시계열 분석 개요

#01. 시계열 자료의 이해

- 시간의 흐름에 따라 관찰된 값들
- 시계열 데이터의 분석을 통해 미래의 값을 예측하고 경향, 주기, 계절성 등을 파악하여 활용한다.

1. 시계열 자료의 종류

종류	설명
비정상성 시계열 자료	시계열 분석을 실시할 때 다루기 어려운 자료로 대부분의 시계열 자료
정상성 시계열 자료	비정상 시계열을 핸들링해 다루기 쉬운 시계열 자료로 변환한 자료

정상성

평균과 분산이 일정하고 공분산도 단지 시차에만 의존하고 특정 시점에는 의존하지 않는 상태

평균이 일정할 경우

- 모든 시점에 대해 일정한 평균을 갖는다
- 실제 대부분의 자료는 평균이 일정하지 않다. 이 경우 차분(Difference)을 통해 정상화 할 수 있다.

분산이 일정할 경우

- 분산도 특정 시점에 의존하지 않고 일정해야 한다.
- 분산이 일정하지 않을 경우 변환(Transformation)을 통해 정상화 할 수 있다.

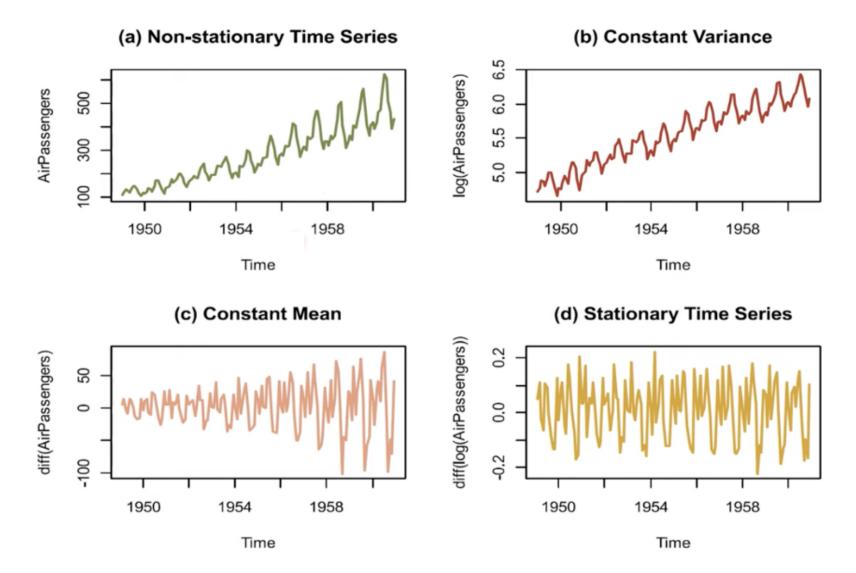
약한 의미의 정상성(약정상성)

모든 시점에 평균 일정, 시점과 분산 독립, 공분산은 시차에만 의존

비정상→정상 : 변환(transformation), 차분(difference)

이름	설명
변환	분산이 일정하지 않은 비정상 시계열에 대해 수행
차분(t1-t0)	평균이 일정하지 않은 비정상 시계열에 대해 수행
일반차분 (regular difference)	바로 전 시점의 자료를 빼는 방법이다
계절차분 (seasonal difference)	여러 시점 전의 자료를 빼는 것 방법 , 주로 계절성을 갖는 자료를 정상화 하는데 사용한다

정상 시계열



구 분	설명
(a)	일반적으로 수집되는 비정상 시계열. 시간에 따라 변동폭이 일정하지 않고, 추세와 계절적 영향이 존재하 는 형태
(b)	(a)상태의 비정상 시계열에 로그 변환을 수행하여 변동폭을 일정하게 변경한 상태 변환을 수행하여 평균을 일정하게 처리함
(c)	(a)상태의 비정상 시계열에 차분을 수행하여 평균을 일정하게 맞춘 상태. 1차 차분으로 정상성을 띄지 않으면 반복 수행한다. (n차)
(d)	(b),(c)를 함께 적용하여 정상 시계열로 변한한 상태

#02. 시계열 분석 모형

1. AR 모형 (자기회귀모형)

- 특정 시점 전의 자료가 현재 자료에 영향을 주는 형태
- 판단조건: 자기상관함수(ACF)가 빠르게 감소하고 부분자기상관함수(PACF)는 어느 시점에 절단점을 갖음

$$\phi_1 Z_{t-1} + \phi_2 Z_{t-2} + \dots + \phi_p Z_{t-p} + a_t$$

2. MA 모형 (이동평균모형)

- 유한한 개수의 백색잡음의 결합.
- 항상 정상성을 만족
- ACF에서 절단점 갖고 PACF가 빠르게 감소

$$Z_{t} = a_{t} - \theta_{1}a_{t-1} - \theta_{2}a_{t-2} - \dots - \theta_{p}a_{t-p}$$

3. ARIMA 모형 (자기회귀누적이동평균모형)

- 가장 일반적인 모형 (이것만 사용)
- 비정상시계열 모형
- 차분이나 변환을 통해 AR/MA/ARMA 모형으로 정상화 가능
- 세가지 지표를 사용하는 \$(p,d,q)\$ 모형이라고도 함

\$p\$는 AR 모형과 관련 있는 차수.

\$a\$는 MA 모형과 관련 있는 차수.

\$p\$와 \$q\$는 ARIMA 모형에서 ARMA로 정상화 할 때 차분한 횟수를 의미

종류	의미
\$p=0\$	\$IMA(d,q)\$ 모형이라고 부르고, \$d\$번 차분하면 \$MA(q)\$ 모형을 따른다
\$d=0\$	\$ARMA(p,q)\$ 모형이라 부르고 , 이모형은 정상성을 만족한다.
\$q=0\$	\$ARI(p,d)\$ 모형이라 부르고, \$d\$번 차분하면 \$AR(p)\$ 모형을 따른다.