

시계열 분석 및 실습 개요

이상조

September 1, 2025

시계열 자료란?

- ◆ 연도별, 계절별, 월별, 일별, 시, 분, 초, tick 별로 시간의 흐름에 따라 관측된 자료
- ◆ 재현 실험이 불가능해 과거 자료가 유일한 정보임
- ◆ 자료의 시간에 따른 구조를 고려하여 분석함

- ① Economic time series: 국민 총 생산액, 물가지수, 총 수출액, 주가지수 등
- ② Physical time series: 일일 강수량, 기온, 태양의 흑점 수, 연도별 지진의 발생 수 등
- ③ Marketing time series: 판매량, 광고액, 재고량, 매출액 등
- ④ Demographic time series: 총 인구, 농가 수, 인구 증가율, 평균 결혼 연령 등

- ⑤ Process control time series: 품질 관리, 생산 관리 등
- ⑥ Engineering time series: (0,1)-확률과정, 음성파 등
- ⑦ Medical/Healthcare time series: 심전도, 뇌파, 연속 혈당 모니터링 등
- ⑧ Social life time series: 월별 교통사고 건수, 월별 범죄 발생 수 등
- ⑨ Mobility/Network time series: GPS 이동 궤적, 교통량, 인터넷 트래픽, 사이버 공격 탐지 등

시계열 자료의 구분

- ◆ 연속 시계열(continuous time series): 모든 순간이 관측 대상
- ◆ 이산 시계열(discrete time series): 특정 시점에서만 or 특정 간격으로만 관측
증가
- ◆ 시차(time lag): 관측 시점들 사이의 간격

시계열 자료의 분석 목적

- ◆ 미래에 대한 예측(forecast)
- ◆ 시스템 또는 확률과정의 이해와 제어(control)

- ◆ 시계열그림(time series plot): 시간의 경과에 따라 시계열 자료의 값이 변하는 것을 그린 그림

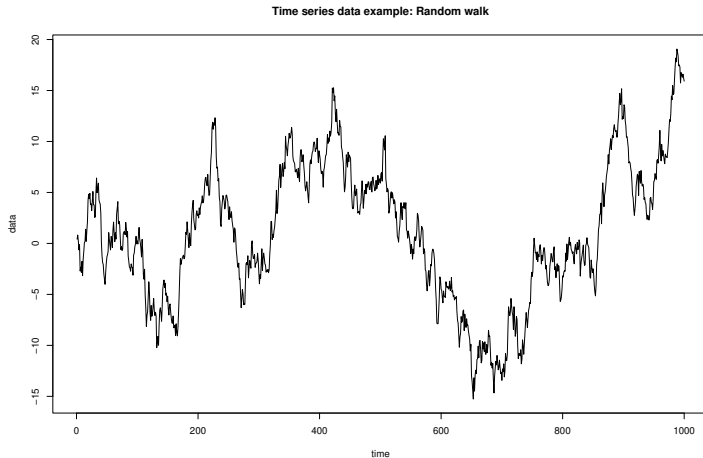


Figure: Random Walk 그림

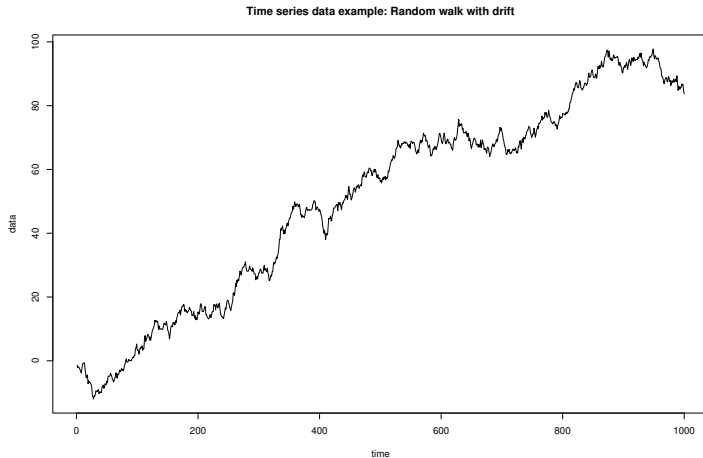


Figure: Random Walk with drift

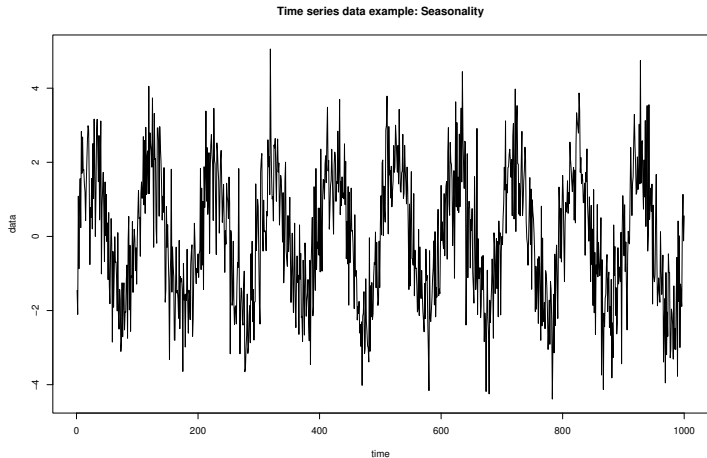


Figure: Seasonal time series

시계열 자료의 성분

- ◆ 불규칙 성분(irregular component)
- ◆ 체계적 성분(systematic component)
- ◆ 각 성분을 분리하여 분석하면 패턴을 더 명확히 이해 가능

- ◆ 시간에 따른 규칙적인 움직임이 아닌, 랜덤한 원인에 의한 변동 성분

- ◆ 추세 성분(trend component): 관측값이 지속적으로 증가하거나 감소하는 추세를 갖는 경우
- ◆ 계절 성분(seasonal component): 계절의 변화(주별, 월별, 계절별, ...)에 따른 주기적인 변동
- ◆ 순환 성분(cyclical component): 주기적인 변화를 갖지만 계절에 의한 것이 아니며 주기가 긴 경우

성분 결합

- ◆ 각 시계열 자료는 추세 및 계절 성분등 여러 성분을 동시에 포함 가능
- ◆ $Y_t = I_t + T_t + S_t + C_t$

곱한게 나올수도있는데 어려워서 안할

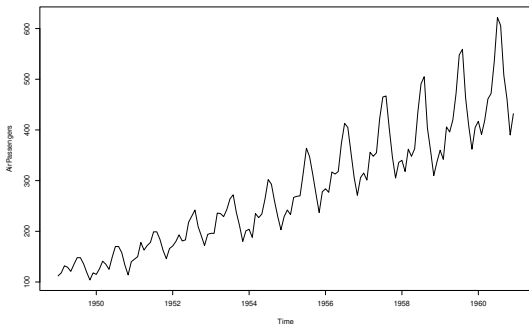


Figure: Air Passengers plot

시간에 따른 변동의 폭 고려

- ◆ 시간에 따라 변동폭이 변화하는 경우: 분산 안정화(로그 또는 제곱근 등)

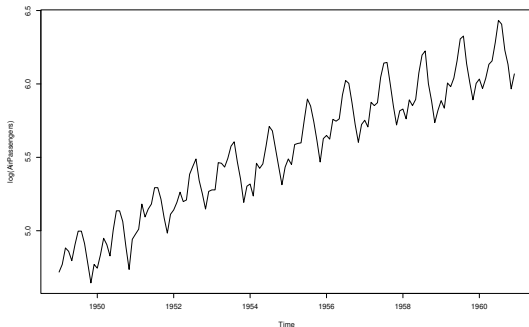


Figure: Log-transformed Air Passengers plot

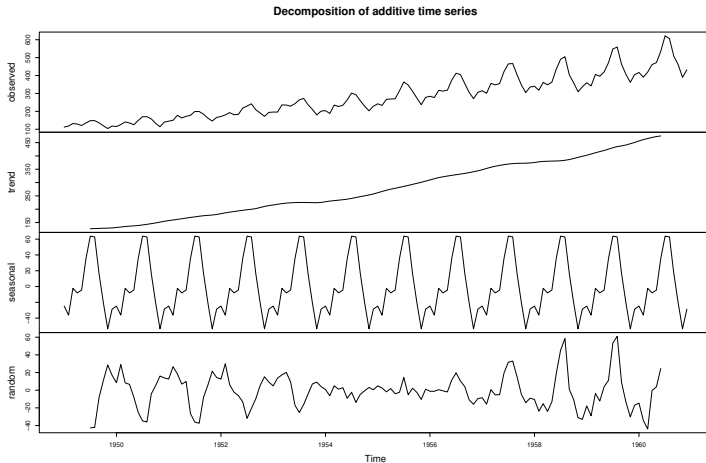


Figure: Air Passengers plot (decomposed)

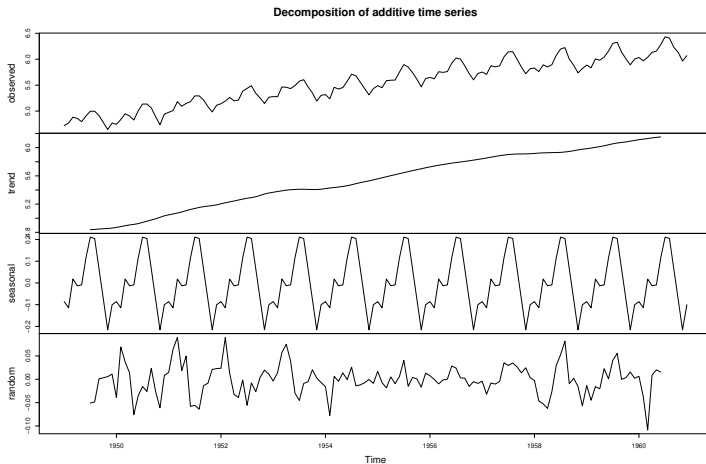


Figure: Log-transformed Air Passengers plot (decomposed)

시계열 그림

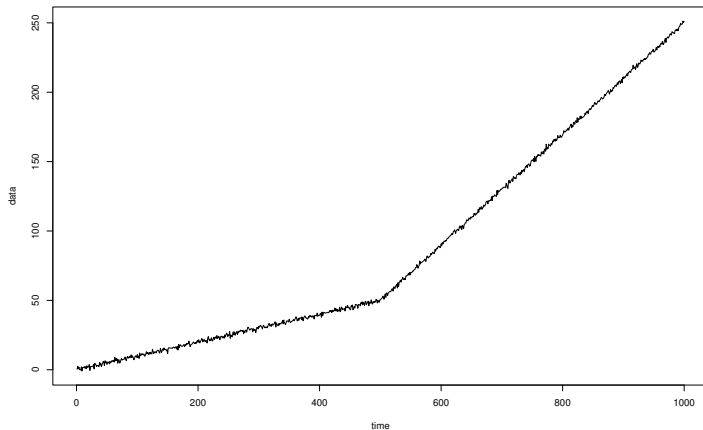


Figure: Time series with two trend lines

예측(forecast) 중심의 분석방법

- ◆ 추세분석(trend analysis)
- ◆ 평활법(smoothing method)
- ◆ 분해법(decomposition method)
- ◆ 자기회귀누적이동평균(AutoRegressive Integrated Moving Average: ARIMA)
- ◆ 이분산자기회귀모형((Generalized) Autoregressive Conditional Heteroskedasticity: (G)ARCH)

이해와 제어 중심의 분석 방법

- ◆ 스펙트럼분석(spectral analysis)
- ◆ 개입분석(intervention analysis)
- ◆ 전이함수모형(transfer function model)
- ◆ 자기상관오차를 갖는 회귀모형(autoregressive error model)

진동수 영역(frequency domain)에서의 분석법

- ◆ 푸리에분석(Fourier analysis)
- ◆ 스펙트럼 밀도(spectral density function) 추정
- ◆ 파형 분석(wavelet analysis)

시간 영역(time domain)에서의 분석법

- ◆ (부분)자기상관함수 ((partial) autocorrelation function)
- ◆ 추세분석, 평활법, 분해법, ARIMA, 전이함수모형 등
- ◆ 이분산모형(GARCH), long memory process 등

- ◆ 결정모형(deterministic): 확률적 요소 없음,

$$y = f(x_1, \dots, x_p; \beta_1, \dots, \beta_m)$$

- ◆ 확률모형(stochastic): 오차 포함,

$$y = f(x_1, \dots, x_p; \beta_1, \dots, \beta_m) + \varepsilon$$

- f , β 는 과거자료로 추정, ε : i.i.d. 확률변수

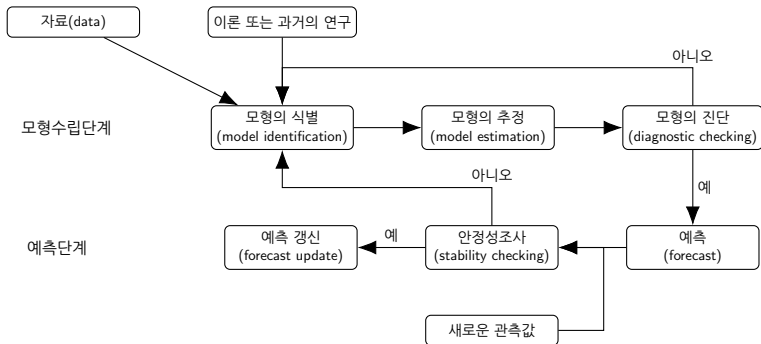
예측모형의 종류

- ◆ 주관적(qualitative / subjective) 모형
- ◆ 객관적(quantitative / objective) 모형

모형적합의 3단계

- ① 식별(identification): 시계열그림, ACF, PACF, 차분
- ② 추정(estimation): OLS, MLE, 비선형최소제곱 등으로 모수 추정
- ③ 진단(diagnostic checking): 잔차 ACF/PACF, Ljung-Box, 과대적합 점검, AIC/SBC 비교

◆ 모형적합식별→추정→진단) ⇒ 예측 ⇒ 안정성 점검/예측 갱신



예측방법의 선택기준

- ◆ 고려: 정확성, 예측기간(forecast horizon), 복잡성, 이용가능 데이터
- ◆ t 시점 기준 n -step 후의 예측: $\hat{y}_t(h)$
- ◆ 오차: $\hat{e}_{t+1} = y_{t+1} - \hat{y}_t(1)$

예측방법의 선택기준

- ◆ 평균제곱예측오차(mean squared prediction error, MSE):

$$\frac{1}{m} \sum_{t=1}^m \hat{\epsilon}_{n-1+t}(1)^2$$

- ◆ 평균절대백분위예측오차(mean absolute percentage prediction error, MAPE):

$$\frac{100}{m} \sum_{t=1}^m \left| \frac{\hat{\epsilon}_{n-1+t}(1)}{y_{n+t}} \right|$$

- ◆ 평균절대예측오차(mean absolute prediction error, MAE):

$$\frac{1}{m} \sum_{t=1}^m |\hat{\epsilon}_{n-1+t}(1)|$$

시계열의 역사 (1) : 푸리에와 웨이블릿

- ◆ 태양 흑점, 밀 가격지수에서 주기 탐색; Fourier(1807), Stokes(1879)
- ◆ Schuster(1906), Beveridge(1922): 주기도(Periodogram)
- ◆ Bartlett(1948,1950), Blackman(1958): 평활
- ◆ Wavelets: Donoho & Johnstone(1994)

시계열의 역사 (2) : 평활법과 계절조정

- ◆ Holt(1957), Brown(1959), Holt 등(1960), Winters(1960)
- ◆ X-11(1967), X-11 ARIMA(Dagum, 1975), X-12 ARIMA(Findley 등, 1998)

시계열의 역사 (3) : ARMA와 확장

- ◆ Yule(1926), Walker(1931): AR 개념; Slutsky(1937): MA 개념; Wold(1954): 분해정리
- ◆ Box-Jenkins(1970): ARIMA 3단계 방법론; Box-Tiao(1975): 개입분석; 이상점 탐지(Chang 등, 1988)
- ◆ 상태공간/칼만필터: Kalman(1960) 외; 베이지 예측: Harrison-Stevens(1971,1976)
- ◆ 단위근/공적분: Dickey-Fuller(1979), Phillips-Perron(1988), Engle-Granger(1987)
- ◆ 이분산 모형: ARCH/GARCH (Engle, 1982; Bollerslev, 1986)