시계열 분석 기법과 응용

여러 경제, 금융 시계열을 동시에 분석

Week 6. 벡터자기회귀모형

- 2가지 축 (분석) ① 정상적 시계열로 변환해 분석
② 비정상적 시계열 자체로 분석

6-1. 벡터자기회귀모형의 식별 및 추정

전치혁 교수
(포항공과대학교 산업경영공학과)

벡터자기회귀모형

- **벡터 시계열**
  - 여러 시계열을 동시에 고려하여 상호 연관성을 분석 ~벡터!
  - 예 (3개의 시계열) → 어떤 상호 관계를 맺는지 분석해 모형 구성

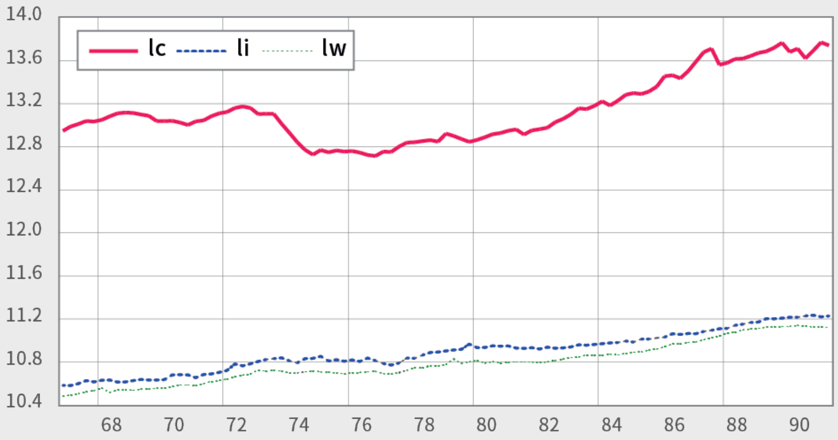
    시계열 1: Z_{1t} – 분기별 소비

    시계열 2: Z_{2t} – 분기별 소득

    시계열 3: Z_{2t} – 분기별 자산
  - 벡터시계열의 표현:

$$\boldsymbol{z}_t = \begin{pmatrix} Z_{1t} \\ Z_{2t} \\ Z_{3t} \end{pmatrix}, t = 1,2, \ldots$$

  - 벡터자기회귀 (vector autoregressive; VAR)모형이 주로 사용



벡터자기회귀모형

- 벡터자기회귀 (VAR) 모형

  - 한 시계열에 대한 AR(1) 모형

$$\underline{Z_t = \phi_1 Z_{t-1} + a_t}$$

  - 두 시계열에 대한 VAR(1) 모형 → 과거 시차 하나만 영향, ∴ VAR(1)

Z_1, Z_2의 과거값에 영향받음!

$$Z_{1t} = \phi_{11} Z_{1,t-1} + \phi_{12} Z_{2,t-1} + a_{1t}$$
$$Z_{2t} = \phi_{21} Z_{1,t-1} + \phi_{22} Z_{2,t-1} + a_{2t}$$

여기서 $a_{1t} \sim Nor(0, \sigma_1^2), a_{2t} \sim Nor(0, \sigma_2^2)$ 이며, $Cov[a_{1t}, a_{2t}] = \sigma_{12}$ ⇒ (벡터로 표현)

$$\boldsymbol{z}_t = \begin{pmatrix} Z_{1t} \\ Z_{2t} \end{pmatrix}, \boldsymbol{a}_t = \begin{pmatrix} a_{1t} \\ a_{2t} \end{pmatrix}, \Phi_1 = \begin{bmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{bmatrix}$$

white noise

계수 → 행렬표현

$$\underline{\boldsymbol{z}_t = \Phi_1 \boldsymbol{z}_{t-1} + \boldsymbol{a}_t}, \boldsymbol{a}_t \sim MVN(\boldsymbol{0}, \Sigma), \Sigma = \begin{bmatrix} \sigma_1^2 & \sigma_{12} \\ \sigma_{12} & \sigma_2^2 \end{bmatrix}$$

AR(1) 꼴
but Φ_1 : matrix
z_{t-1}, a_t : vector

multivariate normal
(vector이므로)

VAR 모형

- m차원 시계열의 VAR(1) 모형

시차1

$$\boldsymbol{z}_t = \begin{pmatrix} Z_{1t} \\ \vdots \\ Z_{mt} \end{pmatrix}; \boldsymbol{z}_t = \underline{\Phi_1 \boldsymbol{z}_{t-1} + \boldsymbol{a}_t}$$ (여기서 Φ_1는 mxm 행렬)

AR(1)꼴 but vector값

- m차원 시계열의 VAR(p) 모형

벡터, 행렬 이용

$$\boldsymbol{z}_t = \Phi_1 \boldsymbol{z}_{t-1} + \Phi_2 \boldsymbol{z}_{t-2} + \cdots + \Phi_p \boldsymbol{z}_{t-p} + \boldsymbol{a}_t$$

$$\Phi(x) = I - \Phi_1 x - \cdots - \Phi_p x^p \Rightarrow \Phi(B)\boldsymbol{z}_t = \boldsymbol{a}_t$$

VAR(1)형태의 표현 → VAR(p)보다 VAR(1)모형이 성질 찾기 쉬우므로 변환해줌

$$\boldsymbol{y}_t = \begin{pmatrix} \boldsymbol{z}_t \\ \boldsymbol{z}_{t-1} \\ \vdots \\ \boldsymbol{z}_{t-p+1} \end{pmatrix}, F = \begin{bmatrix} \Phi_1 & \Phi_2 & \cdots & \Phi_p \\ I & 0 & & 0 \\ & \ddots & & \\ 0 & & I & 0 \end{bmatrix}, \boldsymbol{v}_t = \begin{pmatrix} \boldsymbol{a}_t \\ 0 \\ \vdots \\ 0 \end{pmatrix} \Rightarrow \boldsymbol{y}_t = F\boldsymbol{y}_{t-1} + \boldsymbol{v}_t$$

↑ 새로운 행렬도입

VAR 모형

- 정상성 조건

VAR(1)모형 의 정상성 조건

- 모형: $\boldsymbol{z}_t = \Phi_1 \boldsymbol{z}_{t-1} + \boldsymbol{a}_t$
- 계수행렬 Φ_1의 고유치들의 크기가모두 1보다 작아야 한다.

eigenvalue

- $det\Phi(x) = \det(I - \Phi_1 x) = 0$ 의 근들의 크기가 모두 1보다 커야 한다.

VAR(p)모형 의 정상성 조건

- 모형: $\boldsymbol{z}_t = \Phi_1 \boldsymbol{z}_{t-1} + \Phi_2 \boldsymbol{z}_{t-2} + \cdots + \Phi_p \boldsymbol{z}_{t-p} + \boldsymbol{a}_t$

$$\Phi(B)\boldsymbol{z}_t = \boldsymbol{a}_t,\ \Phi(x) = I - \Phi_1 x - \cdots - \Phi_p x^p$$

- $det\Phi(x) = \det\left(I - \ \Phi_1 x - \cdots - \ \Phi_p x^p\right) = 0$ 의 근들의 크기가 모두 1보다 커야 한다.
- $\boldsymbol{y}_t = F\boldsymbol{y}_{t-1} + \boldsymbol{v}_t$의 형태에서 행렬 F의 고유치들의 크기가 모두 1보다 작아야 한다.

VAR 모형

- 정상성 조건

(예) 다음의 두 시계열에 대한 VAR(2)모형이 정상적인지 알아보자.

var(a_{1t})　cov(a_{1t}, a_{2t})

$$\begin{pmatrix} Z_{1t} \\ Z_{2t} \end{pmatrix} = \begin{bmatrix} 0.3 & -0.1 \\ -0.05 & 0.25 \end{bmatrix} \begin{pmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{pmatrix} + \begin{bmatrix} 0.2 & -0.75 \\ -0.1 & 0.4 \end{bmatrix} \begin{pmatrix} Z_{1,t-2} \\ Z_{2,t-2} \end{pmatrix} + \begin{pmatrix} a_{1t} \\ a_{2t} \end{pmatrix}, \ \Sigma = \begin{bmatrix} 0.6 & 0.05 \\ 0.05 & 1 \end{bmatrix}$$

var(a_{2t})

(풀이) $\boldsymbol{y}_t = F\boldsymbol{y}_{t-1} + \boldsymbol{v}_t$ 형태로 표시하면 행렬 F는 다음과 같다.

$$F = \begin{bmatrix} 0.3 & -0.1 & 0.2 & -0.75 \\ -0.05 & 0.25 & -0.1 & 0.4 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

- 고유치는 다음과 같다. $\lambda_1 = 0.9627, \lambda_2 = 0.2376, \lambda_3 = -0.0356, \lambda_4 = -0.6148$
- 즉, <u>고유치 크기가 모두 1보다 작으므로 위의 VAR(2)모형은 정상적이다.</u>

VAR 모형

- 모형의 식별 - 시차 p 결정

방법 1: 이론적 ACF/PACF를 구하고 표본 ACF/PACF와 비교 → VAR에서 구하기 어려움

- 교차상관관계 등이 존재하므로 쉽지 않음
- 개별 시계열에 대한 표본 ACF/PACF를 관찰하고 충분한 시차 p를 추정할 수도 있음

방법 2: 정보기준 (information criteria) 사용

- 실제적으로 널리 사용되고 있음 likelihood function, #(parameter)로 표현
- 여러 시차에 대해 AIC, BIC, HQ 등을 산출하고 정보기준값이 최소인 시차 선택
- $AIC(p) = ln|\hat{\Sigma}| + \frac{2m^2p}{N}$: Akaike information criteria
- $BIC(p) = ln|\hat{\Sigma}| + \frac{m^2 p lnN}{N}$: Schwarz (Bayesian) information criteria
- $HQ(p) = ln|\hat{\Sigma}| + \frac{2m^2 p ln lnN}{N}$: Hannan –Quinn information criteria

방법 3: 우도비 검정

VAR 모형의 식별

(예) 다음은 세가지 벡터 시계열에 대한 여러 시차에 따른 VAR모형을 추정한 후 정보기준값을 산출한 결과이다.

적절한 시차 p를 선택하라.

- SC 기준으로는 시차 1이 가정 적절하다고 볼 수 있으나 대부분 기준은 시차 2가 적절하다고 선언하므로 결과적으로 시차 2가 최적이라 하겠다.
- 기준마다 서로 다른 최적 시차를 제시하는 경우에는 가장 작은 시차를 최종적으로 선택하는 것도 방법이라 하겠다.

↳ 모형 단순하므로

log likelihood

Lag	LogL	LR	FPE	AIC	SC	HQ
0	-1415.48	NA	5.05e+11	35.46	35.55	35.49
1	-1070.11	656.19	1.13e+08	27.05	**27.41***	27.19
2	-1055.28	**27.07***	**9.74e+07***	**26.90***	27.53	**27.16***
3	-1046.40	15.53	9.79e+07	26.91	27.80	27.27
4	-1039.98	10.75	1.05e+08	26.97	28.14	27.44
5	-1038.50	2.37	1.28e+08	27.16	28.59	27.73

LR: sequential modified LR test statistic (each test at 5% level)
FPE: final prediction error

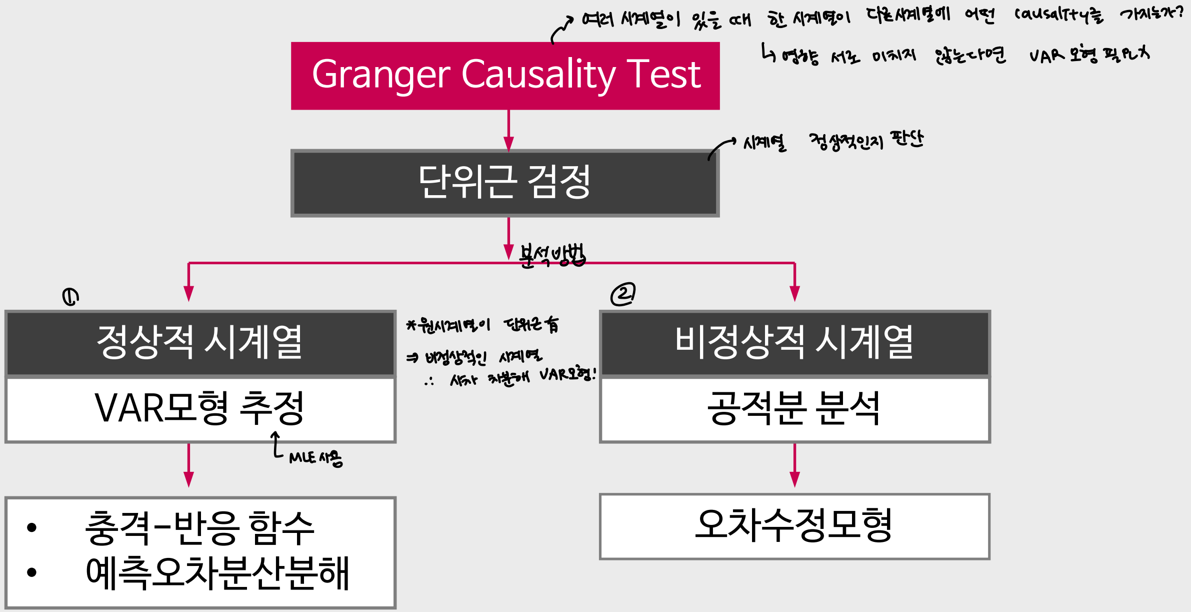
시계열 분석 기법과 응용

Week 6. 벡터자기회귀모형

6-2. 충격-반응 함수와 분산분해

전치혁 교수
(포항공과대학교 산업경영공학과)

VAR모형 분석


여러 시계열이 있을 때 한 시계열이 다른시계열에 어떤 causality를 가지는가?
↳ 영향 서로 미치지 않는다면 VAR 모형 필요X
Granger Causality Test
시계열 정상적인지 판단
단위근 검정
분석방법
①
정상적 시계열
VAR모형 추정
MLE 사용
• 충격-반응 함수
• 예측오차분산분해
*원시계열이 단위근 有
⇒ 비정상적인 시계열
∴ 시차 차분해 VAR모형!
②
비정상적 시계열
공적분 분석
오차수정모형

그래인저 인과관계

X, Y : 두 시계열

(그래인저 인과관계: Granger causality) 시계열 $\{X_t, t \geq 1\}$이 시계열 $\{Y_t, t \geq 1\}$의 미래값을 예측하는 데 도움이 될 때, $\{X_t, t \geq 1\}$이 $\{Y_t, t \geq 1\}$에 영향을 준다 (Granger-cause) 말한다.

↳ granger causal relationship : 반드시 대칭적인 것은 아님

그래인저 인과관계 검정 (가설검정)

- 자기회귀 시차모형

X의 과거값에 영향받음

$$Y_t = \alpha_0 + \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} + \beta_1 X_{t-1} + \cdots + \beta_q X_{t-q} + a_t$$

- 가설

모든 β = 0 일때 Y가 X에 대한 영향 받지 않는 것

$$H_0 : \beta_1 = \cdots = \beta_q = 0$$

- 검정통계량

F검정

$$F = \frac{(SSE_r - SSE_{ur})/q}{SSE_{ur}/(N-p-q-1)}$$

이모형의 SSE

이때의 SSE

여기서 SSE_{ur} 는 상기 모형 (비제한모형)의 SSE이며, SSE_r는 H_0가 옳을 때 (제한모형) SSE임

- 결과 해석 : 쌍방에 대해 검정. 서로 영향을 주지않을 때는 VAR모형 분석의 의미가 없음.

↳ X→Y or Y→X

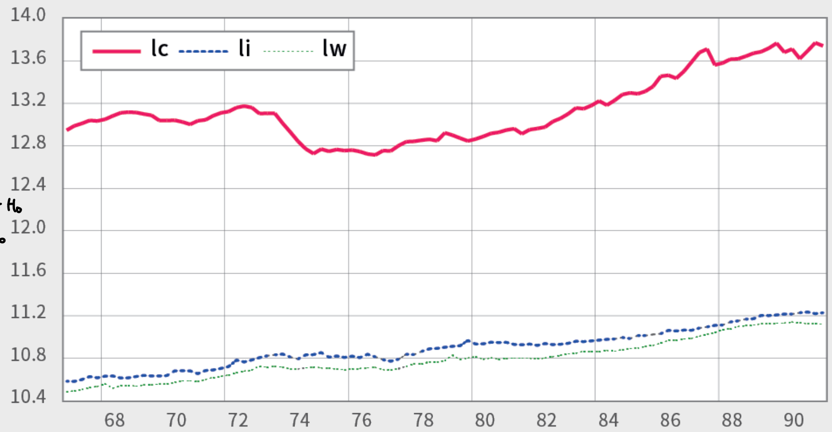
그래인저 인과관계

(예 6-1) 영국 소비(LC)/소득 (LI)/자산 (LW) 데이터에 대한 그래인저 인과관계 검정을 실시하면 아래와 같다.

EViews software

Pairwise Granger Causality Tests
Date: 07/19/19 Time: 11:13
Sample: 1966Q4 1991Q2
Lags: 4

Null Hypothesis:	Obs	F-Statistic	Prob.
LI does not Granger Cause LC → LI가 LC에 영향X	95	1.77913	0.1404 → don't reject H_0
LC does not Granger Cause LI		8.15510	1.E-05 → reject H_0
LW does not Granger Cause LC	95	2.25700	0.0695 → 나름 유의
LC does not Granger Cause LW		0.72713	0.5758
LW does not Granger Cause LI	95	1.34091	0.2614
LI does not Granger Cause LW		0.52232	0.7196



결과해석: 소비가 소득에 유의하게 영향을 미침, 유의수준 $\boxed{10\%}$에서 자산이 소비에 영향을 미침

충격-반응 함수

충격-반응 함수 (impulse-response function; IRF)

한 시계열에 특정시점에서 충격이 발생할 때 다른 시계열에 시간에 따라 어떤 영향을 주는지 분석

ex) 달러환율 상승
→ 유로, 엔화 환율, GDP, … 영향?
분석

- VAR(1)모형 예

$$Z_{1t} = \phi_{11}Z_{1,t-1} + \phi_{12}Z_{2,t-1} + a_{1t}$$

$$Z_{2t} = \phi_{21}Z_{1,t-1} + \phi_{22}Z_{2,t-1} + a_{2t}$$

충격: 시점 1에서 Z_{1t}에만 σ_1의 충격이 있고 Z_{2t}에는 충격이 없음. 다른 시점에는 충격 없음 → 이때, 미래의 Z_1, Z_2의 변화?

↳ 즉, Z_1의 white noise a_{1t}에 σ_1만큼 증가

($a_{11} = \sigma_1$, $a_{1t} = 0, t > 1$, $a_{2t} = 0, t \geq 1$; $Z_{10} = Z_{20} = 0$ 가정)

t=1

반응

t=1: $Z_{11} = a_{11} = \sigma_1$; $Z_{21} = a_{21} = 0$

Z_1에만 충격 but Z_2도 변하

t=1에서만 충격
but
t=2,3에도
Z_1, Z_2 변화!

t=2: $Z_{12} = \phi_{11}Z_{11} + \phi_{12}Z_{21} + a_{12} = \phi_{11}\sigma_1$; $Z_{22} = \phi_{21}Z_{11} + \phi_{22}Z_{21} + a_{22} = \phi_{21}\sigma_1$

t=3: $Z_{13} = \phi_{11}Z_{12} + \phi_{12}Z_{22} + a_{13} = (\phi_{11}^2 + \phi_{12}\phi_{21})\sigma_1$

$$Z_{23} = \phi_{21}Z_{12} + \phi_{22}Z_{22} + a_{23} = (\phi_{21}\phi_{11} + \phi_{22}\phi_{21})\sigma_1$$

- 반응함수: 시간에 따른 Z_{1t}에의 영향, 시간에 따른 Z_{2t}에의 영향

충격-반응 함수

VAR 모형

$$\boldsymbol{z}_t = \Phi_1 \boldsymbol{z}_{t-1} + \Phi_2 \boldsymbol{z}_{t-2} + \cdots + \Phi_p \boldsymbol{z}_{t-p} + \boldsymbol{a}_t$$

MA형태 모형

$$\boldsymbol{z}_t = \boldsymbol{a}_t + \Psi_1 \boldsymbol{a}_{t-1} + \Psi_2 \boldsymbol{a}_{t-2} + \cdots$$

직교오차 MA형태 모형

$$\boldsymbol{z}_t = \boldsymbol{v}_t + \Psi_1 \boldsymbol{v}_{t-1} + \Psi_2 \boldsymbol{v}_{t-2} + \cdots$$

$IRF(i, j, s)$
= j번째 시계열의 충격에 대한 i번째 시계열의 시간 s이후의 반응
= Ψ_s^*의 (i,j)원소

z간 상관관계

시계열간의 상관관계가 있어
IRF산출 어려움 ∴ MA형태 → white noise 독립 ∴ 분석용이

MA꼴

$\boldsymbol{a}_t \sim MVN(0, \Sigma)$ 이므로 여전히 오차항간의 상관관계가 있어 IRF산출 어려움

white noise간 공분산이 없음

직교오차 MA꼴!

$\boldsymbol{v}_t \sim MVN(0, I)$ 이므로
IRF산출 용이

충격-반응 함수

- MA형태 모형

$$\boldsymbol{z}_t = \boldsymbol{a}_t + \Psi_1 \boldsymbol{a}_{t-1} + \Psi_2 \boldsymbol{a}_{t-2} + \cdots,\ \boldsymbol{a}_t \sim MVN(0, \Sigma)$$

- Σ 의 분해 → 직교오차 MA오형 변환 위해 Σ분해 필요!

  - LDL 분해: positive definite 대칭행렬 분해

    $\Sigma = LDL^T$ (L: 대각원소가 1인 아래쪽 삼각행렬; D: 양의 원소를 갖는 대각행렬)

  - Cholesky 분해: 양의 대각원소를 갖는 삼각행렬로 분해

    $\Sigma = PP^T$ ($P = LD^{1/2}$)

  - 직교오차 변환

$$\boldsymbol{v}_t = P^{-1} \boldsymbol{a}_t$$

$$Var[\boldsymbol{v}_t] = Var[P^{-1}\boldsymbol{a}_t] = P^{-1}\Sigma(P^{-1})^T = I$$

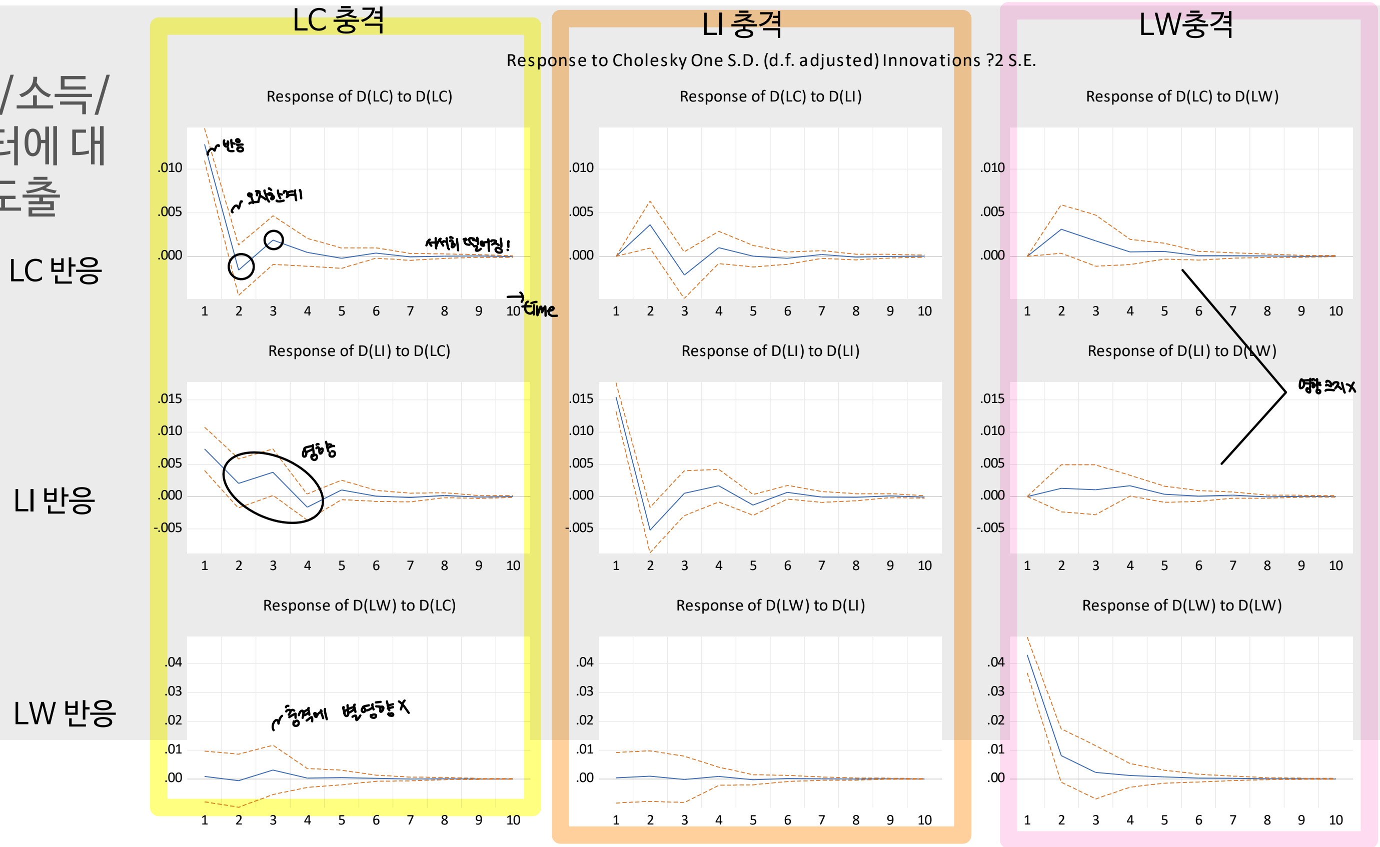
  - 직교오차 MA형태 모형

$$\boldsymbol{z}_t = \boldsymbol{v}_t + \Psi_1^* \boldsymbol{v}_{t-1} + \Psi_2^* \boldsymbol{v}_{t-2} + \cdots$$

충격-반응 함수

→연관성이 많은 시계열일수록 충격에 민감하게 반응

(예 6-2) 영국 소비/소득/자산 1차 차분 데이터에 대한 충격-반응 함수 도출



예측오차 분산분해

- **필요성**
  - 특정 시계열의 미래 불확실성에 다른 여러 시계열의 충격이 영향을 줄 수 있음
  - 이때 - 어떤 시계열이 상대적으로 어떤 영향을 끼치고 있는지 중요도 산출이 의미있음 즉, 미래 불확실성에 대해 시계열이 상대적으로 얼마나 중요한가?
  - 이를 위해 미래값을 예측하고 예측오차의 분산을 시계열별로 분해함
- **예측오차**
  - 모형: $\boldsymbol{z}_t = \boldsymbol{v}_t + \Psi_1^* \boldsymbol{v}_{t-1} + \Psi_2^* \boldsymbol{v}_{t-2} + \cdots$
  - k-단계 예측치 (예측)

$$\boldsymbol{f}_{t,k} = E[\boldsymbol{z}_{t+k} | \boldsymbol{z}_t, \ldots] = \Psi_k^* \boldsymbol{v}_t + \Psi_{k+1}^* \boldsymbol{v}_{t-1} + \cdots$$

  - k-단계 예측오차 (계산)

$$\boldsymbol{e}_{t,k} = \boldsymbol{v}_{t+k} + \Psi_1^* \boldsymbol{v}_{t+k-1} + \Psi_2^* \boldsymbol{v}_{t+k-2} + \cdots + \Psi_{k-1}^* \boldsymbol{v}_{t+1}$$

  - i-번째 시계열의 k-단계 예측오차
  - $e_{t,k}^{(i)} = \sum_{j=1}^{m} v_{j,t+k} + \sum_{j=1}^{m} \psi_{ij,1}^* v_{j,t+k-1} + \cdots + \sum_{j=1}^{m} \psi_{ij,k-1}^* v_{j,t+1}$

ㄴ이 예측오차의 분산을 분해!

예측오차 분산분해

- 예측오차 분산분해 (variance decomposition)

① i-번째 시계열의 k-단계 예측오차 계산

$$e_{t,k}^{(i)} = \sum_{j=1}^{m} v_{j,t+k} + \sum_{j=1}^{m} \psi_{ij,1}^{*} v_{j,t+k-1} + \cdots + \sum_{j=1}^{m} \psi_{ij,k-1}^{*} v_{j,t+1}$$

② i-번째 시계열의 k-단계 예측오차 분산 계산 → 각 시계열의 비중 계산

$$Var[e_{t,k}^{(i)}] = m + \sum_{j=1}^{m} \left(\psi_{ij,1}^{*}\right)^2 + \cdots + \sum_{j=1}^{m} \left(\psi_{ij,k-1}^{*}\right)^2 = \sum_{s=0}^{k-1} \sum_{j=1}^{m} \left(\psi_{ij,s}^{*}\right)^2$$

이중 j-번째 시계열 분산 비중

$$Var[e_{t,k}^{(ij)}] = 1 + \left(\psi_{ij,1}^{*}\right)^2 + \cdots + \left(\psi_{ij,k-1}^{*}\right)^2 = \sum_{s=0}^{k-1} \left(\psi_{ij,s}^{*}\right)^2$$

③ i-번째 예측오차분산 중 j-번째 시계열 기여율

↳ 상대적 기여율 분석 → 어느 시계열이 중요한지 파악

$$R_{ij,k} = \frac{Var[e_{t,k}^{(ij)}]}{Var[e_{t,k}^{(i)}]} \times 100 = \frac{\sum_{s=0}^{k-1} \left(\psi_{ij,s}^{*}\right)^2}{\sum_{s=0}^{k-1} \sum_{j=1}^{m} \left(\psi_{ij,s}^{*}\right)^2} \times 100(\%)$$

i번째 시계열의 k단계 예측오차분산에 대한 j번째 시계열의 기여율

percentage 단위

예측오차 분산분해

(예 6-3) 영국 소비/소득/자산 데이터에 대한 예측오차 분산분해

차분해줌

Variance Decomposition of D(LC):

Period	S.E.	D(LC)	D(LI)	D(LW)
1	0.012824	100.0000	0.000000	0.000000
2	0.013775	87.97436	6.915991	5.109644
3	0.014178	84.78336	8.791400	6.425242
4	0.014230	84.26114	9.232929	6.505934
5	0.014244	84.12672	9.215815	6.657461
6	0.014251	84.11612	9.230617	6.653260
7	0.014253	84.09334	9.250909	6.655749
8	0.014253	84.08904	9.253813	6.657074
9	0.014253	84.08914	9.253787	6.657074
10	0.014253	84.08885	9.253944	6.657206

S.E. → 미래시점에 따른 표준오차

→ 처음에 LC가 100% 기여

미래시점

상대적 중요도 다른 시계열에서 증가 (LC제외)

→ 어느정도 증가 이후 일정하게 유지

↳ 분석을 통해 자체(충격받은) 시계열의 다른 시계열이 얼마나 큰 영향,

즉, 상대적 중요성 볼 수 있음

시계열 분석 기법과 응용

Week 6. 벡터자기회귀모형
6-3. 공적분과 오차수정 모형

전치혁 교수
(포항공과대학교 산업경영공학과)

VAR모형 분석

Granger Causality Test

↓

단위근 검정

↓

- 정상적 시계열
  - VAR모형 추정
    - ↓
    - • 충격-반응 함수
    - • 예측오차분산분해
- 비정상적 시계열
  - 공적분 분석
    - ↓
    - 오차수정모형

비정상적 시계열 분석

필요성

- VAR모형은 정상적 시계열에 대하여 적용
  - 비정상적 시계열에 대해서는 차분을 통해 정상성 확보후에 VAR모형화
- 추세, 계절성, 사이클 존재
  경제/금융관련 시계열들에 각각은 비정상적이나 장기적으로 균형적 관계를 갖는 경우가 있으며 이를 공적분 (cointegration)관계라 함. (예) 소득과 소비 : 연도별 소득 증가 → 소비도 이에 따라 증가
- 이런 경우 각각을 차분하여 정상적 시계열로 변환하여 분석하는 것보다 직접적으로 회귀 모형화하는 것이 보다 많은 정보를 얻음. → 비정상적 시계열 직접적 분석 이유
- 관계없으나 두 시계열이 증가한다는 가정 하 회귀분석 결과로 나오는 상관관계는 가짜상관!
  공적분관계가 없는 비정상적 시계열을 대상으로 회귀분석할 때 가성 회귀 (spurious regression)의 문제가 발생함에 유의
- 공적분 분석을 위해서는 각 시계열이 동일 차수의 (비정상적) 누적 시계열이어야 함.

공적분

*원시계열 x_t : 비정상적 → 1차 차분 → $(1-B)x_t$: 비정상적? → 2차 차분 → $(1-B)^2 x_t$: 정상적 ⇒ 원시계열 x_t : 차수 2의 누적 벡터시계열 = $I(2)$
차수

누적 (integrated) 벡터시계열

- 벡터 시계열 $\{\boldsymbol{x}_t, t \geq 1\}$에 대해 $\{(1-B)^{d-1}\boldsymbol{x}_t, t \geq 1\}$는 비정상적이나 $\{(1-B)^{d}\boldsymbol{x}_t, t \geq 1\}$이 정상적일 때, 차수 d의 누적 벡터시계열이라 하며 $\{\boldsymbol{x}_t, t \geq 1\}\sim I(d)$ 로 표기

공적분 (cointegration)

ex) 2차 누적벡터시계열 $I(2)$ → 선형결합을 통해 정상적 시계열이 될 때 "공적분 관계"

- $I(d)$ 인 누적 벡터시계열에 대해 선형결합 $\boldsymbol{\alpha}^T\boldsymbol{x}_t$이 차수 d 미만의 누적 시계열이 될 때, 공적분 벡터 $\boldsymbol{\alpha}$를 갖는 공적분 관계에 있다 함
  ↳ 즉, 선형결합으로 차수 적어지는 경우 (d=1, 즉, 0차 누적 벡터시계열 : 정상적 시계열)
- 총 m개의 시계열이 있을 때 최대 m−1개의 공적분 벡터가 있을 수 있음. 이 때 존재하는 공적분 벡터의 최대수를 공적분 랭크 (cointegration rank)라 함
- (예) 다음 세 시계열을 고려하자.
  random walk ∴ 1차차분 → 백색잡음
  - $Z_{1t} = Z_{1,t-1} + a_{1t}; Z_{2t} = \beta_1 Z_{1,t} + a_{2t}; Z_{3t} = \beta_2 Z_{1,t} + \beta_3 Z_{2,t} + a_{3t}$
    ↳ 1차누적인 Z_1으로 구성 ↳ 1차누적인 Z_1, Z_2로 구성
  - 각 시계열은 $I(1)$
  선형결합 → 정상적 시계열
  - $-\beta_1 Z_{1,t} + Z_{2t} + 0Z_{3t} = a_{2t}\sim I(0)$; $-\beta_2 Z_{1,t} - \beta_3 Z_{2,t} + Z_{3t} = a_{3t}\sim I(0)$
    2개의 공적분 vector!
  - 따라서 공적분 랭크는 2

오차수정모형

공적분의 활용 사례

- 필요성
  - 여러 시계열이 공적분 관계가 있는 경우 오차수정모형 (error correction model; ECM)을 통하여 시계열 상호간의 미치는 단기 및 장기 효과를 분석할 수 있음.
  - 공적분 관계는 ECM 표현을 위한 필요충분조건임 (그래인저 표현정리-Engle and Granger, 1987).

ㄴ 즉, 공적분 관계 有 $\longleftrightarrow$ ECM 표현 가능

- ECM 유도 예
  - 두 시계열 $\{X_t\}$와 $\{Y_t\}$가 각각 $\mathrm{I}(1)$이며 다음의 공적분 관계가 있다고 하자.

$$\underline{Y_t = \beta X_t + \varepsilon_t}$$

  - 이 관계식을 다음과 같이 쓸수있다.

Y_{t-1}을 빼고 더함 $\qquad$ X_{t-1}을 빼고 더함

$$Y_t - (Y_{t-1} - Y_{t-1}) = \beta X_t + \beta(X_{t-1} - X_{t-1}) + \varepsilon_t$$

$\longrightarrow$ $Y_t - Y_{t-1} = \beta X_t \underline{+ \beta X_{t-1}} - \beta X_{t-1} + \varepsilon_t \underline{- Y_{t-1}}$

실제값 $\qquad$ 예측값

$$\Downarrow\ e_{t-1} = Y_{t-1} - \beta X_{t-1} \text{ (이전 시점 오차)}$$

$$\Delta Y_t = -e_{t-1} + \beta \Delta X_t + \varepsilon_t$$

$\Delta Y_t = -e_{t-1} + \underline{\beta X_t - \beta X_{t-1}} + \varepsilon_t$

$\Delta Y_t = -e_{t-1} + \beta \cdot \Delta X_t + \varepsilon_t$

Y차분 → 이전시점 오차와 X의 차분으로 표현됨

⇒ ECM의 가장 기본적 형태

오차수정모형

- 단순형태 ECM

항상 음수

$$\Delta Y_t = \lambda e_{t-1} + \beta \Delta X_t + \varepsilon_t$$

- λe_{t-1} $(\lambda < 0)$: 오차수정항 (error correction term) – 이전 시점에서 예측오차가 양수이면 (실제값이 예측값보다 크면) 다음 시점에서 Y값을 축소시켜 X와 Y의 장기적관계를 유지시키는 효과 → X, Y 관계식이 유지되도록 오차 커지면 Y값 축소 작아지면 〃 증가 ???
- X값에 변화가 있으면 Y값에 영향을 줌 (단기 효과)

↳ ΔX가 ΔY에 영향

시차 : p일때

- 시차변수를 포함한 ECM

$$\Delta Y_t = \lambda e_{t-1} + \sum_{j=1}^{p} \beta_{1j} \Delta X_{t-j} + \sum_{j=1}^{p} \gamma_{1j} \Delta Y_{t-j} + \varepsilon_{1t} \qquad (e_{t-1} = Y_{t-1} - \beta X_{t-1})$$

$$\Delta X_t = \lambda e_{t-1} + \sum_{j=1}^{p} \beta_{2j} \Delta X_{t-j} + \sum_{j=1}^{p} \gamma_{2j} \Delta Y_{t-j} + \varepsilon_{2t}$$

: ΔX를 ΔY, X 이전 변화량으로 표현 가능

∴ X, Y 동시에 표현하는 벡터에 대해 ECM 고려 O

$$\Rightarrow \begin{pmatrix} \Delta X_t \\ \Delta Y_t \end{pmatrix} = \begin{bmatrix} -\lambda\beta & \lambda \\ -\lambda\beta & \lambda \end{bmatrix} \begin{pmatrix} X_{t-1} \\ Y_{t-1} \end{pmatrix} + \sum_{j=1}^{p} \begin{bmatrix} \beta_{2j} & \gamma_{2j} \\ \beta_{1j} & \gamma_{1j} \end{bmatrix} \begin{pmatrix} \Delta X_{t-j} \\ \Delta Y_{t-j} \end{pmatrix} + \begin{pmatrix} \varepsilon_{2t} \\ \varepsilon_{1t} \end{pmatrix}$$

오차수정모형

- VAR모형과 벡터 ECM (VECM)
  - 벡터 시계열의 경우 공적분관계가 있을 때 벡터 ECM (VECM)으로 확장됨
  - $I(1)$인 VAR(p)모형은 공적분관계 여부와 관계없이 다음의 VECM으로 표현

$$\boldsymbol{z}_t = \Phi_1 \boldsymbol{z}_{t-1} + \Phi_2 \boldsymbol{z}_{t-2} + \cdots + \Phi_p \boldsymbol{z}_{t-p} + \boldsymbol{a}_t$$

$$\Downarrow \quad \Pi = \Phi_1 + \Phi_2 + \cdots + \Phi_p - I, \; \Gamma_k = -\sum_{j=k+1}^{p} \Phi_j$$

~여러시차

$$\Delta \boldsymbol{z}_t = \Pi \boldsymbol{z}_{t-1} + \sum_{j=1}^{p-1} \Gamma_j \, \Delta \boldsymbol{z}_{t-j} + \boldsymbol{a}_t$$

↑ vector 차분

  - 행렬 Π $(m \times m)$의 랭크 크기에 따라 3가지 경우

① $\Pi = 0$인 경우: 공적분 관계 없으며 차분 형태가 정상적인 VAR(p-1)모형으로 변환

② Π의 랭크가 m인 경우: VAR(p)모형이 이미 정상적

~ β에 대한 해석

③ Π의 랭크가 0과 m 사이인 경우: 실제적 VECM. 공적분 관계식이 랭크수만큼 있음. 이 때 $\Pi = AB^T$ 가 성립하며 행렬 B의 열들이 공적분 벡터들이 됨.

공적분 검정 필요
(공적분관계 있는가?)

Engle-Granger 검정

- 회귀분석의 잔차 시계열에 대하여 ADF검정 실시
- 단위근이 없다면 공적분 관계있다고 결론

Johansen 검정 → 주로 사용, 공적분 rank가 몇 개인가

- 트레이스 검정
  - $H_0 : r = r_0$ vs $H_1 : r > r_0$
  - $r_0 = 0$에서 귀무가설 기각되면 다음으로 $r_0 = 1$을 검정하는 과정을 반복
  - 검정통계량은 우도비에 바탕으로 산출
- 최대고유치 검정
  - $H_0 : r = r_0$ vs $H_1 : r = r_0 + 1$
  - 역시 $r_0 = 0, 1, \ldots$ 등으로 변화시키면서 순차적으로 진행

공적분과 오차수정모형 예

(예 6-4) 옆의 그림은 2016.1~ 2018.12 사이 kospi200과 kospi200 선물 지수를 나타낸다. 다음을 알아보자

1. 공적분 검정
2. 공적분 관계
3. 오차수정모형

공적분과 오차수정모형 예

(예 6-4 계속)

다음은 공적분 검정 결과이다.

- 트레이스 검정 과 최대고유치 검정에서 두 시계열간 공적분관계가 있다고 판정됨

Sample (adjusted): 1/11/2016 12/28/2018
Included observations: 728 after adjustments
Trend assumption: Linear deterministic trend
Series: KOSPI200 KOSPI200_FUTURE
Lags interval (in first differences): 1to 4

Unrestricted Cointegration Rank Test (Trace)

Hypothesized No. of CE(s)	Eigenvalue	Trace Statistic	0.05 Critical Value	Prob.**
None *	0.051946	41.25473	15.49471	0.0000
At most1	0.003319	2.420325	3.841466	0.1198

Trace test indicates 1 cointegrating eqn(s)at the 0.05 level
*denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

Unrestricted Cointegration Rank Test (Maximum Eigenvalue)

Hypothesized No. of CE(s)	Eigenvalue	Max-Eigen Statistic	0.05 Critical Value	Prob.**
None*	0.051946	38.83440	14.26460	0.0000
At most 1	0.003319	2.420325	3.841466	0.1198

Max-eigenvalue test indicates 1 cointegrating eqn(s) at the 0.05 level
*denotes rejection of the hypothesis at the 0.05 level
**MacKinnon-Haug-Michelis (1999) p-values

공적분과 오차수정모형 예

(예 6-4계속) 공적분관계식과 오차수정모형의 추정식을 보여준다.

공적분 관계식

$$k200_{t-1} = 0.2155 + \underbrace{0.9982}_{\text{계수} \approx 1} k200\_f_{t-1}$$

오차수정모형

계수 : 오차 발생 시 얼마나 빨리 조정되는가

$$\Delta k200_t = 0.0364 - 0.1183e_{t-1} - 0.1867\Delta k200_{t-1} - 0.0136\Delta k200_{t-2} + 0.1561\Delta k200_{f_{t-1}} + 0.0814\Delta k200_{f_{t-2}}$$

$$\Delta k200\_f_t = 0.0365 + 0.0622e_{t-1} + 0.0811\Delta k200_{t-1} + 0.045\Delta k200_{t-2} - 0.1402\Delta k200_{f_{t-1}} - 0.0133\Delta k200_{f_{t-2}}$$

Cointegrating Eq:	CointEq1
KOSPI200(-1)	1.000000
KOSPI200_FUTURE(-1)	-0.998212 (0.00329) [-303.412]
C	-0.215456

추정됨

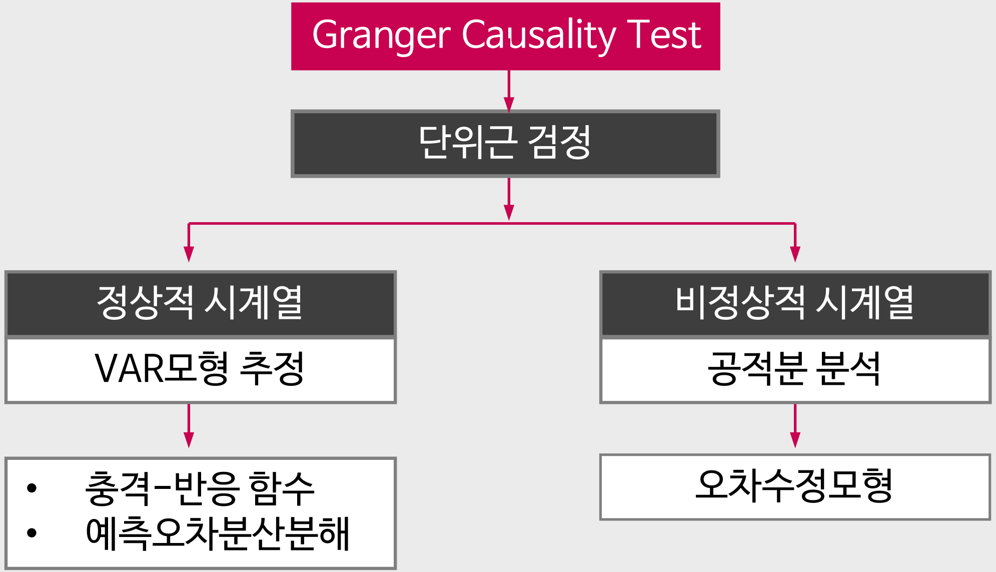
2개의 ECM 추정

Error Correction:	D(KOSPI200)	D(KOSPI20…
CointEq1	-0.118302 (0.12215) [-0.96850]	0.062194 (0.12706) [0.48949]
D(KOSPI200(-1))	-0.186666 (0.18259) [-1.02233]	0.081089 (0.18993) [0.42695]
D(KOSPI200(-2))	-0.013611 (0.17574) [-0.07745]	0.045013 (0.18280) [0.24624]
D(KOSPI200_FUTURE(..	0.156070 (0.17797) [0.87693]	-0.140201 (0.18512) [-0.75734]
D(KOSPI200_FUTURE(..	0.081424 (0.17234) [0.47245]	0.013280 (0.17927) [0.07408]
C	0.036415 (0.08482) [0.42932]	0.036494 (0.08823) [0.41363]

se

t값 (검정가능)

VAR모형 분석


Granger Causality Test
단위근 검정
정상적 시계열
VAR모형 추정
• 충격-반응 함수
• 예측오차분산분해
비정상적 시계열
공적분 분석
오차수정모형