

제 10장 자기상관(autocorrelation)

1. 자기상관의 성격과 문제점

- 자기상관(autocorrelation) 의 성격

→ 고전적 회귀모형의 기본가정중에 오차항(ε_i)들은 서로 (1차함수적) 상관관계에 있지 않다($E[\varepsilon_i \varepsilon_j] = 0, i \neq j$)라고 하는 가정이 성립되지 않은 현상, 따라서 $E[\varepsilon_i \varepsilon_j] \neq 0 (i \neq j)$ 의 경우 자기상관(autocorrelation) 또는 계열상관(serial correlation) 이라 한다

→ 자기상관 현상은 연속적인 일련의 관측치들이 서로 상관되어 있을 때, 하나의 잔차항의 크기가 이웃하는 다른 잔차항의 크기와 서로 일정한 관련이 있을 때 일어난다.

→ 자기상관 현상은 주로 시계열 데이터에서 발생한다

- 자기상관 현상의 발생요인

a) 지속적인 습관을 갖는 행태(habit resistant behavior): 한 시점의 행태가 다음시점에서의 행태와 밀접한 관계를 갖는 경우

⇒ Duesenberry의 상대소득가설(relative income hypothesis)에 의하면, 소비자들의 지속적인 소비행태로 인해 소득이 줄어도 소비는 급격하게 줄지 않는다

b) 외부충격이나 정책의 변화로 인한 효과가 점차 감소하나 상당기간 지속되는 경우

c) 거미집 현상(cobweb phenomenon) 의 경우: 농산물의 공급변화가 시장가격에 일정한 시차를 가지며 조정되는 경우

d) 이전의 종속변수값을 설명변수로 활용하는 자기회귀 모형(autoregressive model)에서 시차(time lag)가 포함되는 경우

⇒ $C_t = \alpha + \beta_1 Y_t + \beta_2 C_{t-1} + \varepsilon_t$: 현재의 소비(C)가 이전의 소비(C_{t-1})에 의존한다

e) 시계열 데이터를 집계할 때 관측치의 평균값을 이용하는 경우

⇒ 일간자료 30 일간을 평균하여 월간 자료로 사용하는 경우

f) 모형설정오류에 의해 자기상관현상이 발생하는 경우가 있다

⇒ 부당하게 제외된 설명변수의 영향이 오차항에 남아 반영됨으로써 마치 오차항끼리 자기상관이 존재하는 것 처럼 보인다

⇒ 회귀함수를 잘못 설정하여 자기상관 현상이 나타난다: 비선형함수를 선형함수로 모형을 설정 한 경우

- 자기상관 현상에 의한 문제점

a) 자기상관이 발생할 경우 최소자승추정량(OLS estimator)은 더 이상 BLUE

가 되지 못한다: OLS 추정량은 불편성(unbiasedness)은 유지하나 효율성(efficiency)은 상실하게 된다

b) 자기상관현상을 무시한 OLS 추정량이 갖는 분산은 하향편의(downward bias)를 갖게 되어 추정량의 신뢰구간을 위축시키고 유의성 검증 통계량(t-통계량)이 부당하게 커짐에 따라 귀무가설($H_0: \beta = 0$)을 부당하게 기각시킬 가능성이 높다

c) 자기상관현상을 무시한 채 OLS 를 적용하여 오차항의 분산값을 산출한 경우 실제 분산값 보다 작게 나타나게 되어 결정계수인 R^2 값은 커지게 된다.

- 자기회귀 모형 (autoregressive model)

→ 자기상관을 나타내는 오차항의 구조로 1차자기회귀모형(first-order autoregressive model)이 가장 많이 사용된다.

→ $\varepsilon_t = \rho\varepsilon_{t-1} + \mu_t$ ($-1 < \rho < 1$, $\mu_t \sim N(0, \sigma_u^2)$)

⇒ $E[\varepsilon_t] = 0$

⇒ $\text{Var}(\varepsilon_t) = \sigma_\varepsilon^2$ (동분산 가정은 성립: homoskedacity)

⇒ $\text{Cov}(\varepsilon_t, \varepsilon_{t-k}) = \rho^k \sigma_u^2$ (자기공분산: autocovariance): $\rho \neq 0$ 인 이상 항상 자기상관을 갖는다.

2. 자기상관 여부의 검정

1) 잔차항의 그래프를 이용한 방법

- 실제 확인하고자 하는 것은 모회귀모형의 확률오차항 (ε_t)이 자기상관현상이 있는 것인지 여부이나 오차항(ε_t)을 실제로 관측 할 수 없으므로 가장 적절한 추정량인 OLS 추정량 e_t 를 활용한다
 - 회귀식에서 도출된 잔차항 (e_t)에 상관관계가 나타나는지를 그래프를 통해 살펴본다
- 서로 이웃하는 잔차항사이에 같이 움직이는 경향이 있으면 자기상관 현상이 있는 것으로 판단되며,
- 잔차항이 고르게 분포되어 있으면, 자기상관현상이 없는 것으로 판단된다.
- 그림10.1 (교과서 218)

2) Durbin-Watson 검정법(DW 검정법)

- 자기상관은 오차항 (ε_t)이 일차자기회귀 함수, $\varepsilon_t = \rho\varepsilon_{t-1} + \mu_t$ 의 형태를 갖는 것으로 가정하고 귀무가설 $H_0: \rho = 0$ 여부를 검정함으로써 자기상관현상의 여부를 판정하는 방법
- 검정 통계치 $d = \frac{\sum_{t=2,n} (e_t - e_{t-1})^2}{\sum_{t=1,n} (e_t)^2}$, e_t 는 원래의 모형을 이용하

여 OLS 에 의해 추정된 잔차항

→ 상관관계의 정도를 나타내는 ρ 값에 대한 추정치는,

$$\hat{\rho} = \sum_{t=2,n} (e_t e_{t-1}) / \sum_{t=2,n} (e_{t-1})^2,$$

→ $d \approx 2(1 - \hat{\rho})$

⇒ $d \rightarrow 2$ 이면 $\hat{\rho} \rightarrow 0$: 자기상관이 없다

⇒ $d \rightarrow 0$ 이면 $\hat{\rho} \rightarrow 1$: (+)의 자기상관

⇒ $d \rightarrow 4$ 이면 $\hat{\rho} \rightarrow -1$: (-)의 자기상관

→ d-검정통계량의 정의식에서 이용된 잔차항(e_{t-1})값이 설명변수 값에 의해 달라지기 때문에 그 분포는 모형에 포함된 설명변수에 따라 달라질 수 있다

⇒ d의 분포가 일정한 범위내(최저한계값 d_L 과 최고 한계값 d_U 사이)에서 만 변화한다

⇒ 이러한 분포를 이용하여 범위검정(bounds test)방법을 적용하여 자기상관 여부에 대한 가설검정을 시행한다.

a) $d < d_L$: (+)의 자기상관 존재

b) $d_U < d < (4-d_U)$: 자기상관이 없다는 귀무가설 채택

c) $(4-d_U) < d$: (-)의 자기상관 관계 존재

d) $d_L < d < d_U, (4-d_U) < d < (4-d_L)$: 검정결과 미정

⇒ d-통계치를 이용하기 위해서는 회귀모형에 상수항이 포함되어야

하고, 설명변수에는 종속변수의 과거값이 포함되지 않아야 한다.