⟨**권이태, 0809**⟩

공적분

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^2값과 회귀계수의 큰 t값, 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, F값은 $F = O_p(T)$로 발산하여 회귀식 역시도 유의하게 관측된다. R^2이 0으로 수렴하지 않는 것과 DW 통계량이 매우 작게 나오는 것 역시 이와 관련된다. 결국 문제는 회귀식에서 사용하는 오차항 ϵ_t가 $I(1)$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 ϵ_t를 $I(0)$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.

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_t가 **공적분 벡터** γ로써 공적분되어있다고 말한다. 만약 그렇다면, 회귀분석의 과정에서 오차항 ϵ_t가 $I(0)$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 ϵ_t만의 변동을 거친다. 따라서 공적분 관계는 장기평형의 관계로 해석할 수 있다. 또한 회귀분석으로써 얻는 α와 β는 이 장기관계를 밝혀내는 데에 도움이 된다. 즉 공적분을 이용하여 회귀분석을 수행하는 것은 허구적 회귀의 위험으로부터 벗어나 모수들의 장기균형을 밝히는 데 도움이 된다.

공적분 관계 검정

어떠한 벡터 시계열 y_t에 대하여 공적분 벡터가 γ로 알려져 있다고 하자. 실제로는 아래의 과정을 거쳐 이를 검정할 수 있다.

1. 벡터 시계열을 이루는 y_t의 각 원소들 중 γ와 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_t의 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.

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 h개의 공적분 관계가 있다는 것은, 랭크가 h인 $n \times h$ 행렬 A가 있어 $A^T y_t \sim I(0)$임을 의미한다. 좋은 조건 하에서, h개의 공적분 관계가 있는 시계열 y_t에 대한 $\Phi(1)$은

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

로 분해될 수 있음이 알려져 있다. 이때 $A, B \in \mathbb{R}^{n \times h}$이며 A는 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 Δy_t에 어떠한 영향을 미치는지를 묘사한다.
- 이때 $BA^T y_{t-1}$은 그 과정에서 오차를 수정하는 역할을 한다.
- A가 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 Δy_t의 과대차분을 보상해주는 항일뿐더러, 장기균형으로부터 벗어난 시계열을 장기균형으로 복원시키는 역할을 한다.

한편 μ_t는 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.

Johansen's Method

VECM의 적합은 Johansen의 알고리즘을 통해 최대가능도에 기반하여 수행한다. 특히 공적분 관계의 개수 h의 결정과 VECM에서의 계수 결정이 동시에 수행된다. 대표적으로 추세가 있는 μ_t를 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 h개 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 μ_t 추세항의 형태에 따라 이 알고리즘은 달라질 수 있다. 위의 알고리즘은 선형 추세를 가정하는 경우의 알고리즘이다.

한편 이를 이용하면 h가 정해졌을 때의 로그가능도를 알 수 있다. 따라서 이는 공적분 개수 h에 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, μ_t의 형태에 따라 그 분포가 달라진다. 이러한 검정법을 **트레이스 검정**이라고 부른다. 일반적으로 μ_t의 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 h를 점점 줄여나가 h를 찾는다. 이 경우 검정통계량은

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

이며, 이러한 방법을 **최대고유값 검정**이라 부른다.

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 μ_t의 개형을 결정한다.
3. 적절한 μ_t를 고른 뒤 그에 맞는 Johansen's method를 적용하고, 트레이스 검정 혹은 최대고유값 검정을 통해 h를 결정한다.
4. 해당 h에 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 A를 통하여 y_t에 존재하는 장기적 균형관계, 혹은 해당 거시경제변수들이 갖는 관계를 확인할 수 있다. 또한 결정된 계수 μ_t와 ξ_i 등을 통하여, Δy_t가 단기적으로 이전 기의 Δy_t들과 $A^T y_{t-1}$에 의하여 어떤 영향을 받는지 확인할 수 있게 된다.