제 1 장 시계열자료

1.1 시계열(time series)자료

년도별, 계절별, 월별, 일별, 시, 분, 초별로 시간의 흐름에 따라 관측된 자료

① 국민 총 생산액, 물가지수, 총 수출액, 주가지수 등과 같이 경제활동과 관련된 시계열(economic time series)

② 일일 강수량, 기온, 태양의 흑점 수, 년간 지진의 발생 수 등과 같이 물리적 현상과 관련된 시계열(physical time series)

③ 상품판매량, 상품광고액, 상품재고량, 상품매출액 등과 같이 회사의 경영활동과 관련된 시계열(marketing time series)

④ 총인구, 농가 수, 인구증가율, 평균결혼연령 등과 같이 인구와 관련된 시계열(demographic time series)

⑤ 품질관리 등과 같은 생산관리와 관련된 시계열(time series in process control)

⑥ (0,1)-확률과정, 음성파와 같이 통신 공학 또는 공학과 관련된 시계열

⑦ 월별 교통사고 건수, 월별 범죄 발생 수와 같이 사회생활과 관련된 시계열

시계열자료의 구분

연속시계열(continuous time series)

이산시계열(discrete time series)

시차(time lag) : 관측시점과 관측시점들 사이의 간격

시계열자료의 분석 목적:

1) 미래에 대한 예측(forecast)

2) 시스템 또는 확률과정의 이해와 제어(control) :

1.2 시계열자료의 여러 형태

시계열그림(time series plot)

시간의 경과에 따라 시계열자료의 값이 변하는 것을 그린 그림

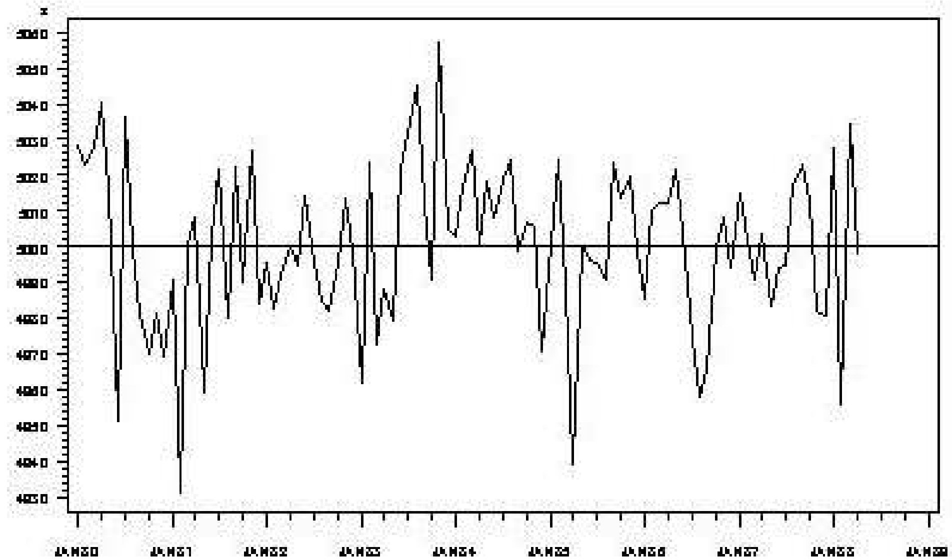
시계열 자료의 성분

불규칙성분(irregular component)

체계적성분(systematic component)

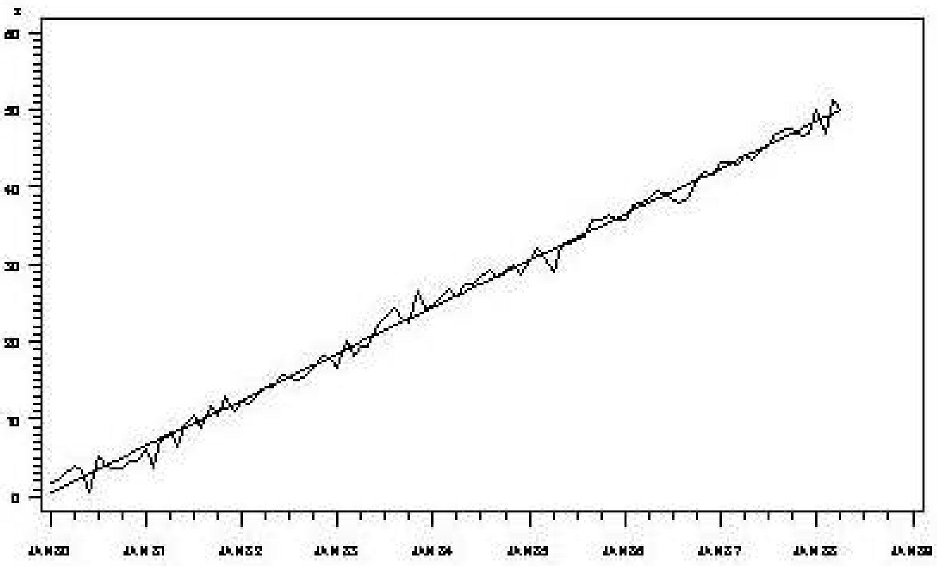
추세성분(trend component), 계절성분(seasonal component), 순환성분(cyclical component)

(1) 불규칙성분 : 시간에 따른 규칙적인 움직임과는 무관하게 랜덤한 원인에 의한 변동성분



<그림 1.1> 불규칙성분으로 구성된 시계열의 시계열그림

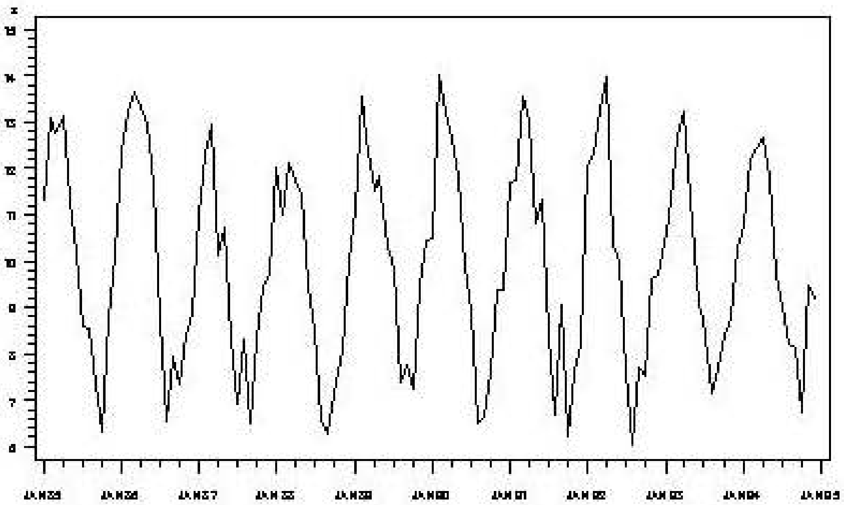
2) 추세성분 : 관측값이 지속적으로 증가 하거나 감소하는 추세를 갖는 경우의 변동



<그림 1.2> 추세성분을 갖는 시계열의 시계열그림

(3) 계절성분 : 계절의 변화에 따른 주기적인 변동

주별, 월별, 계절별과 같이 주기적인 성분에 의한 변동

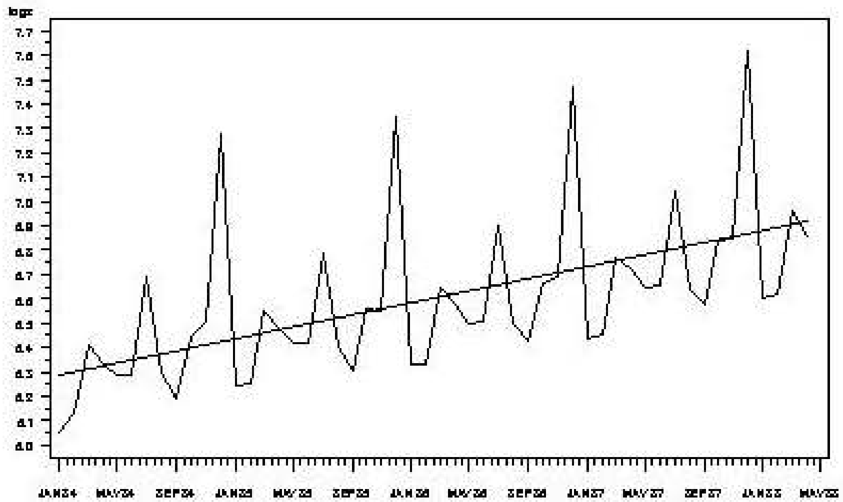


<그림 1.3> 계절성분을 갖는 시계열의 시계열그림

(4) 순환성분

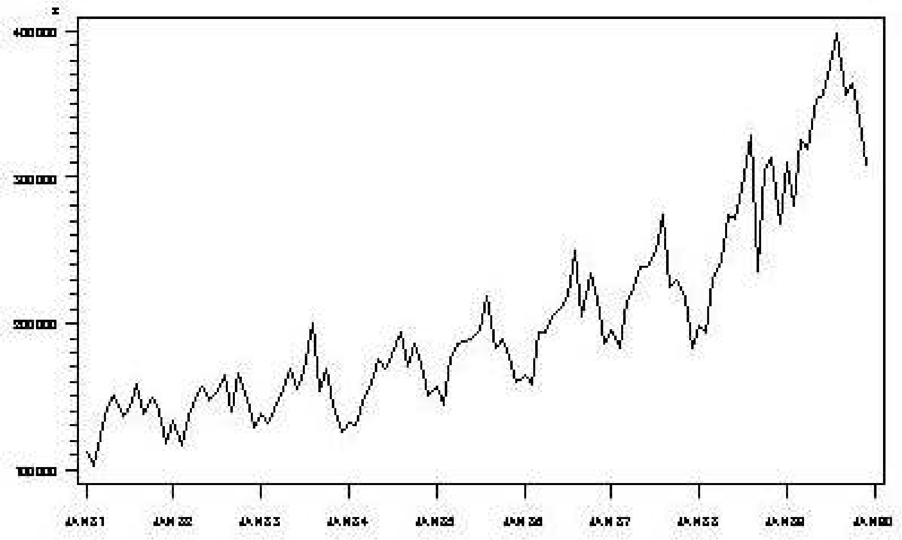
: 주기적인 변화를 가지나 변화가 계절에 의한 것이 아니고, 주기가 긴 경우의 변동
경기순환(business cycle)

(5) 추세성분과 계절성분을 동시에 포함하는 시계열



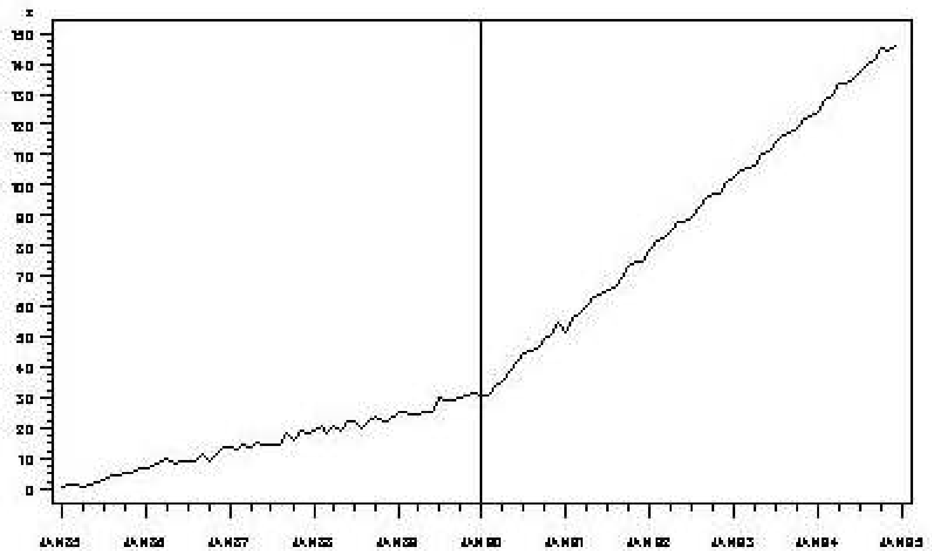
<그림 1.4> 추세성분과 계절성분을 갖는 시계열의 시계열그림

시간의 증가에 따라 변동의 폭이 커지는 시계열
  자료분석 시에 로그(logarithm)변환을 먼저 해주는 것이 일반적



<그림 1.5> 추세성분과 계절성분을 갖고 시간의 변화에 따라 변동폭이 커지는 시계열의 시계열그림

(6) 두 개의 추세선(trend line)을 갖는 시계열



<그림 1.6> 추세선이 두 개인 시계열의 시계열그림

1.3 시계열자료의 분석법 및 모형

1.3.1 분석 목적 및 방법

A. 시계열자료의 분석 목적

(1) 예측(forecast) 목적으로 주로 사용되는 방법
추세분석(trend analysis),
평활법(smoothing method),
분해법(decomposition method),
자기회귀누적이동평균(AutoRegressive Integrated Moving Average: ARIMA)모형
이분산자기회귀모형(Autoregressive Conditional Heterskedasticity)

(2) 시스템의 이해와 제어의 목적으로 주로 사용되는 방법
스펙트럼분석(spectral analysis),
개입분석(intervention analysis),

전이함수모형(transfer function model)
자기상관오차를 갖는 회귀모형(autoregressive error model)

B. 적용 분야에 따른 두 가지 접근법

(1) 진동수영역(frequency domain)에서의 분석법
푸리에분석(Fourier analysis) :
· 태양 흑점(sunspot)의 수 또는 밀 가격지수(index of wheat price)
· 스펙트럼 밀도함수(spectral density function)의 추정
· 공학이나 이학 분야의 자료와 같이 생성되는 시계열이 일정한 패턴을 따르는 정상적 (stationary)인 경우에 많이 사용됨
파엽분석(wavelet Analysis)

(2) 시간영역(time domain)에서의 분석법
· 자기상관함수(autocorrelation function) 등을 이용한 시간에 따른 상관 정도를 파악
· 추세분석, 평활법, 분해법, $ARIMA$ 모형, 전이함수모형, 자기상관오차를 갖는 회귀모형 등, 이분산 모형(GARCH) 모형, long memory process 등

1.3.2 분석모형의 구분

(1) 결정모형(deterministic model) 또는 확정모형

확률적 요소를 전혀 갖지 않는 모형

$y = f(x_1, \ldots, x_p; \beta_1, \ldots, \beta_m)$: $f, \beta_1, \ldots, \beta_m$은 알려져 있음,

(2) 확률모형(stochastic model)

확률오차를 수반하는 모형으로

$y = f(x_1, \ldots, x_p; \beta_1, \ldots, \beta_m) +$ 오차 , $f, \beta_1, \ldots, \beta_m$은 과거의 자료들에 의해 결정됨

오차 : 서로 독립이고 동일한 분포를 따르는 확률변수(random variable)

1.4 예측모형의 종류

· 주관적인 모형(qualitative 또는 subjective model)

· 객관적인 모형(quantitative 또는 objective model)

1.5. 모형적합의 3단계 및 예측시스템

(i) 모형의 식별단계(model identification)

시계열그림, 자기상관함수, 부분자기상관함수 등을 이용

차분의 필요 여부와 모형의 차수를 잠정적으로 결정

(ii) 모형의 추정(model estimation) 단계

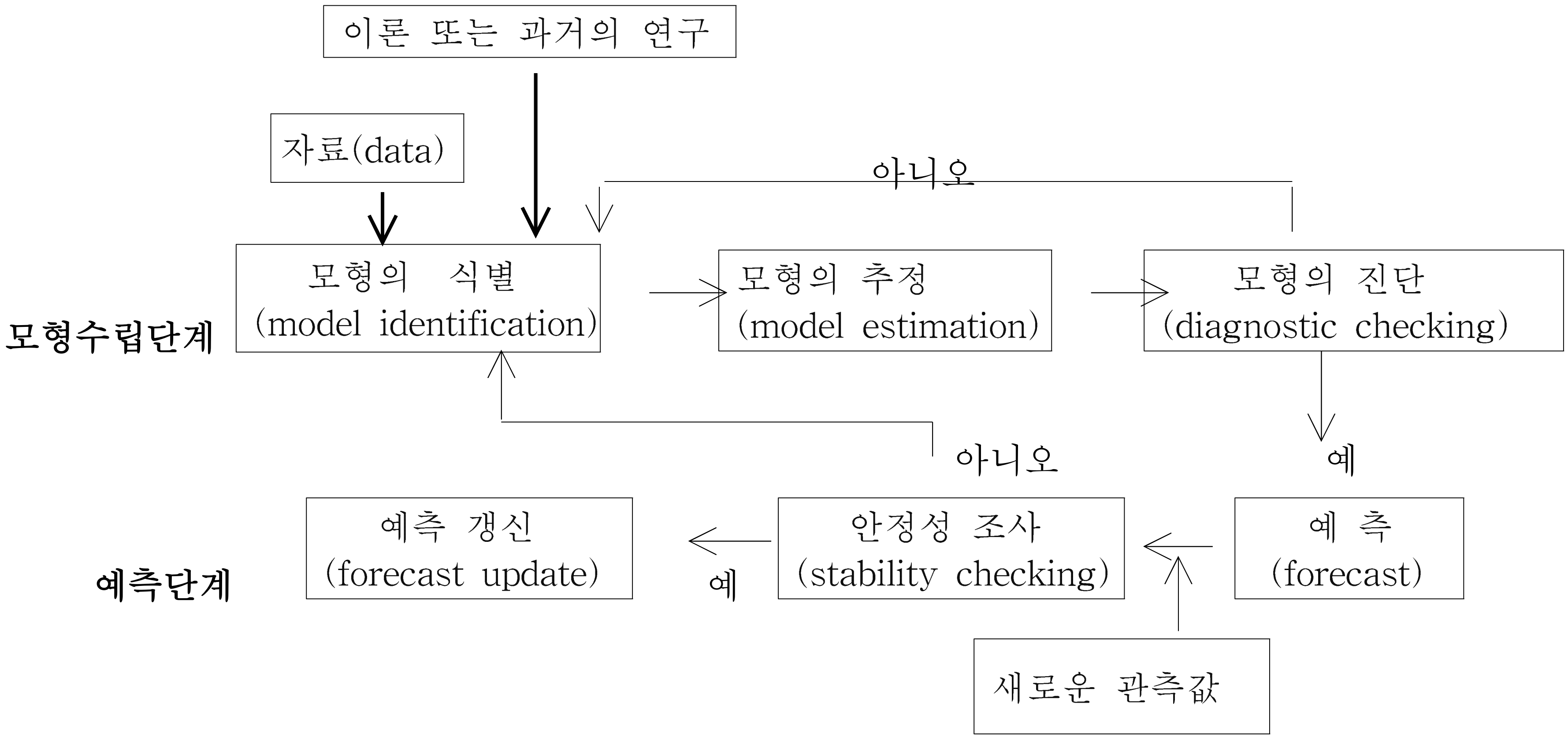
최소제곱법, 최대가능도방법, 비선형최소제곱법(nonlinear least squares method) 등을 이용하여 모수들을 추정

(iii) 모형의 진단(model diagnostic checking) 단계

잔차의 시계열그림, 잔차의 자기상관함수와 부분자기상관함수 및 Ljung과 Box(1978)에 의한 수정된 포트맨토(portmanteau)통계량을 이용한 잔차분석과 과대적합에 의하여 모형의 적합 정도를 진단

AIC(Akaike Information Criterion) 또는 SBC(Schwartz Bayesian Criterion) 통계값 등을 이용하여 가장 설명력이 높은 모형을 선택

예측시스템



<그림 1.7> 예측시스템의 구조

1.6 예측방법의 선택기준

정확성 , 예측기간의 길이(forecast horizon), 복잡성, 이용 가능한 자료의 종류

$Z_t, t = 1, 2, ..., n$에 기초한 시점 t에서의 1-시차 후의 예측값 : $\hat{Z}_t(1)$

1-시차 후의 예측오차 (one-step-ahead prediction error) : $\hat{e}_t(1) = Z_{t+1} - \hat{Z}_t(1)$

① 평균제곱예측오차: MSE(mean square prediction error)

$$MSE = \frac{\sum_{t=1}^{m} \widehat{e_{n-1+t}(1)^2}}{m}$$

② 평균절대백분위예측오차: MAPE(mean absolute percentage prediction error)

$$MAPE = \frac{100}{m} \sum_{t=1}^{m} \left| \frac{e_{n-1+t}(1)}{Z_{n+t}} \right|$$

③ 평균절대예측오차 : MAE(mean absolute prediction error)

$$MAE = \frac{\sum_{t=1}^{m} |\widehat{e_{n-1+t}}(1)|}{m}$$

1.7 시계열의 역사

(1) 푸리에분석(Fourier analysis)과 잔물결(wavelets)

태양의 흑점자료와 밀 가격지수를 분석
감쳐진 주기성의 탐색

Fourier(1807) : 임의의 함수는 sine과 cosine의 선형결합으로 근사
Stokes(1879) : 스펙트럼분석이 제안

스펙트럼밀도함수의 추정

Schuster(1906)와 Beveridge(1922) : 주기도분석(periodogram analysis)

Bartlett(1948, 1950)와 Blackman(1958) : 주기도의 평활 방법

잔물결(wavelets)

Donoho와 Johnstone(1994)

(2) 평활법(smoothing method)과 계절조정(seasonal adjustment)

Holt(1957) : 지수평활법

Brown(1959)과 Holt 등(1960) : 일모수(one paremeter) 지수평활법

Winters(1960) : 계절형(seasonal) 지수평활법

Brown(1961), Brown과 Meyer(1961)

Theil과 Wage(1964) : 지수평활법을 확률모형과 연관시켜 평활계수(smoothing parameter)를 평균제곱예측오차를 최소화하는 방법에 의해 구할 수 있음

이동평균법 : 계절조정(seasonal adjustment)의 목적에 주로 이용

Copeland(1915), Macauley(1930), Shiskin(1957, 1961) 및 McLaughlin(1962) :

미국 상무부(Census Bureau)가 발표한 X-11(Shiskin 등, 1967)
Canada 통계청(Statistics Canada) : X-11 $ARIMA$(Dagum, 1975)
미국 상무부 : X-12 $ARIMA$(Findley 등, 1998)

(3) 자기회귀이동평균모형에 의한 분석법

Yule(1926, 1927) : 시계열을 과거의 관측값들의 함수 형태로 표현
Walker(1931) : 자기회귀(autoregressive)의 개념을 제안
Slutsky(1937) : 이동평균(moving average)의 개념을 제안
Wold(1954) : 월드의 분해정리(Wold's decomposition theory)
Walker(1962) : 자기회귀이동평균(AutoRegressive Moving Average: $ARIMA$) 모형을 제안
Mann과 Wold(1943) : 자기회귀(AR)모형에서의 최대가능도추정법
Durbin(1959, 1960) : AR모형과 이동평균(Moving Average:MA)모형에서의 모수추정법
Jenkins와 Watts(1968) 그리고 Box와 Jenkins(1970) : 효율적인 계산방법 및 3단계 모형적합법이 제안
Box와 Tiao(1975) : 개입분석(intervention analysis)을 제안
Chang 등(1988) : Fox(1972)에 의해 처음으로 제안된 이상점(outlier)을 찾는 방법을 제안

Whittle(1953) : 다변량(multivariate)시계열 분석법
Hannan(1970), Hillmer와 Tiao(1979)와 Tiao와 Box(1981) :
전이함수모형(transfer function model) : 출력시계열과 입력시계열사이의 관계를 모형화, Box 등(1994)과 Wei(1991)

Harvey(1984) : 상태공간모형(state space model)
Kalman(1960)과 Kalman과 Bucy(1961) : Kalman 필터(filter)
Harrison과 Stevens(1971, 1976) : 베이즈(Bayesian) 예측방법
Fuller와 Dickey(1979), Dickey 등(1984) 및 Phillips와 Perron(1988) : 단위근검정(unit root test)
Granger(1983)와 Engle과 Granger(1987) : 공적분검정(cointegration test)
Engle(1982) : GARCH 류의 이분산모형, Bollerslev(1986)