q)$ 과정의 경우 절단점이 없고 특정 점 이후로 서서히 감소하는 형태를 보이기에, SACF, SPACF, SIACF를 통한 모형식별이 어렵다. 따라서 이를 쉽게 하기 위하여, 다양한 **모형선택의 기준**, 혹은 **정보기준**이 개발되었다. 다양한 모형들에 대하여 각각을 계산하고, 해당 값이 가장 작은 모형을 선택하는 것이 고려될 수 있다.

**Definition 16.** *시계열분석에서,* ***AIC****,* ***SBC****는 아래와 같이 정의된다.*

$$\begin{aligned}\text{AIC} &= n\ln(\hat{\sigma}^2) + 2(p+q)\\ \text{SBC} &= n\ln(\hat{\sigma}^2) + \ln n \times (p+q)\end{aligned}$$

AIC와 SBC는 음의 로그가능도항(시계열분석의 경우 양의 로그오차항)에 모수의 개수에 적절한 가중치를 곱한 뒤 더한 값으로, 이 값은 모형의 가능도가 높을수록, 모형이 간단할수록 작아진다. 따라서 우리는 대개 이 값이 작은 모형을 선호하게 된다.

### 8.1.3 상수항 $\delta$의 포함 여부 결정

만약 시계열의 차분을 한 뒤 ARMA 모형을 적합하는 경우, 상수항의 포함 여부가 중요하다. 예를 들어 한 번 차분된 시계열에 상수항 $\delta$를 포함한다면, 이는 차분되기 전에는 $\delta t$였을 것이므로 결정적 선형추세가 있었음을 의미하게 되기 때문이다. 따라서 차분된 시계열 $W_t = (1-B)^d Z_t$의 평균

$$\bar{W} = \frac{1}{n}\sum_{t=1}^{n} W_t$$

을 통해 $\delta = 0$을 검정한다.

**Theorem 9.** $W_t = \mu + \sum_{j=0}^{\infty} a_j \epsilon_{t-j}$*로 표현되는 경우,*

$$\sqrt{n}(\bar{W} - \mu) \xrightarrow{d} N\left(0, \left(\sum_{j=0}^{\infty} a_j\right)^2 \sigma^2\right)$$

*이다.* $\sigma$*는* $\hat{\sigma}$*로 추정할 수 있으며,* $n$*이 충분히 클 경우 그로써 대체하여 검정하여도 무리가 없다.* $a_j$ *역시 적절히 모형을 식별한 뒤 정상성, 가역성 조건을 통해* $\hat{\Phi}, \hat{\Theta}$*로부터 추정할 수 있다.*

## 8.2 모수의 추정

차분의 수 $d$와 정상시계열에서의 모형의 차수 $p, q$, 그리고 상수항 $\delta$의 포함 여부까지 모두 선택했다면 $\Phi$와 $\Theta$를 추정해 주어야 한다. 이제 $p+q+2$개의 모수를 추정해야 하는 ARMA$(p,q)$ 모형

$$Z_t - \mu = \phi_1(Z_{t-1} - \mu) + \cdots + \phi_p(Z_{t-p} - \mu) + \epsilon_t - \theta_1\epsilon_{t-1} - \cdots - \theta_q\epsilon_{t-q}$$

을 고려하자. 추가적인 2개의 모수는 $\mu$와 $\sigma^2$이다.

### 8.2.1 적률추정법

**적률추정법**은 모집단의 적률을 상응하는 표본의 적률로 대체한 뒤 방정식을 풀어 추정량을 구하는 방법이다. 대표적으로 $\mathbb{E}[Z_t] = \mathbb{E}[\bar{Z}] = \mu$이므로, $\mu$의 적률추정량은

$$\hat{\mu} = \bar{Z} = \frac{1}{n}\sum_{t=1}^{n} Z_t$$

이다. 한편 적률추정량은 MA 모형이나 ARMA 모형에 대해서는 구하기 어렵다. 가장 간단한 AR($p$) 모형의 경우, Yule-Walker 방정식

$$\begin{pmatrix} \rho_0 & \rho_1 & \rho_2 & \cdots & \rho_{p-1} \\ \rho_1 & \rho_0 & \rho_1 & \cdots & \rho_{p-2} \\ \rho_2 & \rho_1 & \rho_0 & \cdots & \rho_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \cdots & \rho_0 \end{pmatrix} \begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \\ \phi_p \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_p \end{pmatrix}$$

에서 자기상관함수 $\rho_k$를 표본자기상관함수 $\hat{\rho}_k$로 대체하여

$$\begin{pmatrix} \hat{\rho}_0 & \hat{\rho}_1 & \hat{\rho}_2 & \cdots & \hat{\rho}_{p-1} \\ \hat{\rho}_1 & \hat{\rho}_0 & \hat{\rho}_1 & \cdots & \hat{\rho}_{p-2} \\ \hat{\rho}_2 & \hat{\rho}_1 & \hat{\rho}_0 & \cdots & \hat{\rho}_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \hat{\rho}_{p-1} & \hat{\rho}_{p-2} & \hat{\rho}_{p-3} & \cdots & \hat{\rho}_0 \end{pmatrix} \begin{pmatrix} \hat{\phi}_1 \\ \hat{\phi}_2 \\ \hat{\phi}_3 \\ \vdots \\ \hat{\phi}_p \end{pmatrix} = \begin{pmatrix} \hat{\rho}_1 \\ \hat{\rho}_2 \\ \hat{\rho}_3 \\ \vdots \\ \hat{\rho}_p \end{pmatrix}$$

를 얻고, 이들을 풀어 얻은 $\hat{\phi}_1, \hat{\phi}_2, \cdots, \hat{\phi}_p$를 추정값으로 사용한다. 즉

$$\hat{\phi} = \hat{P}^{-1}\hat{\rho}_p$$

로 표현하기도 한다.

마지막으로

$$\gamma_0 = \phi_1\gamma_1 + \cdots + \phi_p\gamma_p + \sigma^2$$

이므로, 이를 통해

$$\hat{\sigma}^2 = \hat{\gamma}_0(1 - \hat{\rho}_1\hat{\phi}_1 - \hat{\rho}_2\hat{\phi}_2 - \cdots - \hat{\rho}_p\hat{\phi}_p)$$

으로 백색잡음 분산의 적률추정량을 얻을 수 있다. MA 모형이나 ARMA 모형에서도 적절한 방식으로 자기상관함수, 부분자기상광함수와 계수들 사이의 식을 만든 뒤 이와 같은 방법으로 추정한다.

### 8.2.2 조건부 최소제곱추정법

오차의 제곱합이 가장 작게 되도록 하는 최소제곱법을 시계열분석에 적용하려 하면, $\sum_{t=1}^{n}\hat{\epsilon}_t^2$을 최소화하기 위하여

$$\hat{\epsilon}_1 = (Z_1 - \mu) - \phi_1(Z_0 - \mu) - \cdots - \phi_p(Z_{1-p} - \mu) + \theta_1\epsilon_0 + \cdots + \theta_q\epsilon_{1-q}$$

를 얻어야 하는데 이는 $Z$와 $\epsilon$의 이전 시점 정보를 필요로 한다. **조건부 최소제곱추정법**에서는 이전 시점의 정보들이 주어졌을 때의 조건부 오차제곱합을 최소로 하는 $\phi, \theta, \mu$를 찾고 그로써 $\sigma^2$을 추정하는 것을 목적으로 한다. 이전 정보들에 대해 관측되지 않은 $Z_t$의 값들로는 $\bar{Z}$를, 오차항 $\epsilon_t$들의 값들로는 0을 가정한다. 그 다음 이로써 얻은 조건부 오차제곱합을 오차항의 자유도(자료의 개수에서 모수의 개수 $p + q + 1$을 빼어 만듦)으로 나누어 $\hat{\sigma}^2$을 얻는다.

예를 들어 AR(1) 모형

$$(Z_t - \mu) = \phi(Z_{t-1} - \mu) + \epsilon_t$$

을 고려하고, $(Z_1, Z_2, \cdots, Z_n)$이 관측되었다고 하여 보자. 그렇다면 조건부 오차제곱합은

$$S*(\phi, \mu) = \sum_{t=2}^{n}\hat{\epsilon}_t^2 = \sum_{t=2}^{n}((Z_t - \mu) - \phi(Z_{t-1} - \mu))^2$$

이며

$$\frac{\partial S*}{\partial \mu} = -2\sum_{t=2}^{n}(1-\phi)((Z_t-\mu)-\phi(Z_{t-1}-\mu))$$
$$\frac{\partial S*}{\partial \phi} = -2\sum_{t=2}^{n}[(Z_t-\mu)-\phi(Z_{t-1}-\mu)](Z_{t-1}-\mu)$$

이다. 일계조건으로부터

$$\hat{\mu} = \frac{\sum_{t=2}^{n} Z_t - \phi\sum_{t=2}^{n} Z_{t-1}}{(n-1)(1-\phi)} \approx \bar{Z}$$
$$\hat{\phi} = \frac{\sum_{t=2}^{n}(Z_t-\mu)(Z_{t-1}-\mu)}{\sum_{t=2}^{n}(Z_{t-1}-\mu)^2} \approx \hat{\rho}_1$$

을 얻는다. 따라서 위와 같이 최소제곱법의 아이디어를 통해 추정량을 구할 수 있으며, 이는 $n$이 커질 때 적률추정량과 거의 유사하다. 마지막으로

$$\hat{\sigma}^2 = \frac{S*(\hat{\phi},\hat{\mu})}{(n-1)-(1+0+1)} = \frac{S*(\hat{\phi},\hat{\mu})}{n-3}$$

이 될 것이다.

### 8.2.3 비조건부 최소제곱추정법

비조건부 최소제곱추정법에서는 조건부 최소제곱추정법과 달리 이전 시점의 관측값들을 $\bar{Z}$으로 대체하는 대신, **후방예측**하여 얻은 값으로 대체한 뒤 제곱합을 최소화하는 모수를 구한다. 이는 초기 조건이 오차제곱합에 미치는 영향을 줄일 수 있어 조건부 최소제곱추정법보다 더 효율적이다.

**Definition 17.** *$ARMA(p,q)$ 모형은* ***전진 형태***

$$\Phi(B)(Z_t-\mu) = \Theta(B)\epsilon_t$$

*또는* ***후진 형태***

$$\Phi(F)(Z_t-\mu) = \Theta(F)a_t$$

*으로 표현될 수 있다. 이때 $F$는* ***전진작용소****이며, $a_t$는 분산이 $\epsilon_t$와 같은 백색잡음과정이다. 후진 형태를 이용하면 관측값을 통해 이전 시기의 관측값을 후방예측하는 데 사용할 수 있다.*

예를 들어, AR(1) 모형에서, $(Z_1,Z_2,Z_3,Z_4,Z_5) = (0.1,0.3,0.2,0.5,0.3)$으로 주어졌다고 하자. 편의상 $\mu=0$이고 $\phi=0.2$라고 가정하자. 그렇다면

$$\epsilon_t = Z_t - 0.2Z_{t-1}$$
$$Z_t = 0.2Z_{t+1} + a_t$$

가 성립하므로,

$$Z_0 = 0.2Z_1 + a_0 = 0.2\times 0.1 = 0.02$$

로 추정할 수 있고,

$$Z_{-1} = 0.2Z_0 + a_{-1} = 0.2\times 0.02 = 0.004$$

로 추정할 수 있다. 이처럼 이전 시점의 $Z_t$들을 추정하고,

$$\hat{\epsilon}_0 = Z_0 - 0.2Z_{-1} = 0.02 - 0.2\times 0.004 = 0.0192$$

등으로 얻을 수 있다.
비조건부 오차제곱합에서는

$$S(\phi, \theta, \mu) = \sum_{t=-\infty}^{n} \hat{\epsilon}_t^2 \approx \sum_{t=-Q}^{n} \hat{\epsilon}_t^2$$

을 최소화하는 모수들을 찾는다. $-Q$는 $\hat{\epsilon}_t$가 매우 작아지는 점으로 잡는다. 다양한 모수들에 대해 이를 계산하고, 최소화하는 모수를 찾음으로써 비조건부 최소제곱추정법을 수행한다. 마찬가지로 $\hat{\sigma}^2$은 오차제곱합을 자유도 $(n+1+Q)-(p+q+1)$으로 나누어 계산한다.

### 8.2.4 최대가능도추정법

최대가능도추정법은 오차의 분포가 정규분포를 따른다고 가정하고, 가능도함수를 만든 뒤 해당 가능도함수를 최대화하는 모수를 선택하는 방법을 의미한다. 예를 들어 AR(1) 모형

$$Z_t - \mu = \phi(Z_{t-1} - \mu) + \epsilon_t$$

를 적합하고자 한다면, 가능도함수는

$$f(Z_2, Z_3, \cdots, Z_n | Z_1) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n-1}{2}} \exp\left(-\frac{1}{2\sigma^2}\sum_{t=2}^{n}((Z_t - \mu) - \phi(Z_{t-1} - \mu))^2\right)$$

으로 주어진다. 여기에서 모수를 구하면 **조건부 최대가능도추정법**이라 한다.
한편 $Z_1$에 대한 확률밀도함수를 곱하여 **비조건부 최대가능도추정법** 역시 수행할 수 있다.

$$Z_t - \mu = (1 - \phi B)^{-1}\epsilon_t = \sum_{j=0}^{\infty} \phi^j \epsilon_{t-j}$$

이기에 $Z_t - \mu$의 분포가 $N(0, \frac{\sigma^2}{1-\phi^2})$임을 감안하면,

$$f(Z_1, Z_2, \cdots, Z_n) = \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{n}{2}} (1-\phi^2)^{\frac{1}{2}} \exp\left(-\frac{1}{2\sigma^2}S(\phi, \mu)\right)$$

이다. 이때 $S(\phi, \mu) = \sum_{t=2}^{n}((Z_t - \mu) - \phi(Z_{t-1} - \mu))^2 + (1-\phi^2)(Z_1 - \mu)^2$이다. 따라서 가능도함수 혹은 로그가능도함수를 미분하여 모수들을 찾으면 그것이 최대가능도추정량이 되며, $\sigma^2$에 대해서도 마찬가지로

$$\hat{\sigma}^2 = \frac{1}{n}S(\hat{\phi}, \hat{\mu})$$

를 최대가능도추정량으로 사용할 수 있다. 일반적인 ARMA 모형에서도 이를 사용할 수 있으나, 매우 복잡하다. 특히 어떤 추정 방법이든 복잡한 ARMA 모형에서는 비선형적인 추정을 해야 해 반복법 등 다양한 알고리즘이 필요하다.

### 8.2.5 추정량의 분포

최대가능도추정량 혹은 최소제곱추정량의 성질을 고려하면, AR($p$) 모형에서의 모수 $\phi = (\phi_1, \cdots, \phi_p)^T$의 점근적 분포는

$$\sqrt{n}(\hat{\phi} - \phi) \xrightarrow{d} N(0, \sigma^2\Gamma^{-1})$$

이다. 이때 $\Gamma$는 $p-1$차까지의 자기공분산함수들로 이루어진 행렬을 의미한다. 역과정을 이용하면 MA($q$) 모형에서도 이와 같은 추론이 가능하다.

## 8.3 모형의 진단

마지막으로는 모형을 진단한 후 합당하다고 평가되면 해당 모형을 최종모형으로 선택하고, 그렇지 않다면 다시 식별과 추정을 수행하여 더 좋은 모형을 찾을 수 있도록 하는 과정이 필요하다.

### 8.3.1 잔차분석

정상 ARMA$(p, q)$ 모형

$$\Phi(B)(Z_t - \mu) = \Theta(B)\epsilon_t$$

를 적합하였다면, $\epsilon_t$에 대한 백색잡음 가정이 성립하여야 한다. 따라서 잔차

$$e_t = Z_t - \hat{Z}_t = \hat{\Theta}(B)^{-1}\hat{\Phi}(B)(Z_t - \hat{\mu})$$

로부터 얻은 $e_t$들이 오차 $\epsilon_t$와 같은 성질을 따를 것으로 기대된다. 이와 같이 잔차의 독립성, 정규성 등을 판별하는 과정을 **잔차분석**이라 한다. 보통 잔차들의 그림을 그린 후 문제가 있다고 생각되는 부분에 대하여 검정한다.

먼저 정규성 검정을 위해서는 QQ-plot을 그려 시각적으로 확인하거나, 히스토그램을 그린 후 콜모고로프-스미르노프 검정을 하는 등의 방법이 있다. 시계열분석에서는 아래의 **자크-베라 검정** 역시 많이 이용한다.

$$JB = \frac{n}{6}\left(\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\hat{e}_t}{\hat{\sigma}}\right)^3\right)^2 + \frac{1}{24}\left(\frac{1}{n}\sum_{t=1}^{n}\left(\frac{\hat{e}_t}{\hat{\sigma}}\right)^4 - 3\right)^2$$

이는 첨도와 왜도를 이용하여 정규성을 검정하는 방법으로, JB값이 크면 클수록 정규분포로부터 멂을 알 수 있다.

백색잡음 가정에서 자기상관성이 존재하는지 확인하기 위해서는 잔차의 SACF인 **RSACF**나 잔차의 SPACF인 **RSACF**를 그려 볼 수 있다. 이들이 0에 가까운지 확인하여 백색잡음 가정을 확인한다. 이들은 백색잡음 가정에서 분산이 점근적으로 $\frac{1}{n}$이므로, 기각역을 $\frac{2}{\sqrt{n}}$으로 하여 검정한다. 단, AR(1) 모형에서처럼 시차가 작은 경우에는

$$\mathrm{Var}(\hat{\rho}_1) \approx \frac{|\phi|}{n}$$

으로, 모수에 의한 어느 정도의 변화가 존재할 수는 있다. 한편 전체 시차의 정보를 종합하여

$$H_0 : \rho_1 = \rho_2 = \cdots = \rho_K = 0$$

을 검정하기 위해 **포트맨토 검정**을 이용하기도 한다. 그 검정통계량은

$$Q = n(n+2)\sum_{k=1}^{K}\hat{\rho}_k^2/(n-k) \sim \chi^2(K - p - q)$$

이다. 이때 1은 평균에 대한 것으로, 평균이 없으면 $K - p - q$가 자유도이다. 이를 **룽-박스** 검정이라 하기도 하며, $K$는 유의해 보이는 $\rho_k$까지에 대해 적절히 선택한다. 상황에 따라서는 1차 자기상관관계의 검정을 위하여 **더빈-왓슨 검정**을 이용하기도 한다.

### 8.3.2 과대적합

**과대적합**이란 잠정모형에 모수를 추가하여 모형을 적합한 뒤, 시계열이 얼마나 더 잘 설명되는지를 확인하는 과정이다. 만약 과대적합 결과 모수추정량이 많이 달라진다면, 혹은 잔차들의 분산이 작아진다면, 과대적합된 모형을 새롭게 선택할 수 있다. 이때 **모수과잉**을 방지하기 위하여 AR 항과 MA 항을 동시에 추가하지는 않으며, 둘 중 하나를 추가하는 식으로 반복하여 진단한다.

# Chapter 9

# 예측

시계열자료의 생성모형을 분석하는 것은 결국 이후 시계열이 어떻게 흘러갈지 예측하는 것을 최종적인 목적으로 한다. 여기에서는 앞서 선택된 ARIMA 모형이 참모형이고 진단을 통과하여 해당 모형이 선택되었을 경우, 어떻게 이후 시점의 시계열을 예측하는지에 대해 다룬다.

## 9.1 최소 평균제곱오차 예측

최소 평균제곱오차를 통한 예측은 $Z_1, \cdots, Z_n$을 이용하여 $Z_{n+l}$의 평균제곱오차

$$\mathbb{E}[(Z_{n+l} - g(Z_1, \cdots, Z_n))^2]$$

을 최소화하는 추정량을 찾는 것을 목적으로 한다. 잘 알려져 있듯이, 이는

$$Z_n(l) = \mathbb{E}[Z_{n+l} | Z_1, Z_2, \cdots, Z_n]$$

이 된다. 또한 예측오차는

$$e_n(l) - Z_{n+l} - Z_n(l)$$

이다.

### 9.1.1 차분방정식을 이용하는 방법

$$Z_n(l-i) = \mathbb{E}[Z_{n+l-i} | Z_1, \cdots, Z_n] = \begin{cases} Z_n(l-i) & i \le i-1 \\ Z_{n+l-i} & i \ge i \end{cases}$$

$$\mathbb{E}[\epsilon_{n+l-i} | Z_1, \cdots, Z_n] = \begin{cases} 0 & i \le i-1 \\ \epsilon_{n+l-i} & i \ge i \end{cases}$$

이므로, ARIMA$(p, d, q)$ 모형

$$\Phi(B)(1-B)^d Z_t = \Theta(B)\epsilon_t$$

에서 **일반화된 자기회귀작용소** $\Psi(B) = \Phi(B)(1-B)^d$를 정의한다면

$$Z_{n+l} = \psi_1 Z_{n+l-1} + \cdots + \psi_{p+d} Z_{n+l-(p+d)} + \epsilon_{n+l} - \theta_1 \epsilon_{n+l-1} - \cdots - \theta_q \epsilon_{n+l-q}$$

이 될 것이고, 양변에 $Z_1, \cdots, Z_n$에 대한 조건부 평균을 세우면,

$$Z_n(l) = \psi_1 Z_n(l-1) + \cdots + \psi_{p+d} Z_n(l-p-d) - \theta_l \epsilon_n - \cdots - \theta_q \epsilon_{n+1-q}$$

으로 추정할 수 있고, $l = 1, 2, \cdots$에 대해 단계적으로 계산이 가능함을 알 수 있다. $\epsilon$은 잔차를 사용한다.

### 9.1.2 무한이동평균의 형태로 표현하는 방법

만약 시계열이

$$Z_n = \sum_{j=0}^{\infty} a_j \epsilon_{n-j}$$

처럼 무한이동평균으로 나타내어 진다고 하자. ARIMA 모형의 경우

$$\Phi(B)(1-B)^d A(B) = \Theta(B)$$

이게 하는 $A(B)$가 $\sum_{j=0}^{\infty} a_j B^j$와 같다. 그렇다면, $Z_{n+l}$은 오차를 이용하여 표현가능하며, 추정의 측면에서는

$$Z_n(l) = \sum_{j=l}^{\infty} a_j \epsilon_{n+l-j}$$

가 된다. 따라서 잔차 $\epsilon$들을 이용하여 $Z_n(l)$을 얻을 수 있고, 예측오차는

$$e_n(l) = \sum_{j=0}^{l-1} a_j \epsilon_{n+l-j}$$

으로 주어진다. 이때 장점은 예측오차의 성질을 알 수 있다는 것이다.

$$\mathbb{E}[e_n(l)] = 0$$
$$\mathrm{Var}(e_n(l)) = \left(\sum_{j=0}^{l-1} a_j^2\right) \sigma^2$$

또한 예측오차들 사이에 존재하는 상관성 역시 확인이 가능하다.

## 9.2 궁극적 예측함수

**Definition 18.** *$l$이 커질 경우 $Z_n(l)$이 어떻게 변할지를 나타내는 함수를* ***궁극적 예측함수****라고 한다. $l > q$ 인 경우*

$$Z_n(l) = \sum_{i=1}^{p+d} \psi_j Z_n(l-i)$$

*로 표현이 가능하므로, $l > q-(p+d)$인 경우 $Z_n(l)$은 $l$에 대한 선형차분방정식으로 표현이 가능하며, 그 $l \to \infty$일 때의 해가 궁극적 예측함수가 된다. 정상 확률과정인 경우 $Z_n(l)$의 형태는 $l$차 다항식, 지수함수, 삼각함수의 결합 형태로 주어진다.*

예를 들어 $(Z_t - \mu) = \phi(Z_{t-1} - \mu) + \epsilon_t$인 AR(1) 모형에서는,

$$Z_n(l) - \mu = \phi^l (Z_n - \mu)$$

가 됨이 잘 알려져 있다. 따라서

$$\lim_{l\to\infty} Z_n(l) = \lim_{l\to\infty}(\mu + \phi^l(Z_n - \mu)) = \mu$$

가 궁극적 예측함수가 될 것이다.

## 9.3 예측구간

앞서 무한이동평균의 형태로 예측을 수행할 때, 예측오차의 평균과 분산을 구하였다. 만약 오차가 정규분포를 따른다면,

$$-z_{\alpha/2} \le \frac{Z_{n+l} - Z_n(l)}{\left(\sum_{j=0}^{l-1} a_j^2\right)^{1/2}\sigma} \le z_{\alpha/2}$$

의 포함확률이 $1-\alpha$가 될 것이므로 $100(1-\alpha)$퍼센트 예측구간은

$$Z_n(l) \pm z_{\alpha/2} \times \sigma\sqrt{\sum_{j=0}^{l-1} a_j^2}$$

이 된다. 정상시계열인 경우, 루트 안의 값은 특정한 값으로 수렴하기에 예측구간의 너비 역시 수렴하게 된다.

## 9.4 예측의 갱신

새로운 시계열자료가 관측되면, 이를 이용하여 과거의 예측값을 갱신할 수 있다. $n+1$ 시점에서 $l-1$ 시차 이후의 예측값은 $Z_{n+l}$에 대한 예측값인데, 이는

$$Z_{n+1}(l-1) = \sum_{j=l-1}^{\infty} a_j\epsilon_{n+l-j} = \sum_{j=l}^{\infty} a_j\epsilon_{n+l-j} + a_{l-1}\epsilon_{n+1} = Z_n(l) + a_{l-1}(Z_{n+1} - Z_n(1))$$

으로 갱신될 수 있다. 즉 $n$시점에서 $n+1$시점으로 감에 따라, $Z_{n+l}$의 예측값은 $a_{l-1}$에 예측오차를 곱한 만큼을 더하여 보정된다.

전반적으로 예측은 모형의 종류에 따라 다르기 때문에, 일반론적인 이야기를 하기는 어렵다.

# Chapter 10

# 계절형 자기회귀누적이동평균모형

**Definition 19.** *$ARIMA(p,d,q)(P,D,Q)_s$ 모형은* ***승법계절 ARIMA 모형*** *혹은* ***SARIMA*** *모형이라는 이름으로 불리며,*

$$\phi(B)\Phi(B^s)(1-B^s)^D(1-B)^d Z_t = \delta + \theta(B)\Theta(B^s)\epsilon_t$$

*와 같은 형태를 가진다.*

**Theorem 10.** *승법계절모형에서 계절 시차를 기준으로 ACF들이 대칭인 형태를 보인다. 반면 이를*

$$\Phi^*(B)(1-B^s)^D(1-B)^d Z_t = \delta + \Theta^*(B)\epsilon_t$$

*처럼 비승법계절모형의 형태로 표현하는 경우, 그렇지 않다. 따라서 ACF를 이용하여 계절성이 승법적으로 반영되는지를 대략적으로 확인해볼 수 있다.*

계절형 자기회귀누적이동평균모형의 적합은 일반적인 ARIMA 모형의 적합과 유사하다. 먼저 SACF, SPACF 나 분야에 대한 지식을 바탕으로 주기 $s$를 고른다. 그 다음 $B$ 대신 $B^s$를 이용하여 ARIMA 모형을 적합한다. 그 다음 그렇게 만들어진 잔차 $u_t$에 대해 다시 ARIMA 모형을 적합한다. 그렇다면 승법성에 의하여 모든 모수들을 식별/추정할 수 있다.

# Chapter 11

# 전이함수모형

## 11.1 전이함수모형

여러 개의 입력시계열을 갖는 전이함수 모형(ARMAX 모형)의 형태는 아래와 같다.

$$Y_t = \sum_i v_i(L)X_{i,t} + n_t$$

여기에서 $X_{i,t}$는 $i$번째 입력시계열, $v_i(L) = \sum_{j=-\infty}^{\infty} v_{i,j}L^j$는 $j$번째 전이함수, $n_t$는 잡음과정으로 $n_t = \psi(L)\epsilon_t = \frac{\theta(L)}{\phi(L)}\epsilon_t$처럼 나타난다. 특히 $v_{i,j}$는 **충격반응가중값**, $\psi(L)$은 **선형필터**라 불리기도 한다.

그 적합을 위해서는 아래처럼 추정해야 할 모수의 개수를 제한시키는 방식이 많이 사용된다. 특히 입력시계열이 하나인 경우,

$$Y_t = \frac{w_s(L)}{\delta_r(L)}L^b X_t + n_t$$

를 적합하는 것을 고려할 수 있다. 이때

$$\begin{aligned} w_s(L) &= w_0 - w_1 L - \cdots - w_s L^s \\ \delta_r(L) &= 1 - \delta_1 L - \cdots - \delta_r L^r \end{aligned}$$

이며 $b$는 지연모수이다.

**Definition 20.** *만약 $\sum_{j=0}^{\infty} |v_j| < \infty$가 만족된다면, 전이함수가* ***안정적****이라 한다. 이는 입력시계열이 유한인 경우 출력시계열 또한 유한하게 됨을 의미한다.*

**Definition 21.** *$j < 0$일 때의 충격반응가중값이 0이면, 즉 $v_j = 0$이면 $Y_t = \sum_{j=0}^{\infty} v_j X_{t-j} + n_t$와 같이 쓸 수 있으며, 이때 전이함수를* ***인과적****이라 한다. 이는 입력시계열의 외생성을 나타낸다.*

## 11.2 교차상관함수

**Definition 22.** *정상인 두 시계열 $X_t$와 $Y_t$ 사이의* ***교차상관함수****는 다음과 같이 정의된다.*

$$\gamma_{XY}(h) := \mathrm{Cov}(X_t, Y_{t+h})$$

**Definition 23.** *정상인 두 시계열 $X_t$와 $Y_t$ 사이의* ***교차상관계수****는 다음과 같이 정의된다.*

$$\rho_{XY}(h) = \mathrm{Corr}(X_t, Y_{t+h}) = \frac{\gamma_{XY}(h)}{\sqrt{\gamma_X(0)\gamma_Y(0)}}$$

인과성을 만족하는 전이함수모형에서,

$$\begin{aligned}\gamma_{XY}(h) &= \mathrm{Cov}(X_t, Y_{t+h})\\ &= \mathrm{Cov}(X_t, \sum_{j=0}^{\infty} v_j X_{t+h-j} + n_{t+h})\\ &= \sum_{j=0}^{\infty} v_j \mathrm{Cov}(X_t, X_{t+h-j}) + \mathrm{Cov}(X_t, n_{t+h})\\ &= \sum_{j=0}^{\infty} v_j \gamma_X(h-j)\end{aligned}$$

만약 $X_t$가 백색잡음이라면, $h-j \neq 0$일 때 $\gamma_X(h-j) = 0$이므로,

$$\rho_{XY}(h) = \frac{\sigma_X}{\sigma_Y} v_h \rho_X(0)$$

이며 이를 정리하면

$$v_h = \frac{\sigma_Y}{\sigma_X} \rho_{XY}(h)$$

를 얻는다. 따라서 전이함수를 구할 때에는 입력시계열 $X_t$가 ARMA 모형

$$\phi_X(L)X_t = \theta_X(L)\alpha_t,\ \alpha_t \sim WN(0, \sigma_\alpha^2)$$

을 따른다고 한 다음 **사전백색화**된 백색잡음 $\alpha_t = \frac{\phi_X(L)}{\theta_X(L)} X_t$를 이용한다. 전이함수모형

$$Y_t = v(L)X_t + n_t$$

의 양변에 $\frac{\phi_X(L)}{\theta_X(L)}$을 가하여

$$\begin{aligned}\beta_t &= \frac{\phi_X(L)}{\theta_X(L)} Y_t\\ e_t &= \frac{\phi_X(L)}{\theta_X(L)} n_t\end{aligned}$$

로 놓으면 새 전이함수모형 $\beta_t = v(L)\alpha_t + e_t$에서 입력시계열 $\alpha_t$가 백색잡음이므로,

$$v_h = \frac{\sigma_\beta}{\sigma_\alpha} \rho_{\alpha\beta}(h)$$

로 구할 수 있게 된다.

## 11.3 전이함수모형의 적합

1. 전이함수모형에서 $X_t$와 $Y_t$는 모두 정상시계열을 따른다고 가정한다. 따라서 먼저 $X_t$와 $Y_t$에 대한 정상성 검정을 수행하고, 정상성이 만족되지 않는다면 차분이나 분산안정화변환 등을 수행한다.

2. 전이함수모형의 식별을 수행하기 위해

$$Y_t = v(L)X_t + n_t$$

를 세우고,

(a) 입력시계열 $X_t$에 ARMA 모형 $\phi_X(L)X_t = \theta_X(L)\alpha_t$를 적당한 시차를 선택하여 적합하여 사전백

색화된 시계열 $\alpha_t = \frac{\hat{\phi}_X(L)}{\hat{\theta}_X(L)} X_t$를 구한다.

(b) 사전백색화 필터 $\frac{\hat{\phi}_X(L)}{\hat{\theta}_X(L)}$를 적용하여 변환된 전이함수모형

$$\beta_t = v(L)\alpha_t + e_t$$

를 구한다.

(c) $\alpha_t$와 $\beta_t$ 사이의 표본교차상관계수 $\hat{\rho}_{\alpha\beta}(h)$를 구한 다음, 이를 이용하여

$$\hat{v}_h = \frac{\hat{\sigma}_\beta}{\hat{\sigma}_\alpha} \hat{\rho}_{\alpha\beta}(h)$$

를 구한다.

(d) 만약 $\alpha_t$와 $\beta_t$가 서로 상관이 없고 $\alpha_t$가 백색시계열이라면 $\text{Var}(\hat{\rho}_{\alpha\beta}(h)) \approx (T-h)^{-1}$임을 이용하여, 유의성 검정을 수행한다.

(e) 지연모수가 $b$일 때

$$\begin{aligned}
v_j &= 0 && j < b \\
v_j &= \delta_1 v_{j-1} + \cdots + \delta_r v_{j-r} + w_0 && j = b \\
v_j &= \delta_1 v_{j-1} + \cdots + \delta_r v_{j-r} - w_{j-b} && j = b+1, \cdots, b+s \\
v_j &= \delta_1 v_{j-1} + \cdots + \delta_r v_{j-r} && j > b+s
\end{aligned}$$

의 관계가 성립함을 이용하여, $r, s, b$를 잠정적으로 결정한 후 $\hat{\delta}_j, \hat{w}_k$를 구하고 이로써 초기 추정값 $\hat{v}(L) = \hat{w}_s(L)\hat{\delta}_r^{-1}(L)L^b$을 구한다.

3. 오차모형의 식별을 위하여

$$\hat{n}_t = Y_t - \hat{v}(L)X_t$$

를 구한 뒤, 여기에 적절한 차수의 ARMA 모형

$$\phi(B)n_t = \theta(B)\epsilon_t$$

을 적합시킨 뒤

$$Y_t = \frac{\hat{w}_s(L)}{\hat{\delta}_r(L)} X_{t-b} + \frac{\hat{\theta}(L)}{\hat{\phi}(L)} \epsilon_t$$

을 모형으로 사용한다.

4. 잠정적으로 추정된 모델

$$Y_t = \frac{w_s(L)}{\delta_r(L)} X_{t-b} + \frac{\theta(L)}{\phi(L)} \epsilon_t$$

에서 $\delta = (\delta_1, \cdots, \delta_r)^T, w = (w_0, w_1, \cdots, w_s)^T, \phi = (\phi_1, \cdots, \phi_p)^T, \theta = (\theta_1, \cdots, \theta_q)^T$과 $\sigma_\epsilon^2$을 조건부 가능도함수법, 조건부최소제곱법 등을 통하여 구하고, 최종적으로 계수들을 결정하여 전이함수모형을 적합한다.

## 11.4 전이함수모형의 진단

전이함수모형이 적합되면 잔차를 통해 진단을 수행한다.

1. $\epsilon_t$의 백색잡음 여부 판정을 위하여 자기상관을 검토할 수 있다. 적합된 잔차 $\hat{\epsilon}_t$로부터 그 표본자기상관

계수 $\hat{\rho}_\epsilon$을 구한 다음, 포트맨토 검정의 Q검정통계량

$$Q = T(T+2)\sum_{j=1}^{K}\frac{1}{T-j}\hat{\rho}_\epsilon^2(j) \sim \chi^2(T-p-q)$$

을 통하여 검정한다.

2. 오차와 입력시계열 사이의 독립 여부를 판정하기 위해 교차상관을 검토할 수도 있다.

$$\hat{\rho}_{\alpha\epsilon}(h) = \frac{\sum_{t=h+1}^{T_{\max}}\alpha_{t-h}\hat{\epsilon}_t}{\left(\sum_{t=1}^{T_{\max}}\alpha_t^2\sum_{t=1}^{T_{\max}}\hat{\epsilon}_t^2\right)^{1/2}}$$

을 $T_{\max} = T - p - \max(r, s+b)$에 대해 계산한 뒤, 그 분산이 근사적으로 $(T-h)^{-1}$임을 이용해 검정하거나 포트맨토검정의 검정통계량

$$Q = T_{\max}(T_{\max}+2)\sum_{j=0}^{K}\frac{1}{T_{\max}-j}\hat{\rho}_{\alpha\epsilon}(j)^2 \sim \chi^2(K-(r+s))$$

을 이용해 검정할 수 있다. 자유도는 $K+1$에서 전이함수측에 해당하는 모수의 개수 $r+s+1$을 감하여 $K-(r+s)$로 나타난다.

## 11.5 전이함수모형의 예측

전이함수모형에서

$$Y_t = \frac{w_s(L)}{\delta_r(L)}L^b X_t + \frac{\theta(L)}{\phi(L)}\epsilon_t, \quad \phi_X(L)X_t = \theta_X(L)\alpha_t$$

임을 이용하여

$$u(L) = \frac{w_s(L)L^b\theta_X(L)}{\delta_r(L)\phi_X(L)}$$
$$\psi(L) = \frac{\theta(L)}{\phi(L)}$$

을 정의하면,

$$Y_t = u(L)\alpha_t + \psi(L)\epsilon_t = \sum_{j=0}^{\infty}u_j\alpha_{t-j} + \sum_{j=0}^{\infty}\psi_j\epsilon_{t-j}$$

이므로 $t$ 시점에서 $Y_{t+l}$의 최적예측값은

$$\hat{Y}_t(l) = \sum_{j=0}^{\infty}u_{l+j}\alpha_{t-j} + \sum_{j=0}^{\infty}\psi_{l+j}\epsilon_{t-j}$$

으로 주어지고 예측오차는

$$Y_{t+l} - \hat{Y}_t(l) = \sum_{j=0}^{l-1}u_j\alpha_{t+l-j} + \sum_{j=0}^{l-1}\psi_j\epsilon_{t+l-j}$$

이다. 이로부터 그 MSE, 혹은 분산은 아래처럼 나타난다.

$$V(l) = \sum_{j=0}^{l-1}(\sigma_\alpha^2 u_j^2 + \sigma_\epsilon^2\psi_j^2)$$

# Chapter 12

# 벡터시계열모형

## 12.1 VAR

VAR은 벡터자기회귀 모형으로, 단순히 단변량 자기회귀모형의 벡터 버전이라 보기는 어렵다. 이는 변수 간의 동적인 상호작용을 묘사한다. 특히 이는 구조방정식모형의 축소된 형태이기도 하다. 아래의 $n$차원 VAR 모형을 고려하여 보자.

$$X_t = \mu + \Phi_1 X_{t-1} + \cdots + \Phi_p X_{t-p} + \epsilon_t$$

여기에서 사용되는 가정은 아래와 같다.

1. $\epsilon_t = (\epsilon_{1t}, \epsilon_{2t}, \cdots, \epsilon_{nt})^T$는 벡터잡음과정으로 $\epsilon_t \sim (0, \Omega)$를 따른다.

2. 정상성 조건은 특성방정식 $\Phi(z) = |I - \Phi_1 z - \cdots - \Phi_p z^p| = 0$의 근 $z$가 모두 단위원 밖에 있는 것이다.

## 12.2 시차의 결정

아래의 가설을 생각하자.

$$H_0 : p = p_0, \quad \text{v.s.} \quad H_1 : p = p_1$$

그렇다면 $p_0 < p_1$이라 할 경우 이 귀무가설과 대립가설의 차이는

$$(\Phi_{p_0+1})_{ij} = \cdots = (\Phi_{p_1})_{ij} \ \forall (i,j) \in \{1, 2, \cdots, n\}^2$$

이 제약조건으로 주어지는지 여부에 달려 있다. 이 제약에서는 총 $n^2(p_1 - p_0)$개의 변수를 조정한다.

그렇다면 가능도비 검정에서의 검정통계량은

$$LR = T(\log|\hat{\Omega}_0| - \log|\hat{\Omega}_1|) \ \dot{\sim} \ \chi^2(n^2(p_1 - p_0))$$

으로 주어진다. 이때 $\hat{\Omega}_0, \hat{\Omega}_1$은 각각 귀무가설과 대립가설에서의 잔차를 이용해 계산한 $\Omega$의 추정량이다. 조신섭 책에서는 $p_1 = p_0 + 1$일 때에 맞추어

$$M(p_0) = -(T - n - p_0 - 2.5)\log\left(\frac{|\hat{\Omega}_1|}{|\hat{\Omega}_0|}\right) \ \dot{\sim} \ \chi^2(n^2)$$

을 제시하고 있는데, 어차피 점근적인 검정이고 $T$가 큰 상황이면 큰 차이가 나지는 않는다. 여기에서는 Hamilton 책의 표현을 따르기로 한다. 일반적으로는 bottom-up 방식으로 $p_1 = p_0 + 1$처럼 둔 뒤 $p_0 = 0$부터

시작하여 더이상 귀무가설을 기각하지 못할 때까지 이 검정을 수행하고 그 시점에서의 시차를 VAR 모형의 시차로 사용한다.

혹은 AIC나 SBC와 같은 정보량 기준을 사용하여줄 수도 있다. 주의할 점은 AIC나 SBC의 부호나 더해지는/곱해지는 상수가 책마다 살짝씩 다르다는 것이다.

$$\begin{aligned}\text{AIC}(p_0) &= T\log(|\hat{\Omega}_0|) + 2n^2 p_0 \\ \text{SBC}(p_0) &= T\log(|\hat{\Omega}_0|) + \ln(T)n^2 p_0\end{aligned}$$

이들이 작은 $p_0$를 선택하는 것 또한 하나의 방법이 될 수 있다. 혹은 시계열도나 표본부분자기회귀행렬의 형태를 봐도 된다.

**부분자기회귀 행렬**은 아래의 다변량 Yule-Walker 방정식으로부터 구해지는 행렬 $\Phi_{k,k}$를 의미한다.

$$\begin{pmatrix}\Gamma(0) & \Gamma^T(1) & \cdots & \Gamma^T(k-1) \\ \Gamma(1) & \Gamma(0) & \cdots & \Gamma^T(k-2) \\ \vdots & \vdots & \ddots & \vdots \\ \Gamma(k-1) & \Gamma(k-2) & \cdots & \Gamma(0)\end{pmatrix}\begin{pmatrix}\Phi_{k,1}^T \\ \Phi_{k,2}^T \\ \vdots \\ \Phi_{k,k}^T\end{pmatrix} = \begin{pmatrix}\Gamma(1) \\ \Gamma(2) \\ \vdots \\ \Gamma(k)\end{pmatrix}$$

이때 $\Gamma(k)$는 $k$차 교차공분산 행렬 $\text{Cov}(X_t, X_{t-k})$이다. $\Gamma(k) \neq \Gamma^T(k)$인 대신 $\Gamma(k) = \Gamma^T(-k)$인 점에 유의하라. 이렇게 구한 $\Phi_{k,k}$는 VAR 모형이 참일 때

$$\Phi_{k,k} = \begin{cases}\Phi_k & k \le p \\ O & k > p\end{cases}$$

임이 알려져 있으므로, $\Phi_{k,k}$가 특정 차수 이후 영행렬이 되는 모습이 보인다면 그 차수를 VAR 모형을 적합하기 위한 차수 $p$로 사용할 수 있다.

## 12.3 VAR 모형의 적합과 진단

시차를 결정하였다면, 주어진 자료를 해당 시차 $p$에 맞추어 VAR($p$) 모형을 적합할 수 있다. 단변량에서와 같이, 최소제곱추정량과 최대가능도추정량을 사용한다. 이때 초기값 $X_{-p+1}, \cdots, X_0$은 주어졌다고 가정하자. 즉 조건부최소제곱추정량 또는 조건부최대가능도추정량을 얻는다고 가정하자.

1. **최소제곱추정량**

   최소제곱추정량에서 추정해야 할 모수는 절편 $n$개와 계수행렬의 계수 $n^2p$개이다. 여기에서 풀고자 하는 문제는

   $$\begin{cases}\text{minimize} & \sum_{t=1}^T \|X_t - \Phi_1 X_{t-1} - \cdots - \Phi_p X_{t-p}\|_F^2\end{cases}$$

   으로 주어진다. 이를 풀어내기 위해서

   $$Z_t = \begin{pmatrix}1 \\ X_t \\ \vdots \\ X_{t-p+1}\end{pmatrix} \in \mathbb{R}^{np+1}$$

   $$Z = \begin{pmatrix}| & | & & | \\ Z_0 & Z_1 & \ldots & Z_{T-1} \\ | & | & & |\end{pmatrix} \in \mathbb{R}^{(np+1)\times T}$$

$$
\begin{aligned}
Y &= \begin{pmatrix} | & | & & | \\ X_1 & X_2 & \dots & X_T \\ | & | & & | \end{pmatrix} \in \mathbb{R}^{n\times T} \\
y &= \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_T \end{pmatrix} \in \mathbb{R}^{nT} \\
B &= \begin{pmatrix} \delta & \Phi_1 & \cdots & \Phi_p \end{pmatrix} \in \mathbb{R}^{n\times(np+1)} \\
\beta &= \begin{pmatrix} B_1 \\ B_2 \\ \vdots \\ B_{np+1} \end{pmatrix} \in \mathbb{R}^{n(np+1)}
\end{aligned}
$$

을 정의하면, 우리의 VAR 모형이

$$Y = BZ + E$$

혹은

$$y = (Z^T \otimes I_n)\beta + \epsilon$$

처럼 써질 수 있기에 최소제곱추정량이

$$
\begin{aligned}
\hat{\beta} &= ((ZZ^T)^{-1}Z \otimes I_n)y \\
\hat{\Omega}_\epsilon &= \frac{1}{T-(np+1)}(Y-\hat{B}Z)(Y-\hat{B}Z)^T
\end{aligned}
$$

으로 주어진다.

2. **최대가능도추정량** 정규성 가정 하에서, $B$와 $\Omega_\epsilon$에 대한 로그가능도함수는 아래와 같이 주어진다.

$$\ln L(B, \Omega_\epsilon) = -\frac{nT}{2}\ln(2\pi) - \frac{T}{2}\ln|\Omega_\epsilon| - \frac{1}{2}\mathrm{tr}((Y-BZ)^T\Omega_\epsilon^{-1}(Y-BZ))$$

따라서 미분을 통해 간단하게

$$
\begin{aligned}
\hat{B} &= YZ^T(ZZ^T)^{-1} \\
\hat{\Omega}_\epsilon &= \frac{1}{n}(Y-\hat{B}Z)(Y-\hat{B}Z)^T
\end{aligned}
$$

를 얻을 수 있다.

한편 좋은 조건 하에서, 최소제곱추정량과 최대가능도추정량은 모두 일치추정량이며, 점근적으로는

$$\sqrt{T}(\hat{\beta}-\beta) \xrightarrow{d} N_{n(np+1)}(0, (\Omega_\epsilon \otimes \mathbb{E}[Z_tZ_t^T]^{-1}))$$

이다.

이렇게 추정을 마무리하고 나면, 잔차를 통해 모형을 진단할 수 있다. 진단에서는 정보량 기준을 이용할 수도 있고, 각 일변량 시계열에 대한 잔차로써 일변량 시계열의 진단을 수행해도 된다. 그러나 전 시차에서 전 시계열의 잔차의 적합성을 보기 위해서는 **다변량 포트맨토검정**을 이용할 수 있다. 잔차의 교차상관계수를 $\rho(k)$처럼 쓸 때,

$$H_0 : \rho(1) = \cdots = \rho(K) = 0, \quad \text{v.s.} \quad H_1 : \exists k \text{ s.t. } \rho(k) \neq 0$$

을 검정하기 위한 포트맨토검정통계량은

$$Q_n(K) = T^2 \sum_{k=1}^{K} \frac{1}{n-k} \mathrm{tr}(\hat{\rho}(k)^T \hat{\rho}(0)^{-1} \hat{\rho}(k) \hat{\rho}(0)^{-1})$$

으로 주어지며 이는 점근적으로 자유도가 $n^2K - n^2p$인 카이제곱분포를 따른다. 이때 $K$는 잔차의 형태를 보고 선택해야 하는 값이다. 일반적으로 $\ln(T)$ 따위를 사용한다.

잔차진단을 통해 모형이 유효한 것으로 나타났다면, 그 이후 예측은 일반적인 일변량 시계열에서와 같다. 너무 간단하므로 여기에서는 다루지 않는다.

## 12.4 Granger-Causality

VAR에서 인과성이란 아래와 같이 정의된다. $X$가 $Y$를 **그레인저 인과**하지 않는다는 것은

$$\mathbb{E}[Y_{t+s}|Y_t, Y_{t-1}, \cdots, X_t, X_{t-1}, \cdots] = \mathbb{E}[Y_{t+s}|Y_t, Y_{t-1}, \cdots]$$

처럼 $X$의 현재/과거값이 $Y$의 미래값을 예측하는 데 어떠한 도움도 되지 않는다는 것을 의미한다. 이는 $Y$의 미래값에 대한 정보가 $X$의 현재/과거값에 포함되지 않는다는 것을 의미하기도 한다.

이변량 시계열에 대한 아래의 VAR(2) 모형을 고려해 보자.

$$\begin{pmatrix} x_t \\ y_t \end{pmatrix} = \begin{pmatrix} \phi_{11}(1) & \phi_{12}(1) \\ \phi_{21}(1) & \phi_{22}(1) \end{pmatrix} \begin{pmatrix} x_{t-1} \\ y_{t-1} \end{pmatrix} + \begin{pmatrix} \phi_{11}(2) & \phi_{12}(2) \\ \phi_{21}(2) & \phi_{22}(2) \end{pmatrix} \begin{pmatrix} x_{t-2} \\ y_{t-2} \end{pmatrix} + \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}$$

그렇다면

$$\mathbb{E}[x_t|x_{t-1}, x_{t-2}, y_{t-1}, y_{t-2}] = \phi_{11}(1)x_{t-1} + \phi_{12}(1)y_{t-1} + \phi_{11}(2)x_{t-2} + \phi_{12}(2)y_{t-2}$$

이므로 $y_t$가 $x_t$를 그레인저 인과하지 않을 필요충분조건은 $H_0 : \phi_{12}(1) = \phi_{12}(2) = 0$이다. 동일한 이유로, $x_t$가 $y_t$를 그레인저 인과하지 않을 필요충분조건은 $H_0 : \phi_{21}(1) = \phi_{21}(2) = 0$이 된다. 즉 귀무가설은 그레인저 인과관계가 없다는 쪽이다. 그 검정은 왈드검정을 이용한다. 예를 들어 $y_t$가 $x_t$를 그레인저 인과하는지 확인할 때,

$$x_t = \phi_{11}(1)x_{t-1} + \phi_{12}(1)y_{t-1} + \phi_{11}(2)x_{t-2} + \phi_{12}(2)y_{t-2} + \epsilon_{1t}$$

을 회귀분석하여 $\phi_{12}(1) = \phi_{12}(2) = 0$인지를 검정하는 것이다. 따라서 검정통계량은 회귀분석을 진행한 뒤

$$F = \frac{(\mathrm{SSE(RM)} - \mathrm{SSE(FM)})/2}{\mathrm{SSE(FM)}/(T-5)} \sim F(2, T-5)$$

임을 이용하여 검정할 수 있다. 더욱 일반적인 VAR($p$) 모형의 경우에는

$$F = \frac{(\mathrm{SSE(RM)} - \mathrm{SSE(FM)})/p}{\mathrm{SSE(FM)}/(T-2p-1)} \sim F(p, T-2p-1)$$

임을 이용하여 검정할 수 있다.

한편 이변량 모형을 넘어서, 더욱 일반적인 $n$차원 시계열 $X_t$를 $n_1$, $n_2$차원 $X_{1t}$와 $X_{2t}$차원 시계열로 나누어 보자. 그렇다면

$$\begin{pmatrix} X_{1,t} \\ X_{2,t} \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} \Phi_{11}(1) & \Phi_{12}(1) \\ \Phi_{21}(1) & \Phi_{22}(1) \end{pmatrix} \begin{pmatrix} X_{1,t-1} \\ X_{2,t-1} \end{pmatrix} + \cdots + \begin{pmatrix} \Phi_{11}(p) & \Phi_{12}(p) \\ \Phi_{21}(p) & \Phi_{22}(p) \end{pmatrix} \begin{pmatrix} X_{1,t-p} \\ X_{2,t-p} \end{pmatrix} + \begin{pmatrix} \epsilon_{1,t} \\ \epsilon_{2,t} \end{pmatrix}$$

에서 그레인저 인과의 개념을 확장해 **블럭 외생성**을 고려하자. $X_1$이 $X_2$에 외생적이라는 것은, $X_1$의 미래값

이 $X_2$의 과거나 현재값에 의해 영향을 받지 않는다는 것을 의미한다. 즉 $X_1$이 모형에서 외생변수로 작용함을 의미한다. 그렇다면 앞에서와 같이, $X_2$가 $X_1$을 그레인저 인과하지 않는다는 것은

$$H_0 : \Phi_{12}(1) = \Phi_{12}(2) = \cdots = \Phi_{12}(p) = O_{n_1 \times n_2}$$

를 검정하는 것과 동일하다. $\Phi_{12}(k)$는 $n_1 \times n_2$행렬이다. 따라서 여기에서의 왈드 검정은

$$F = \frac{(|\hat{\Omega}_{11}(0)| - |\hat{\Omega}_{11}(1)|)/pn_1n_2}{|\hat{\Omega}_{11}(1)|/(T - np - n_1)} \sim F(pn_1n_2, T - np - n_1)$$

으로 주어질 것이다. 한편 대표본 하에서

$$pn_1n_2F \dot{\sim} \chi^2(pn_1n_2)$$

이므로 이를 이용할 수도 있고, 가능도비검정에서처럼

$$LR = T(\log(|\hat{\Omega}_{11}(0)|) - \log(|\hat{\Omega}_{11}(1)|)) \dot{\sim} \chi^2(pn_1n_2)$$

임을 이용해도 된다. 주의할 것은 이러한 그레인저 인과검정은 $X_t$가 정상시계열일 때에만 유효하다.

## 12.5 충격반응함수

경제학에서는 오차항을 백색잡음으로 보기도 하지만 혁신으로 보기도 한다. 그렇다면 각 측면에서의 충격이 각종 거시경제변수에 어떠한 반응을 불러오는지 궁금할 수 있다. 이를 위해 충격반응함수를 확인하는 경우가 많다. 먼저 $X_t$가 VAR($p$) 모형을 따르는 정상시계열이라고 하자.

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \cdots + \Phi_p X_{t-p} + \epsilon_t$$

그렇다면 이를 VMA($\infty$) 모형

$$X_t = \epsilon_t + \Psi_1 \epsilon_{t-1} + \cdots$$

으로 변환시키면 과거의 충격 $\epsilon_{t-s}$가 현재의 거시경제변수 $X_t$에 미치는 영향을 확인할 수 있을 것이다. 그렇다면

$$(\Psi_s)_{ij} = \frac{\partial X_{t,i}}{\partial \epsilon_{t-s,j}}$$

으로, $\Psi_s$의 $i$행 $j$열 원소는 $X_t$의 $i$번째 변수 $X_{t,i}$에 $s$기 이전의 $j$번째 충격 $\epsilon_{t-s,j}$가 미치는 영향을 의미하게 된다.

그러나 각 변수에 존재하는 혁신은 여러 측면에서의 충격이 혼재해 만들어진 것일 수 있다. 예를 들어 가격변수(인플레율)을 거시경제변수 중 하나로 넣어 VAR을 수행하는 경우, 공급측 충격과 수요측 충격이 모두 인플레율에 혁신을 불러올 수 있다. 올바른 분석을 위해서는 이들 변수의 $\epsilon_t$를 근본적인, 그리고 서로 직교하는 충격 여러 개로 분해할 필요가 있다. 대표적으로

$$\epsilon_t = Au_t$$

처럼 쓸 수 있다. 이때 $u_t$는 직교화된 충격들이며, $A$는 이들이 $X_t$에 미치는 구조적 효과를 묘사하는 $n \times n$ 행렬이다. 우리가 $\mathrm{Var}(\epsilon_t) = \Omega$로 썼다면, $\Omega$는 대칭인 양정치행렬이기에 $\Omega = ADA^T$처럼 분해할 수 있다. 즉 충격의 분해는 먼저 $\epsilon_t$의 공분산 구조 $\Omega$를 안 다음, 직교대각화를 수행하고, $A$를 찾은 뒤, $\mathrm{Cov}(u_t) = D$로 직교화된 $u_t = A^T\epsilon_t$를 얻음으로써 이루어진다.

그러나 여기에서 문제는 직교대각화가 유일하지 않다는 데 있다. 따라서 많은 경우에는 모형에 기반하여 적절한 제약조건을 걺으로써 $A$를 특정한다. 대표적으로는 Cholesky 분해를 이용하는 방법이 있다. Cholesky

분해는 유일하므로, $PP^T = \Omega$인 하삼각행렬 $P$를 찾을 수 있다. 대각원소를 1로 조정하기 위하여 $P = AD^{1/2}$로 쓰면, $ADA^T = PP^T = \Omega$이며, $A$는 하삼각행렬이 된다. 그리고 유일하다. 이때에는 $\epsilon_t = Au_t$이므로, 충격 $\epsilon_t$가 특정한 종속구조를 가지게 되는다. $n = 3$일 때를 예로 들면,

$$
\begin{aligned}
\epsilon_{t,1} &= u_{t,1} \\
\epsilon_{t,2} &= a_{21}u_{t,1} + u_{t,2} \\
\epsilon_{t,3} &= a_{31}u_{t,1} + a_{32}u_{t,2} + u_{t,3}
\end{aligned}
$$

이다. 따라서 Cholesky 분해를 이용해 identification을 하는 경우에는, 다양한 충격에 의해 즉각적으로 영향을 받는 변수를 $X_t$에서 최대한 아래에 두어야 한다. 반대로 특정 한 충격에만 즉각적으로 반응하고 나머지에는 반응하지 않는 경우, 이를 가장 위에 두어야 한다. 이는 보통 모델 하에서 적절한 가정을 통해 해결한다. 일반적으로 연방이 결정하는 이자율 등의 경우 다양한 충격을 고려해 결정되므로 가장 아래에, 투자수요나 가격수준같이 시장 혁신에 따라 결정되는 변수들을 가장 위에 두고는 한다.

물론 순서가 다르면 Cholesky 분해를 사용하는 대신 그에 맞는 분해를 사용해도 되지만, 귀찮기 때문에 보통 이처럼 한다. Cholesky 분해를 사용하는 식별 방법을 **recursive structure**라 부른다. 추후 구조적 VAR을 다룰 때 더욱 다양한 구조화 방법에 대해 알아볼 것이다.

다시 충격반응함수로 돌아오면, 직교화된 충격(공급측 충격, 수요측 충격 등)이 각 거시경제변수에 어떤 영향을 미칠지 확인하고 싶을 수 있다. 그렇다면

$$
\frac{\partial X_{t,i}}{\partial u_{t-s,j}} = \frac{\partial X_{t,i}}{\partial A_{j\cdot}^{-1}\epsilon_{t-s}} = \sum_{k=1}^{n} \frac{\partial X_{t,i}}{\partial \epsilon_{t-s,k}} \times \frac{\partial \epsilon_{t-s,k}}{\partial A_{j\cdot}^{-1}\epsilon_{t-s}} = \sum_{k=1}^{n} (\Psi_s)_{ik} \times A_{kj} = (\Psi_s A)_{ij}
$$

를 얻는다. 더 나아가 표준오차를 1로 하는 충격(표준화된 충격) $v_t = D^{-1/2}u_t$에 대하여 확인하면,

$$
\frac{\partial X_{t,i}}{\partial v_{t-s,j}} = (\Psi_s)_{i\cdot} A_{\cdot j}\sqrt{D_{ii}} = (\Psi_s AD)_{ij}
$$

를 얻게 된다. 즉 $s$시점 과거에서 일어난 1-SD인 $j$번째 충격 $v_{t-s,j}$가 $X_t$의 $i$번째 변수 $X_{t,i}$에 미치는 영향은 위처럼 쓸 수 있는 것이다. 이를 $s = 1, 2, \cdots$,에 대해 수행한다면 우리가 흔히 IRF라 부르는 impulse-response function(curve)를 얻게 되는 것이다.

## 12.6 국소충격반응

충격반응함수는 전체 피리오드에 VAR을 적합한 뒤 일반적인 충격(표준화된 충격)에 대한 반응함수를 분석한다. 이는 일반적이고 평균적인 이야기이다. 그러나 우리는 이따금 특정 충격이 미친 영향 자체가 궁금할 수 있다. 곡물의 공급측 충격이 가격수준에 미치는 영향 자체를 분석하는 것은 충격반응함수로 가능하지만, 러-우 전쟁으로 인한 공급측 충격이 가격에 어떠한 영향을 미쳤는가?에 대한 답은 할 수 없다. 이때 필요한 것이 국소충격반응이다.

먼저 VAR을 적합하면, $\hat{u}_t = \hat{A}^{-1}\hat{\epsilon}_t$를 얻을 수 있을 것이다. 이때 $\hat{\epsilon}_t$는 잔차로,

$$
\hat{\epsilon}_t = X_t - \hat{\Phi}_1 X_{t-1} - \cdots - \hat{\Phi}_p X_{t-p}
$$

로써 얻는다. 그렇다면 VMA($\infty$) 표현을 수행하면

$$
X_{t+s} = \hat{\epsilon}_{t+s} + \hat{\Psi}_1\hat{\epsilon}_{t+s-1} + \cdots + \hat{\Psi}_{s-1}\hat{\epsilon}_{t+1} + \hat{\Psi}_s\hat{\epsilon}_t + \hat{\Psi}_{s+1}\hat{\epsilon}_{t-1} + \cdots
$$

일 것인데, 만약 이 충격이 없었더라면?

$$
X_{t+s}|_{u_t=0} = \hat{\epsilon}_{t+s} + \hat{\Psi}_1\hat{\epsilon}_{t+s-1} + \cdots + \hat{\Psi}_{s-1}\hat{\epsilon}_{t+1} + 0 + \hat{\Psi}_{s+1}\hat{\epsilon}_{t-1} + \cdots
$$

이었을 것이다. 따라서 $u_t$라는 충격의 기여는

$$X_{t+s} - X_{t+s}|_{u_t=0} = \hat{\Psi}_s \hat{\epsilon}_t = \hat{\Psi}_s \hat{A} \hat{u}_t$$

로 계산할 수 있다. 사실 이는 충격반응함수에서 1-SD를 가정함에 따라 $D^{1/2}$만을 곱했던 것과 달리, 표준편차가 $D^{1/2}$인 실제 충격 $\hat{u}_t$를 집어넣는 것이나 다름없다.

한편 이는 적당한 identification이 되었다는 가정 하에, 해당 시기에 발생한 충격 $\hat{u}_t$의 크기를 앎으로써 $s$기 후에 어떠한 반응이 발생하였는지를 추정한다는 점에서 특정한 사건이나 정책이 미친 영향을 평가하는 데 사용될 수 있다.

## 12.7 예측오차 분산분해

우리는 앞서 특정한 충격이 어떠한 변수에 미치는 영향을 보았다. 반대로 어떠한 변수의 변동에 가장 영향을 많이 미치는 충격이 무엇인지 궁금할 수도 있다. 이를 위해서는 예측오차를 구성하는 변동성분을 분해할 필요가 있다. $X_{t+s}$에 대한 $t$시점에서의 추정량은

$$\hat{X}_{t+s|t} = \mathbb{E}[X_{t+s}|I_t] = \Psi_s \epsilon_t + \Psi_{s+1} \epsilon_{t-1} + \cdots$$

으로 주어진다. 그렇다면 예측오차는

$$X_{t+s} - \hat{X}_{t+s|t} = \epsilon_{t+s} + \Psi_1 \epsilon_{t+s-1} + \cdots + \Psi_{s-1} \epsilon_{t+1}$$

으로 주어지며, 그 MSE는

$$\text{MSE} = \text{Var}(X_{t+s} - \hat{X}_{t+s|t}) = \Omega + \Psi_1 \Omega \Psi_1^T + \cdots + \Psi_{s-1} \Omega \Psi_{s-1}^T$$

으로 써줄 수 있다. 그런데 $\epsilon_t = A u_t$로 쓰고 $\text{Cov}(u_t) = D$라고 하면,

$$\Omega = ADA^T = A_{\cdot 1} D_{11} A_{\cdot 1}^T + \cdots + A_{\cdot n} D_{nn} A_{\cdot n}^T$$

이므로,

$$\begin{aligned}
\text{MSE} = & \, D_{11}(A_{\cdot 1} A_{\cdot 1}^T + \Psi_1 (A_{\cdot 1} A_{\cdot 1}^T \Psi_1^T + \cdots + \Psi_{s-1} (A_{\cdot 1} A_{\cdot 1}^T \Psi_{s-1}) \\
& + D_{22}(A_{\cdot 2} A_{\cdot 2}^T + \Psi_1 (A_{\cdot 2} A_{\cdot 2}^T \Psi_1^T + \cdots + \Psi_{s-1} (A_{\cdot 2} A_{\cdot 2}^T \Psi_{s-1}) \\
& \vdots \\
& + D_{nn}(A_{\cdot n} A_{\cdot n}^T + \Psi_1 (A_{\cdot n} A_{\cdot n}^T \Psi_1^T + \cdots + \Psi_{s-1} (A_{\cdot n} A_{\cdot n}^T \Psi_{s-1})
\end{aligned}$$

처럼 쓸 수 있고, 이 중

$$D_{ii}(A_{\cdot i} A_{\cdot i}^T + \Psi_1 (A_{\cdot i} A_{\cdot i}^T \Psi_1^T + \cdots + \Psi_{s-1} (A_{\cdot i} A_{\cdot i}^T \Psi_{s-1})$$

가 $i$번째 직교화된 충격 $u_{t,i}$가 $s$기 이후의 $X_t$를 예측하는 데 생긴 오차에 미친 기여를 뜻하게 된다. 많은 경우 한 변수에 대한 MSE를 구하고 1로 표준화한 뒤, 각 충격들이 미친 영향을 비율로써 표시하고는 한다.

## 12.8 국소적 예측오차 분산분해

국소충격반응에서와 마찬가지로 국소적인 분산분해를 하고플 수 있다. 앞에서와 같이

$$X_{t+s} = \hat{\epsilon}_{t+s} + \hat{\Psi}_1\hat{\epsilon}_{t+s-1} + \cdots + \hat{\Psi}_{s-1}\hat{\epsilon}_{t+1} + \hat{\Psi}_s\hat{\epsilon}_t + \hat{\Psi}_{s+1}\hat{\epsilon}_{t-1} + \cdots$$

로 쓴다면, 이번에는 $t$기 이후의 모든 혁신이 없다고 가정해 보자. 그러면

$$\hat{X}_{t+s|u_t,\cdots,u_{t+s}=0} = 0 + \hat{\Psi}_{s+1}\hat{\epsilon}_{t-1} + \cdots$$

이며, 예측오차는

$$X_{t+s} - \hat{X}_{t+s|u_t,\cdots,u_{t+s}=0} = \hat{\epsilon}_{t+s} + \hat{\Psi}_1\hat{\epsilon}_{t-1} + \cdots + \hat{\Psi}_s\hat{\epsilon}_t = \sum_{h=0}^{s}\hat{\Psi}_h\hat{A}\hat{u}_{t+s-h}$$

이는 $t$기부터 $t+s$기까지 발생한 국소적인 충격 $\hat{u}_{t+s-h}$들이 $X_{t+s}$의 변동에 미친 영향을 분해한다. 각각의 계수는 $\hat{\Psi}_h\hat{A}$로 주어진다. 마찬가지로 비율로써 각 충격이 특정 변수의 변동에 기여한 바를 알 수 있다.

## 12.9 구조적 VAR

앞의 모든 논의에서 볼 수 있듯, $A$의 역할이 매우 막중하다. 여기에서는 경제학적 모형에 기반하여 $A$를 식별하는 문제에 대해 다룰 것이다.

일반적으로 구조방정식 모형에서는 아래의 모형을 고려한다.

$$A_0X_t = G + A_1X_{t-1} + A_2X_{t-2} + \cdots + A_pX_{t-p} + Ce_t$$

여기에서 $e_t \sim (0, D)$으로, 서로 직교화된 원초적인 충격들을 의미한다. $A_0$는 $X_t$에 존재하는 변수들이 갖는 즉각적인 관계를 묘사하며, $C$는 직교화된 $e_t$가 각 거시경제변수 $X_t$의 충격에 어떻게 포함되는지를 묘사한다. 이제 양변에 $A_0^{-1}$을 곱하면

$$X_t = A_0^{-1}G + A_0^{-1}A_1X_{t-1} + \cdots + A_0^{-1}A_pX_{t-p} + A_0^{-1}Ce_t$$

를 얻고, $\mu = A_0^{-1}G$, $\Phi_i = A_0^{-1}A_i$, $\epsilon_t = A_0^{-1}Ce_t$로 정의하면 이는 우리가 그간 봐왔던 VAR 모형

$$X_t = \mu + \Phi_1X_{t-1} + \cdots + \Phi_pX_{t-p} + \epsilon_t$$

로 변환될 수 있다. 이제 문제는 VAR을 적합한 뒤 직교화된 $e_t$를 포함하는 SEM으로 돌아가기 위해 $\mu, \Phi_1, \cdots, \Phi_p, \epsilon_t$로부터 $G, A_0, \cdots, A_p, C, e_t$를 찾는 식별의 문제이다. 이는 $A_0$과 $C$에 적절한 조건을 가함으로써 이루어진다. 혹은 추가적으로 $A_1$이나 $G$에 제약이 주어질 수도 있다.

대표적인 제약의 예시 중 하나는 우리가 이미 보았다. 바로 Cholesky decomposition이다. Cholesky decomposition에서는 공분산행렬의 Cholesky factor $P$가 하삼각행렬이라는 제약을 걸었었다. 이제 여기에서는

$$P = \underbrace{A_0^{-1}C}_{A\text{ 의 역할}} \times D^{1/2} = A_0^{-1}CD^{1/2}$$

이 하삼각행렬이라는 제약이 생기게 된다. 이는 $A_0$와 $C$에 제약을 겂으로써 식별이 가능하게 된다. 단 앞서 언급하였듯 이론을 통해 변수 간에 recursive structure가 있다고 생각되는 경우에만 이러한 제약을 걸 수 있을 것이다.

더욱 일반적으로는 우리가 특정한 이론을 통해 이를 확인한 경우,

$$\begin{aligned}\Omega &= A_0^{-1}CDC^TA_0^{-T}\\ \mu &= A_0^{-1}G\\ \Phi_i &= A_0^{-1}A_i\end{aligned}$$

를 해당 제약조건 하에서 풂으로써 $A_0^{-1}C$를 특정해줄 수 있다.
예시를 보기 위해 IS-LM-PC 모형을 고려하자.

$$\begin{aligned} y &= \alpha + u^s - \sigma(i - \mathbb{E}[\Delta p_{+1}]) + u^{is} && \text{(IS equation)} \\ m - p &= \phi y - \lambda i + u^{md} && \text{(LM equation)} \\ \Delta m &= u^{ms} && \text{(money supply process)} \\ \Delta p &= \Delta p_{-1} + \beta(y - u^s) && \text{(Okun's Law + Phillips curve)} \end{aligned}$$

이때 $y$는 산출량의 로그값, $p$는 가격수준의 로그값, $m$은 통화량의 로그값이다. 더불어, $u^s, u^{ms}, u^{md}, u^{is}$는 각각 공급, 화폐공급, 화폐수요, IS곡선 측면에서의 충격을 의미하게 된다. 여기에 합리적 기대가 성립한다면, $\mathbb{E}[\Delta p_{+1}] = \Delta p$이다. 따라서 SEM 모형으로 이를 쓰면,

$$\begin{pmatrix} 1 & 0 & \sigma & -\sigma \\ -\phi & 1 & \lambda & -1 \\ 0 & 1 & 0 & 0 \\ -\beta & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_t \\ m_t \\ i_t \\ p_t \end{pmatrix}$$
$$= \begin{pmatrix} \alpha \\ 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & -\sigma \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} y_{t-1} \\ m_{t-1} \\ i_{t-1} \\ p_{t-1} \end{pmatrix} + \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} y_{t-2} \\ m_{t-2} \\ i_{t-2} \\ p_{t-2} \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ -\beta & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} u_t^s \\ u_t^{md} \\ u_t^{ms} \\ u_t^{is} \end{pmatrix}$$

으로 쓸 수 있을 것이다. 이는 일반적인 $A_0$와 $C$를 모두 허용하는 것이 아니라, $A_0, G, A_1, A_2, C$의 구조를 허용하고 $\sigma, \phi, \lambda, \beta, \alpha$만 이동할 수 있게 함으로써 5개의 변수만 결정하면 SVAR 모형이 식별될 수 있게 한다. SVAR 모형은 경제학적 모형에 맞는 VAR 모형을 적합할 수 있게 한다는 점에서 장점이 있다. 단, 제약조건의 설정이나 변수설정 등에서 단순히 통계적으로만 해석하기는 어렵고, 도메인 널리지가 충분한 상태에서 정상시계열에 대한 분석이 필요하다.
한편 SVAR의 적합은 먼저 VAR 모형의 적합을 통해 $\hat{\mu}, \hat{\Phi}_i, \hat{\Omega}$ 등을 얻은 뒤 $A_0, C$ 등에 대한 제약조건 하에서 최대가능도법과 GMM 등을 통해 수행할 수 있다. 이는 특정한 $A_0$나 $C$ 구조 하에서만 잘 밝혀져 있고 그 외에는 직접 얻어야 하는데, 복잡한 경우가 다반사이므로 문제가 나올리가 없다고 생각한다. 따라서 여기에서는 생략한다.

# Chapter 13

# 공적분분석과 벡터오차수정모형

## 13.1 허구적 회귀

아래의 $y_{1,t}$와 $y_{2,t}$는 서로 독립인 확률보행과정을 따른다.

$$y_{1,t} = y_{1,t-1} + v_t,\ v_t \sim_{i.i.d.} (0,1)$$
$$y_{2,t} = y_{2,t-1} + u_t,\ u_t \sim_{i.i.d.} (0,1)$$

그렇다면 $y_{1,t}$와 $y_{2,t}$는 서로 독립적으로 생성되기에 관련이 없다. 그러나 단순회귀모형

$$y_{2,t} = \alpha + \beta y_{1,t} + \epsilon_t$$

를 적합하는 경우, 일반적으로 높은 $R^2$값과 회귀계수의 큰 $t$값, 그리고 낮은 DW통계량이 관측된다. 이는 두 시계열 간 아무 관계가 없음에도(참모형에서 $\alpha = \beta = 0$임에도) 회귀계수가 유의하게 나타난다는 점에서, **허구적 회귀**라 부른다.

더 나아가, 회귀계수로서 얻은 $\hat{\alpha}$와 $\hat{\beta}$에 대하여 아래가 성립한다.

$$\begin{pmatrix} T^{-1/2}\hat{\alpha} \\ \hat{\beta} \end{pmatrix} = O_p(1)$$

우리가 일반적으로 회귀분석을 진행하는 경우 $\sqrt{T}(\hat{\alpha} - \alpha)$ 따위로 검정을 진행함을 돌이켜 보면, 일반적인 회귀분석에서는

$$\begin{pmatrix} T^{1/2}\hat{\alpha} \\ T^{1/2}\hat{\beta} \end{pmatrix} = O_p(1)$$

여야 합니다. 즉 추정량이 $O_p(T^{-1/2})$로, $T \to 0$일 때 일치성을 가진다. 그러나 우리의 경우 이 값이 발산하여 일치성을 가지지 않는 추정량이 된다. 더 나아가, $F$값은 $F = O_p(T)$로 발산하여 회귀식 역시도 유의하게 관측된다. $R^2$이 0으로 수렴하지 않는 것과 DW 통계량이 매우 작게 나오는 것 역시 이와 관련된다. 결국 문제는 회귀식에서 사용하는 오차항 $\epsilon_t$가 $I(1)$임으로부터 기인한다.

허구적 회귀를 해결하기 위해서는 회귀식을

$$y_{1,t} = \alpha + \phi y_{1,t-1} + \beta y_{2,t} + \delta y_{2,t-1} + \epsilon_t$$

처럼 설정해 lagged variable을 넣음으로써 오차항 $\epsilon_t$를 $I(0)$으로 만드는 것이 하나의 대안이 될 수 있다. 이는 동태적 패널 모형 등에서 자주 사용되는 해결법 중 하나이다. 둘째는 차분을 수행하여

$$\Delta y_{1,t} = \alpha + \beta \Delta y_{2,t} + \epsilon_t$$

를 적합하는 것이다. 마찬가지로 차분을 하는 경우 $\Delta y_{1,t}$와 $\Delta y_{2,t}$는 $I(0)$이므로 유효한 회귀모형이 된다. 마지막으로는 이러한 자기상관관계를 고려하여 OLS 대신 적절한 GLS를 수행할 수 있다.

## 13.2 공적분

차분을 이용해 허구적 회귀를 해결하려는 경우, 모든 변수를 동일하게 차분한다. 이 경우 몇몇 변수가 과대차분됨에 따라, 시계열의 분산이 커지는 등의 문제가 발생할 수 있다. 만약 시계열에 **공적분 관계**가 존재하는 경우, 이를 보정함으로써 쉽게 문제를 해결할 수 있다.

만약 벡터확률과정 $y_t \sim I(1)$에 대하여 어떤 $\alpha \neq 0$가 존재해 $\gamma^T y_t \sim I(0)$이라면, $y_t$가 **공적분 벡터** $\gamma$로써 공적분되어있다고 말한다. 만약 그렇다면, 회귀분석의 과정에서 오차항 $\epsilon_t$가 $I(0)$일 수 있게 되므로 허구적 회귀의 문제가 발생하지 않는다.

예를 들어 $y_{1,t}$와 $y_{2,t}$가 $(1, -\beta)^T$라는 공적분 벡터로써 공적분되어있다고 하자. 이 경우

$$\begin{pmatrix} y_{1,t} \\ y_{2,t} \end{pmatrix} \sim C(1, -\beta)$$

처럼 쓰기도 한다. 그렇다면

$$y_{1,t} - \beta y_{2,t} = u_t \sim I(0)$$

이므로, 우리가

$$y_{1,t} = \alpha + \beta y_{2,t} + \epsilon_t$$

와 같은 단순회귀모형을 적합할 때 문제점이 없다. 특히 정상성을 가지는 $I(0)$인 $\epsilon_t = u_t - \alpha$ 과정은 오차항, 혹은 혁신으로 판단가능함을 고려한다면 장기적으로는 $y_{1,t} - \beta y_{2,t} = \alpha$에서 $\epsilon_t$만의 변동을 거친다. 따라서 공적분 관계는 장기평형의 관계로 해석할 수 있다. 또한 회귀분석으로써 얻는 $\alpha$와 $\beta$는 이 장기관계를 밝혀내는 데에 도움이 된다. 즉 공적분을 이용하여 회귀분석을 수행하는 것은 허구적 회귀의 위험으로부터 벗어나 모수들의 장기균형을 밝히는 데 도움이 된다.

## 13.3 공적분 관계 검정

어떠한 벡터 시계열 $y_t$에 대하여 공적분 벡터가 $\gamma$로 알려져 있다고 하자. 실제로는 아래의 과정을 거쳐 이를 검정할 수 있다.

1. 벡터 시계열을 이루는 $y_t$의 각 원소들 중 $\gamma$와 관련된 단변량 시계열들이 각각으로는 비정상 시계열임을 밝혀야 한다. 여기에는 적절한 모형 하에서의 DF 검정, ADF 검정, PP 검정 등이 사용될 수 있다.

2. 다음으로는 $\gamma^T y_t$가 정상성을 갖는지 검정하기 위하여, DF 검정, ADF 검정, PP 검정을 수행할 수 있다.

지난 시간에 다룬 것과 같이, $y_t$의 성질에 따라 적절한 **Case** 하에서 정상성 검정을 수행해 주어야만 한다. 특히 **Case 3**과 같은 경우는 추세를 허용하기 때문에 정확히 말하면 정상성을 가지지는 않는다. 그러나 추세가 선형이기 때문에, 그 계수를 안다면 추세를 제거할 경우 정상시계열을 얻게 된다. 그러므로 공적분은 엄밀하게는 $\gamma^T y_t \sim I(0)$인 경우를 의미하지만, 프랙티컬하게는 공적분 검정에서 장기균형에 추세가 있다고 보고 이를 제거한 뒤 공적분을 이용한 회귀분석을 수행하기도 한다.

## 13.4 VECM

추세가 포함된 벡터자기회귀모형

$$y_t = \mu_t + \Phi_1 y_{t-1} + \cdots + \Phi_p y_{t-p} + \epsilon_t,\ \epsilon_t \sim_{i.i.d.} (0, \sigma^2)$$

을 고려하자. 이때 이를 ADF representation의 형태로 다시 쓰면,

$$y_t = \mu_t + \rho y_{t-1} + \xi_1 \Delta y_{t-1} + \cdots + \xi_{p-1} \Delta y_{t-p+1} + \epsilon_t$$

이며

$$\rho = \sum_{i=1}^{p} \Phi_i, \quad \xi_i = -\sum_{i=j+1}^{p} \Phi_i$$

이다. 한편 양변에서 $y_{t-1}$를 빼면,

$$\Delta y_t = \mu_t + (\rho - I) y_{t-1} + \xi_1 \Delta y_{t-1} + \cdots + \xi_{p-1} \Delta y_{t-p+1} + \epsilon_t$$

를 얻는다. 한편, 우리의 벡터자기회귀모형이

$$y_t - \Phi_1 y_{t-1} - \cdots - \Phi_p y_{t-p} = \Phi(L) y_t = \mu_t + \epsilon_t$$

으로 써진다고 하면

$$\begin{aligned} \rho - I &= \sum_{i=1}^{p} \Phi_i - I \\ &= -(I - \Phi_1 - \cdots - \Phi_p) \\ &= -\Phi(1) \end{aligned}$$

이다.

한편 이 시계열에 $h$개의 공적분 관계가 있다는 것은, 랭크가 $h$인 $n \times h$ 행렬 $A$가 있어 $A^T y_t \sim I(0)$임을 의미한다. 좋은 조건 하에서, $h$개의 공적분 관계가 있는 시계열 $y_t$에 대한 $\Phi(1)$은

$$\Phi(1) = BA^T$$

로 분해될 수 있음이 알려져 있다. 이때 $A, B \in \mathbb{R}^{n \times h}$이며 $A$는 공적분 관계 행렬과 같다. 따라서 아래와 같이 **VECM**을 정리할 수 있다.

$$\Delta y_t = \mu_t + \xi_1 \Delta y_{t-1} + \cdots + \xi_{p-1} \Delta y_{t-p+1} - \underbrace{BA^T y_{t-1}}_{\text{error correction}} + \epsilon_t$$

- VECM은 단기적인 변동 $\Delta y_{t-1} \cdots \Delta y_{t-p+1}$들이 이번 기의 변동 $\Delta y_t$에 어떠한 영향을 미치는지를 묘사한다.
- 이때 $BA^T y_{t-1}$은 그 과정에서 오차를 수정하는 역할을 한다.
- $A$가 공적분 관계 행렬이므로, $A^T y_{t-1}$은 $I(0)$일 것으로 기대된다. 또한 다르게 말하면 그 값은 장기균형으로부터 벗어난 정도를 의미하기도 한다.
- 오차수정항의 계수가 마이너스인 이유도 이와 관련이 있다. 장기균형으로부터 시계열이 많이 벗어나는 경우 $BA^T y_{t-1}$ 값이 커진다. 이는 $\Delta y_t$의 과대차분을 보상해주는 항일뿐더러, 장기균형으로부터 벗어난 시계열을 장기균형으로 복원시키는 역할을 한다.

한편 $\mu_t$는 시계열의 형태를 보고 추세 모양으로부터 적절히 결정해 주어야 한다. 단위근검정에서와 마찬가지로 $\mu_t = 0$인 절편이 없는 경우, $\mu_t = \mu_0$인 절편이 있는 경우, $\mu_t = \mu_0 + \mu_1 t$인 선형 추세가 있는 경우 등이 자주 사용된다. 이외에도 다양한 방법이 있다.

## 13.5 Johansen's Method

VECM의 적합은 Johansen의 알고리즘을 통해 최대가능도에 기반하여 수행한다. 특히 공적분 관계의 개수 $h$의 결정과 VECM에서의 계수 결정이 동시에 수행된다. 대표적으로 추세가 있는 $\mu_t$를 이용하는 경우,

1. 보조회귀식
$$\Delta y_t = \eta_t + \Pi_1 \Delta y_{t-1} + \cdots + \Pi_{p-1}\Delta y_{t-p+1} + u_t$$
를 적합한다. (오차수정항이 없는 형태)

2. 보조회귀식
$$y_{t-1} = \nu_t + \Theta_1 \Delta y_{t-1} + \cdots + \Theta_{p-1}\Delta y_{t-p+1} + v_t$$
를 적합한다. (오차수정항이 없고 반응변수가 $y_{t-1}$)

3. 잔차 $\hat{u}_t$와 $\hat{v}_t$를 이용하여, 아래의 세 공분산행렬
$$\hat{\Sigma}_{UU} = \frac{1}{T-p}\sum_{t=p+1}^{T} \hat{u}_t\hat{u}_t^T, \hat{\Sigma}_{UV} = \frac{1}{T-p}\sum_{t=p+1}^{T} \hat{u}_t\hat{v}_t^T, \hat{\Sigma}_{VV} = \frac{1}{T-p}\sum_{t=p+1}^{T} \hat{v}_t\hat{v}_t^T$$
를 추정한다.

4. $\hat{\Sigma}_{VV}^{-1}\hat{\Sigma}_{VU}\hat{\Sigma}_{UU}^{-1}\hat{\Sigma}_{UV}^{-1}$의 고유값 $\hat{\lambda}_1 > \hat{\lambda}_2 > \cdots > \hat{\lambda}_m$과 상응하는 고유벡터 $e = (e_1, e_2, \cdots, e_n)$을 구한다. 단, $e^T\hat{\Sigma}_{VV}e = I$를 만족하도록 한다.

5. 이를 바탕으로 아래와 같이 모수들의 최대가능도추정량을 얻는다. 공적분 관계가 $h$개 있을 때,
$$\begin{aligned}
\hat{A} &= (e_1, e_2, \cdots, e_h) \in \mathbb{R}^{n\times h} \\
\hat{B} &= \hat{\Sigma}_{UV}\hat{A} \in \mathbb{R}^{n\times h} \\
\hat{\mu}_t &= \hat{\eta}_t - \hat{B}\hat{A}^T\hat{\nu}_t \\
\hat{\xi}_i &= \hat{\Pi}_i - \hat{B}\hat{A}^T\hat{\Theta}_i \\
\hat{\Sigma}_\epsilon &= \frac{1}{T-p}\sum_{t=p+1}^{T}(\hat{u}_t - \hat{B}\hat{A}^T\hat{v}_t)(\hat{u}_t - \hat{B}\hat{A}^T\hat{v}_t)^T
\end{aligned}$$

6. 이때의 로그가능도는
$$l^* = -\frac{Tn}{2}\log(2\pi) - \frac{Tn}{2} - \frac{T}{2}\log|\hat{\Sigma}_{UU}| - \frac{T}{2}\sum_{i=1}^{h}\log(1-\lambda_i)$$
으로 주어진다.

주의할 점은 $\mu_t$ 추세항의 형태에 따라 이 알고리즘은 달라질 수 있다. 위의 알고리즘은 선형 추세를 가정하는 경우의 알고리즘이다.

한편 이를 이용하면 $h$가 정해졌을 때의 로그가능도를 알 수 있다. 따라서 이는 공적분 개수 $h$에 대한 검정을 수행하는 데에도 사용할 수 있다. 먼저 아래의 가설을 고려하여 보자.

$$H_0 : h \text{ cointegrations v.s. } H_1 : \text{more than } h \text{ cointegrations}$$

그렇다면 이 경우의 로그가능도비는

$$LR = 2(l_1^* - l_0^*) = -T\sum_{i=h+1}^{n}\log(1-\lambda_i)$$

이며, $\mu_t$의 형태에 따라 그 분포가 달라진다. 이러한 검정법을 **트레이스 검정**이라고 부른다. 일반적으로 $\mu_t$의 형태에 따라 그 분포표가 주어지고, 이를 이용하여 결정해야 한다.
또 다른 방법은 가설을

$$H_0 : h \text{ cointegrations v.s. } H_1 : h+1 \text{ cointegrations}$$

으로 설정하는 것이다. 이 검정방법에서는 $h = n-1$이라는 큰 쪽부터 시작하여 귀무가설을 기각하지 못할 때마다 $h$를 점점 줄여나가 $h$를 찾는다. 이 경우 검정통계량은

$$LR = -T\log(1-\lambda_{h+1})$$

이며, 이러한 방법을 **최대고유값 검정**이라 부른다.

## 13.6 Summary

따라서 오차수정모형은 아래와 같이 적합한다.

1. AIC나 SBC와 같은 정보량 기준을 이용하여 VAR($p$)의 차수를 정한다.

2. 벡터 시계열에 대한 선행지식과 분석을 통하여 결정적 추세 $\mu_t$의 개형을 결정한다.

3. 적절한 $\mu_t$를 고른 뒤 그에 맞는 Johansen's method를 적용하고, 트레이스 검정 혹은 최대고유값 검정을 통해 $h$를 결정한다.

4. 해당 $h$에 맞게 Johansen's method를 이용해 오차수정모형

$$\Delta y_t = \mu_t + \xi_1 \Delta y_{t-1} + \cdots + \xi_{p-1}\Delta y_{t-p+1} - BA^T y_{t-1} + \epsilon_t$$

   를 적합한다. 이때 $A, B \in \mathbb{R}^{n\times h}$이다.

5. 적합한 오차수정모형의 잔차 $\hat{\epsilon}_t$를 이용하여 잔차검정을 수행해 모형의 타당성을 검토한다.

한편 그 결과에서 $A$를 통하여 $y_t$에 존재하는 장기적 균형관계, 혹은 해당 거시경제변수들이 갖는 관계를 확인할 수 있다. 또한 결정된 계수 $\mu_t$와 $\xi_i$ 등을 통하여, $\Delta y_t$가 단기적으로 이전 기의 $\Delta y_t$들과 $A^T y_{t-1}$에 의하여 어떤 영향을 받는지 확인할 수 있게 된다.

# Chapter 14

# 개입분석

시계열자료가 항상 하나의 모형에 의해 설명되는 것은 아니다. 시계열자료가 어떠한 사건이나 개입에 의하여 그 기저의 패턴이 달라지는 경우, 해당 효과를 추정하고 그 유의성 여부를 판단하는 **개입분석**을 수행할 수 있다. 한편 개입의 유무나 시점을 모르는 경우에는 시계열자료에서 **이상점**을 파악하는 것이 도움될 수 있다. 이상점을 판단하는 것은 안정적인 추론을 가능하게 할 뿐더러 그 원인을 고려해볼 수 있게 함으로써 분석에 도움을 준다.

## 14.1 개입분석

**Definition 24.** *정책의 변화와 같이 시계열자료에 영향을 미칠 가능성이 있는 사건들을* ***개입****이라 부르며, 이들을 입력변수로 사용하는 시계열모형을* ***개입모형****이라 한다. 개입변수는 사건이 지속되는 기간에 따라 두 종류로 나뉜다.*

- ***펄스함수****: 펄스함수 $P_t(T)$는 어떤 사건이 $T$ 시점에 발생하여 해당 시점에만 영향을 미칠 때 사용한다.*
- ***계단함수****: 계단함수 $S_t(T)$는 어떤 사건이 $T$ 시점에 발생하여 $T$ 시점 이후에 모두 영향을 미칠 때 사용한다.*

개입분석에서는 개입변수 $I_t$에 대하여 그 효과를

$$\frac{w(B)}{\delta(B)}B^b I_t$$

처럼 쓴다. 이때 $b$는 개입이 지연된 시차이며, $w_j$는 개입의 초기 기대효과를, $\delta_j$는 개입의 지속적인 효과를 나타낸다. $\delta(z) = 0$의 근들은 1 이상의 절대값을 가진다. 그 반응되는 형태는 다섯 가지로 구분할 수 있다.

- 개입의 효과가 $b$시차 이후에 반영되며, 그 크기가 $w$로 일정한 경우: $wB^bP_t(T)$ 혹은 $wB^bS_t(T)$가 개입의 효과가 된다.
- 개입의 효과가 $b$시차 이후에 반영되며, 그 크기가 점차로 줄어드는 경우: 초기에는 개입의 효과가 있으나 시장 참여자들의 행태 변화로 인하여 개입의 효과가 줄어드는 경우로, $\frac{w}{1-\delta B}B^bI_t$처럼 표현한다.
- 개입의 효과가 $b$시차 이후에 반영되며, 선형으로 증가하는 경우: 개입의 효과가 누적되어 선형추세를 가지는 상황으로, $\frac{w}{1-B}B^bS_t(T)$처럼 묘사된다.
- 반응변수에 반영되는 효과는 점차로 줄어드나 개입 발생 이전과 다른 수준으로 수렴하는 경우: 장기균형에 변화가 생기는 상황으로,

$$\left(\frac{w_0}{1-\delta B} + \frac{w_1}{1-B}\right)BP_t(T), \quad \left(w_0 + \frac{w_1 B}{1-\delta B} + \frac{w_2 B}{1-B}\right)P_t(T)$$

등이 대표적이다.

- 여러 개의 개입이 결합되는 경우: **다중개입모형**이라 부르기도 하며,

$$Z_t = \sum_{i=1}^{k} \frac{w_i(B)B^{b_i}}{\delta_i(B)} I_{i,t} + \frac{\theta(B)}{\phi(B)}\epsilon_t$$

개입모형을 적합하고자 하는 경우 위의 다중개입모형을 기본으로 하여 적절한 개입 $I_{i,t}$과 $k$, $w_i(B)$의 차수를 결정하고 적합한 뒤 오차항에 ARMA 구조를 적합하는 등의 방식으로 진행한다. 이후 모수들의 유의성을 판단하고 잔차분석으로써 진단하여 마무리한다.

## 14.2 이상점 탐지

### 14.2.1 시계열에서의 이상점과 그 종류

$X_t$는 원래 시계열로 오염되지 않은 시계열이고 ARMA$(p, q)$ 모형

$$\Phi(B)X_t = \Theta(B)\epsilon_t$$

를 따른다고 하자. 여기에 $wv(B)I_t(T)$만큼이 더해진 오염된 시계열

$$Z_t = X_t + wv(B)I_t(T)$$

를 고려하자. $w$는 이상점의 효과, $v(B)$는 이상점의 종류를 묘사하는 함수, $I_t(T)$는 $t = T$이면 1이고 아니면 0인 지시함수이다.

1. **Additive Outlier**(AO):

   AO는 $w_A$만큼의 이상점 효과가 $T$ 시점에서 오염되지 않은 시계열에 정해진 형태로, $Z_t = X_t + w_A I_t(T)$와 같이 표현될 수 있다. 이는 한 시점에서의 관측값에만 영향을 미치게 되므로 관측오차 등을 설명할 수 있다.

2. **Innovational Outlier**(IO):

   IO는 이상점이 발생한 이후 관측값에 지속적인 영향을 미치는 형태로,

   $$Z_t = \frac{\Theta(B)}{\Phi(B)}(\epsilon_t + w_I I_t(T))$$

   와 같이 표현된다. 즉 $\frac{\Theta(B)}{\Phi(B)}$의 가중치를 통해 $Z_t$에 이상점의 효과가 지속적으로 더해지는 상황을 의미한다. 이는 오차에 영향을 주는 내생적인 원인을 반영할 수 있다.

3. **Level Shift** (LS):

   특정 시점 이후의 모든 관측값 수준에 영구적인 변화가 발생하는 경우, 수준변화를 묘사하기 위하여

   $$Z_t = X_t + \frac{w_L}{1-B} I_t(T)$$

   처럼 나타낼 수 있다.

4. **Temporary Change** (TO):

   이상점에 의해 관측값 수준에 생긴 변화가 단기적으로 지속되며 그 크기도 점진적으로 줄어드는 경우를 묘사하기 위하여

   $$Z_t = X_t + \frac{w_C}{1-\delta B} I_t(T)$$

처럼 쓸 수 있다. 이때 $\delta$를 감쇠효과라 부르며, 지속 기간 역시 이에 의해 좌우된다. 만약 $\delta$가 0에 가깝다면 AO, 1에 가깝다면 LS가 이에 대응된다.

5. Patches of Outliers:

   AO가 연속적으로 발생하여 효과를 미치는 경우로,

$$Z_t = X_t + \sum_{j=1}^{m} w_j I_t(T_j)$$

   처럼 쓸 수 있다.

가장 일반적으로는,

$$Z_t = X_t + \sum_{j=1}^{m} w_j v_j(B) I_t(T_j)$$

처럼 모형을 만들 수 있다.

### 14.2.2 이상점 영향의 추정

이상점 탐지를 수행할 때에는 $X_t$의 모형은 이미 알고 있다고 가정한다. 즉 $\Phi(B)$와 $\Theta(B)$는 알고 있는 상황에서, 그 발생 시점과 종류를 밝혀내는 것이 이상점 탐지의 목적이다. 잔차는

$$\begin{aligned} e_t &= \frac{\Phi(B)}{\Theta(B)} Z_t \\ &= \frac{\Phi(B)}{\Theta(B)} (X_t + wv(B) I_t(T)) \\ &= \epsilon_t + w \frac{\Phi(B) v(B)}{\Theta(B)} I_t(I) \end{aligned}$$

이므로, 이상점의 종류에 따라 오차항에 붙는 항이 달라진다. $\Pi(B) = \frac{\Phi(B)}{\Theta(B)}$으로 정의하면, $Y_{T+k}$를

$$Y_{T+k} = \begin{cases} -\pi_k & AO \\ 0 & IO \\ 1 - \sum_{j=1}^{k} \pi_j & LS \\ \delta^k - \sum_{j=1}^{k-1} \delta^{k-j} \pi_j - \pi_k & TC \end{cases}$$

로 할 때 $e_t = wY_t + \epsilon_t$가 되는 것이다. 따라서 이들을 설명변수로 하고 $w$를 모수로 하는 회귀모형을 적합하면, 이상점의 종류와 위치를 알 때 각 이상점들의 효과와 그 분산을 추정할 수 있는 것이다.

**Theorem 11.** *AO의 경우*

$$\hat{w}_A = \frac{e_T - \sum_{j=1}^{n-T} \pi_j e_{T+j}}{\sum_{j=0}^{n-T} \pi_j^2}, \quad \mathrm{Var}(\hat{w}_A) = \frac{\sigma^2}{\sum_{j=0}^{n-T} \pi_j^2}$$

*이며, IO의 경우*

$$\hat{w}_I = e_T, \quad \mathrm{Var}(\hat{w}_I) = \sigma^2$$

*이다.*

### 14.2.3 이상점의 탐지

이상점을 안다는 가정 하에서는 위처럼 분석이 가능하다. 따라서 실제 분석 시에는 이상점이 어디서 나타나고 어떤 종류인지를 확인해 보아야만 한다. 대표적으로, 아래 가설을 생각하자.

$$H_0 : Z_T\text{는 이상점이 아니다}$$
$$H_1 : Z_T\text{는 AO이다},\ H_2 : Z_T\text{는 IO이다},\ H_3 : Z_T\text{는 LS이다},\ H_4 : Z_T\text{는 TC이다}$$

이를 확인하기 위해서는 $i = 1, 2, 3, 4$에 대하여 가능도비 검정통계량

$$\lambda_i = \frac{\hat{w}}{\sqrt{\text{Var}(\hat{w})}}$$

를 이용할 수 있다.

**Definition 25.** ***반복 추정 절차**는 아래의 과정으로 이루어진다.*

1. *먼저 이상점이 없다는 가정 하에서,*

$$\hat{e}_t = \hat{\Pi}(B)Z_t = \frac{\hat{\Phi}(B)}{\hat{\Theta}(B)}, \quad \hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n} \hat{e}_t^2$$

   *을 구한다.*

2. *이상점의 탐지*

   (a) *각 $i$와 시점 $t$에서 $\hat{\lambda}_{i,t}$를 구한다.*

   (b) *절대값이 가장 큰 검정통계량 $\hat{\lambda}_T = \max_{i,t} |\hat{\lambda}_{i,t}|$를 선택한다. 이 값을 기각역과 비교하여 더 크다면 이상점이 발생하였다고 판단한다.*

   (c) *그 경우 해당 시점 $T$에서 해당 이상점의 종류가 발생하였다고 생각하고,*

$$\begin{aligned}\tilde{Z}_t &= Z_t - \hat{w}\hat{v}(B)I_t(T)\\ \tilde{e}_t &= \hat{e}_t\hat{v}(B)\hat{\Pi}(B)I_t(T)\\ \tilde{\sigma}^2 &= \frac{1}{n}\sum_{t=1}^{n}\tilde{e}_t^2\end{aligned}$$

   *으로 조정한다. 이때 $\hat{v}(B)$는 대개 ARMA에서의 비처럼 $\frac{\hat{w}(B)}{\hat{\delta(B)}}$으로 추정한다.*

3. *앞서 수행한 과정을 모든 이상점이 탐지될 때까지 반복하여 수행한다.*

4. *$T_1, \cdots, T_k$ 시점에서 나타나는 이상점의 종류를 판단하고,*

$$Z_t = \sum_{j=1}^{k} w_j v_j(B) I_t(T_j) + \frac{\Theta(B)}{\Phi(B)}\epsilon_t$$

   *를 적합하여 이상점의 효과 $w_1, \cdots, w_k$와 모수를 모두 추정한다. 그 다음 새로운 잔차를 구하고, 그들로써 분산을 추정한다.*

5. *위의 네 과정을 다시 반복하여 추가적인 이상점이 없는 순간까지 수행한다.*

# Chapter 15

# 이분산시계열모형

앞선 모든 모형은 동일한 분산의 시계열을 가정하는 반면, 이분산시계열모형에서는 정상시계열이되 조건부 분산이 관측값에 의하여 변할 수 있게 함으로써 **변동성집중**과 같은 현상을 묘사할 수 있다.

## 15.1 ARCH 모형

**Definition 26.** *차수가 $q$인* ***자기회귀이분산*** *모형은 $ARCH(q)$처럼 쓰며,*

$$\begin{aligned} Z_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0,1) \\ \sigma_t^2 &= \zeta + \alpha_1 Z_{t-1}^2 + \cdots + \alpha_p Z_{t-q}^2 \end{aligned}$$

*에 의해 생성되는 $Z_t$를 의미한다.*

**Theorem 12.** *ARCH 시계열에 대하여 비조건부 분산은*

$$\sigma^2 = \frac{\zeta}{1 - \alpha_1 - \alpha_2 - \cdots - \alpha_q}$$

*으로 시간에 무관하며, 정상성이 만족될 조건은*

$$\zeta > 0, \quad \alpha_i \geq 0, \quad \alpha_1 + \alpha_2 + \cdots + \alpha_q < 1$$

*이 된다.*

ARCH 모형 하에서 $Z_t$의 조건부 분포를 구하면

$$f(z) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi}\sigma_t} \exp\left(-\frac{1}{2}\frac{Z_t^2}{\sigma_t^2}\right)$$

이며, 이를 $\sigma_t^2$에 대한 모형 하에서 최적화하는 모수들을 찾으면 적합을 수행할 수 있다. 또한 이렇게 찾은 최대가능도추정량은 일치추정량이며, 점근적으로 정규분포를 따름이 알려져 있다. 오차항 분포가 정규분포가 아니어도 quasi-MLE가 유효함이 알려져 있다. 혹은 $\nu_t = Z_t^2 - \sigma_t^2$이 martingale difference sequence임을 이용하여 $\nu_t$를 오차항으로 하는 최소제곱추정을 해도 된다.

검정하는 경우, 잔차 $\hat{\epsilon}_t$에 대해 AR($q$) 모형을 적합한 뒤, LM-ARCH 검정통계량 $nR^2 \sim \chi^2(q)$임을 이용하여 결정계수가 클 때 ARCH효과가 존재한다고 판단할 수 있다. 혹은 회귀모형에서의 F통계량을 이용하는 것 역시 가능하다. 이 경우에는 $qF \sim \chi^2(q)$임을 이용한다. 이외에도 $\epsilon_t^2$에 대해 륭-박스 검정을 하거나, 자크-베라 검정 등을 수행하여 모형을 따르는지 확인해줄 수 있다.

## 15.2 GARCH 모형

**Definition 27.** *$GARCH(p, q)$ 모형은 아래에 의해 생성되는 $Z_t$를 의미한다.*

$$\begin{aligned} Z_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0,1) \\ \sigma_t^2 &= \zeta + \alpha_1 Z_{t-1}^2 + \cdots + \alpha_q Z_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_p \sigma_{t-p}^2 \end{aligned}$$

*이때 $\zeta > 0, \alpha_i \geq 0, \beta_j \geq 0$이다. 이 모형이 정상시계열일 조건은 $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j < 1$이며, 비조건부 분산은 $\frac{\zeta}{1-\sum_{i=1}^q \alpha_i - \sum_{j=1}^p \beta_j}$이다.*

## 15.3 GARCH 모형의 변형 형태

### 15.3.1 IGARCH 모형

**Definition 28.** *만약 정상성 조건의 끝에 있어 $\sum_{i=1}^q \alpha_i + \sum_{j=1}^p \beta_j = 1$이라면* ***Integrated GARCH*** *모형이라 부른다. 이는 어느 한 시점에서의 영향력이 그 이후 지속적으로 영향을 미치는 경우로, $\sigma_{t+k}^2$을 예측하는데 $\sigma_t^2$가 지속적으로 필요함을 의미한다.*

### 15.3.2 threshold GARCH 모형

**Definition 29.** *$Z_t$가*

$$\begin{aligned} Z_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0,1) \\ \sigma_t^2 &= \zeta + \alpha_{11}(Z_{t-1}^+)^2 + \alpha_{12}(Z_{t-1}^-)^2 + \beta_1 \sigma_{t-1}^2 \\ \zeta &> 0, \alpha_{11}, \alpha_{12}, \beta_1 \geq 0 \end{aligned}$$

*을 만족하면,* ***threshold GARCH(1,1)*** *모형을 따른다고 말한다. 이는 주식시장에서의 레버리지 효과와 같은 비대칭성을 반영하기 위한 모형이다.*

### 15.3.3 ARMA-GARCH 모형

**Definition 30.** *$Z_t$가*

$$\begin{aligned} Z_t &= \phi_1 Z_{t-1} + \cdots + \phi_m Z_{t-m} + \eta_t - \theta_1 \eta_{t-1} - \cdots - \theta_n \eta_{t-n} \\ \eta_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0,1) \\ \sigma_t^2 &= \zeta + \alpha_1 Z_{t-1}^2 + \cdots + \alpha_q Z_{t-q}^2 + \beta_1 \sigma_{t-1}^2 + \cdots + \beta_p \sigma_{t-p}^2 \\ \zeta &> 0, \alpha_i \geq 0, \beta_j \geq 0 \end{aligned}$$

*이라면 이는 $ARMA(m, n) - GARCH(p, q)$ 모형을 따른다고 말한다. 이를 확장하면 threshold AR(1)-GARCH(1,1)과 같은 구성도 가능하게 된다.*

### 15.3.4 EGARCH 모형

**Definition 31.** *만약 $Z_t$가 아래처럼 결정된다면, EGARCH(1,1) 모형을 따른다고 말한다. 이 역시 비대칭성을 표현하기 위한 모형이다.*

$$\begin{aligned} Z_t &= \sigma_t \epsilon_t, \quad \epsilon_t \sim N(0,1) \\ \log(\sigma_t^2) &= \zeta + \theta(Z_{t-1}/\sigma_{t-1}) + \gamma(|Z_{t-1}/\sigma_{t-1}| - \mathbb{E}[|Z_{t-1}/\sigma_{t-1}|]) + \beta \log(\sigma_{t-1}^2) \end{aligned}$$

# Chapter 16

# 분위수회귀와 변화점 탐지

## 16.1 분위수회귀

$X_1, \cdots, X_n$이 $F_X$로부터의 랜덤표본이다. 우리가 궁금한 것은 VaR 등을 구하기 위하여, $\tau \in (0,1)$에 대해 퀀타일

$$q_X(\tau) = F_X^{-1}(\tau) = \inf\{x : F_X(x) \geq \tau\}$$

를 추정하는 것이다. 그런데

$$\rho_\tau(x, a) = (x-a)(\tau - I(x < a))$$

로 정의하면

$$\begin{aligned}
\mathbb{E}[\rho_\tau(X, a)] &= \mathbb{E}[(X-a)(\tau - I(X < a))] \\
\frac{\partial}{\partial a}\mathbb{E}[\rho_\tau(X, a)] &= \mathbb{E}[-(\tau - I(X < a)) - \delta_a(X-a)] \\
&= P(X < a) - \tau
\end{aligned}$$

이므로 일계조건에 의하여 $a = q_X(\tau)$가 이를 최소로 만든다. 따라서

$$q_X(\tau) = \operatorname{argmin}_a \mathbb{E}[\rho_\tau(X, a)]$$

이며 추정은

$$\hat{q}_X(\tau) = \operatorname{argmin}_a \frac{1}{n}\sum_{i=1}^n \rho_\tau(X_i, a)$$

로써 가능하다.

GARCH(1,1) 모형 하에서, $t-1$ 시점까지의 정보를 바탕으로 $Z_t$의 분위수를 예측하는 문제를 생각해 보자. 그렇다면

$$\begin{aligned}
\hat{q}_{Z_t|I_{t-1}} &= \operatorname{argmin}_a \mathbb{E}[\rho_\tau(Z_t, a)|I_{t-1}] \\
&= \sigma_t \hat{q}_{\epsilon_t|I_{t-1}}(\tau) \\
&= \sigma_t \times q_\epsilon(\tau) \\
&= \sqrt{\zeta + \alpha Z_{t-1}^2 + \beta\sigma_{t-1}^2} \times \xi := g_t(\alpha, \beta, \zeta, \xi)
\end{aligned}$$

처럼 퀀타일이 $\alpha, \beta, \zeta, \xi$의 함수로 주어짐을 안다. 따라서 이로부터

$$\text{argmin}\frac{1}{n}\sum_{t=1}^{n}\rho_\tau(Z_t - g_t(\alpha, \beta, \zeta, \xi))$$

을 찾아준 다음, 추정된 모수들을 통해 $g_t$를 계산하면 그것이 $I_{t-1}$ 하에서 $Z_t$의 $\tau$ 퀀타일 예측치가 된다. 이는 $\epsilon$에 정규분포 가정을 하지 않고 어떠한 분포든 분산이 1인 IID 분포기만 하면 적용 가능하다. 분포족을 안다면, $\xi$와 같은 값은 이미 아는 값이 될 것이다.

## 16.2 CUSUM 검정

레짐변화모형이나 지수평활법 등에서는 시계열의 평균이 시간에 따라 변할 수 있는 상황을 고려한다. 예를 들어,

$$X_t \sim N(\mu_t, \sigma^2)$$

으로 평균 수준이 국지적으로는 상수이면서 전반적으로는 변할 수 있는 시계열이 있다. 이 경우 관측된 영역 내에서 $\mu_t$의 변화가 있었는지 검정하고 싶을 수 있다. 이러한 **변화점 탐지**에서 사용할 수 있는 것이 CUSUM 검정이다.

$$\begin{aligned} H_0 &: X_1, \cdots, X_n \sim N(\mu_0, \sigma^2) \\ H_1 &: X_1, \cdots, X_k \sim N(\mu_0, \sigma^2) \\ &\quad X_{k+1}, \cdots, X_n \sim N(\mu_1, \sigma^2) \end{aligned}$$

와 같은 상황이 대표적이다. 이때 $k$는 알려지지 않은 상황이라 하자. 이 검정의 검정통계량은

$$\begin{aligned} T_n(s) &= \frac{1}{\sqrt{n}\sigma}\left|\sum_{i=1}^{[ns]}X_i - \left(\frac{[ns]}{n}\right)\sum_{i=1}^{n}X_i\right| \\ &= \frac{1}{\sqrt{n}\sigma}\left|\sum_{i=1}^{[ns]}(X_i - \mu_0) - \left(\frac{[ns]}{n}\right)\sum_{i=1}^{n}(X_i - \mu_0)\right| \\ &\approx W(s) - sW(1) = W^0(s) \\ T_n &= \sup_s |T_n(s)| \approx \sup_s |W^0(s)| \end{aligned}$$

이며, 그 분포가 잘 알려져 있기에 이 값이 크다면 귀무가설을 기각하고, 최대가 되는 $s$에서 $[ns]$를 변화점으로 삼는다.

동일한 원리로 분산 변화를 보고 싶으면

$$T_n = \max_s\left\{\frac{1}{\sqrt{n}\hat{\tau}_n}\left|\sum_{i=1}^{[ns]}X_i^2 - \left(\frac{[ns]}{n}\right)\sum_{i=1}^{n}X_i^2\right|\right\}$$

을 검정통계량으로, GARCH 모형에서의 모수 변화를 보고 싶으면

$$T_n = \max_s\left\{\frac{1}{\sqrt{n}\hat{\tau}_n}\left|\sum_{i=1}^{[ns]}\hat{\epsilon}_t^2 - \left(\frac{[ns]}{n}\right)\sum_{i=1}^{n}\hat{\epsilon}_t^2\right|\right\}$$

을 검정통계량으로 이용할 수 있다. 이때 $\hat{\tau}_n$은 표준편차의 추정량으로, 표본 4차적률에서 표본 2차적률의 제곱을 빼서 추정한다.