

시계열 자료 분석

시계열 자료

- **정의** : 시간의 흐름에 따라 관측된 자료
- **종류**
 - 연속형 시계열 (Continuous Time Series)
 - 이산형 시계열 (Discrete Time Series)
 - 우리가 접하는 많은 시계열들은 실제로는 연속적으로 생성되고 있으나, 일정한 시간간격을 두고 관측되므로 이산 시계열자료의 형태를 지니고 있다.
- $\{X_t : t = 1, 2, \dots\}$ or X_1, X_2, \dots
- 시차(time lag) : 관측시점들 사이의 시간 간격

시계열 자료의 예

- Economic time series (경제시계열)
 - 국민 총 생산액, 물가지수, 총 수출액, 주가지수 등
- Physical time series (물리적시계열)
 - 일일 강수량, 기온, 태양의 흑점 수, 연간 지진의 발생 수 등
- Marketing time series (경영시계열)
 - 품판매량, 상품 광고액, 상품 재고량, 상품 매출액 등
- Demographic time series (인구시계열)
 - 총인구, 농가 수, 인구증가율, 평균결혼연령 등
- (0,1)-확률과정, 음성파와 같이 통신공학 또는 전자공학과 관련된 시계열
- 월별 교통사고 건수, 월별 범죄발생 수와 같이 사회생활과 관련된 시계열

시계열 자료의 분석 목적

■ 예측 (Prediction/Forecasting)

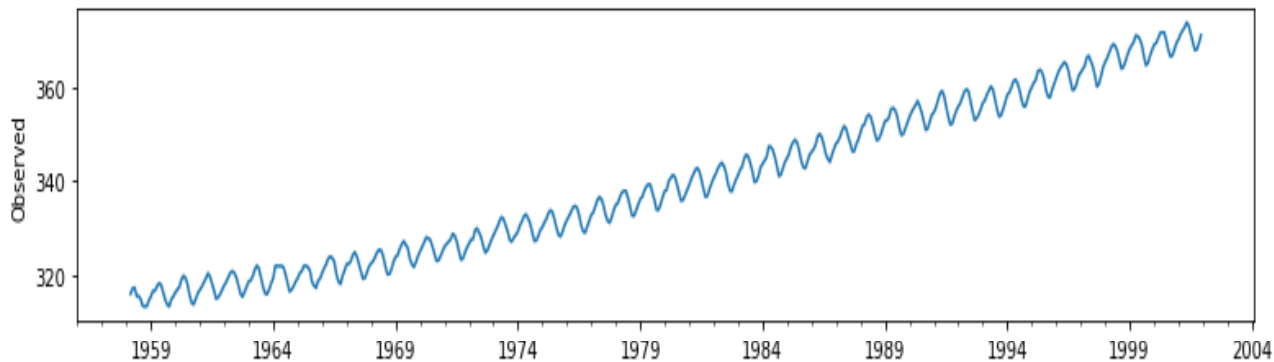
- 과거 시계열 자료의 패턴이 미래에도 지속적으로 유지된다는 가정 하에서 현재까지 수집된 자료들을 분석하여 미래 시점의 값을 추정
- 추세분석
- 평활법(Smoothing Method)
- 분해법(Decomposition Method)
- ARIMA(Auto-Regressive Integrated Moving Average)

■ 제어 (Control)

- 관심 변수 값을 특정 목표 값(Target Value)에 유지
- 스펙트럼 분석 (Spectral Analysis)
- 개입분석 (Intervention Analysis)
- 전이함수 모형 (Transfer Function Model)

■ 시계열 그림 (Time Series plot)

- 시계열 자료의 분석의 시작 : 시계열 그림 그리기



- 시간의 경과에 따라 시계열 값이 변하는 것을 그린 그림
 - : 가로축 - 시간(t)
 - : 세로축 - 관측값 X_t
- 시계열자료의 특성 파악

■ 분해법

- 불규칙 성분 (Irregular component)
- 체계적 성분 (Systematic component)
 - 1) 추세성분 (Trend component)

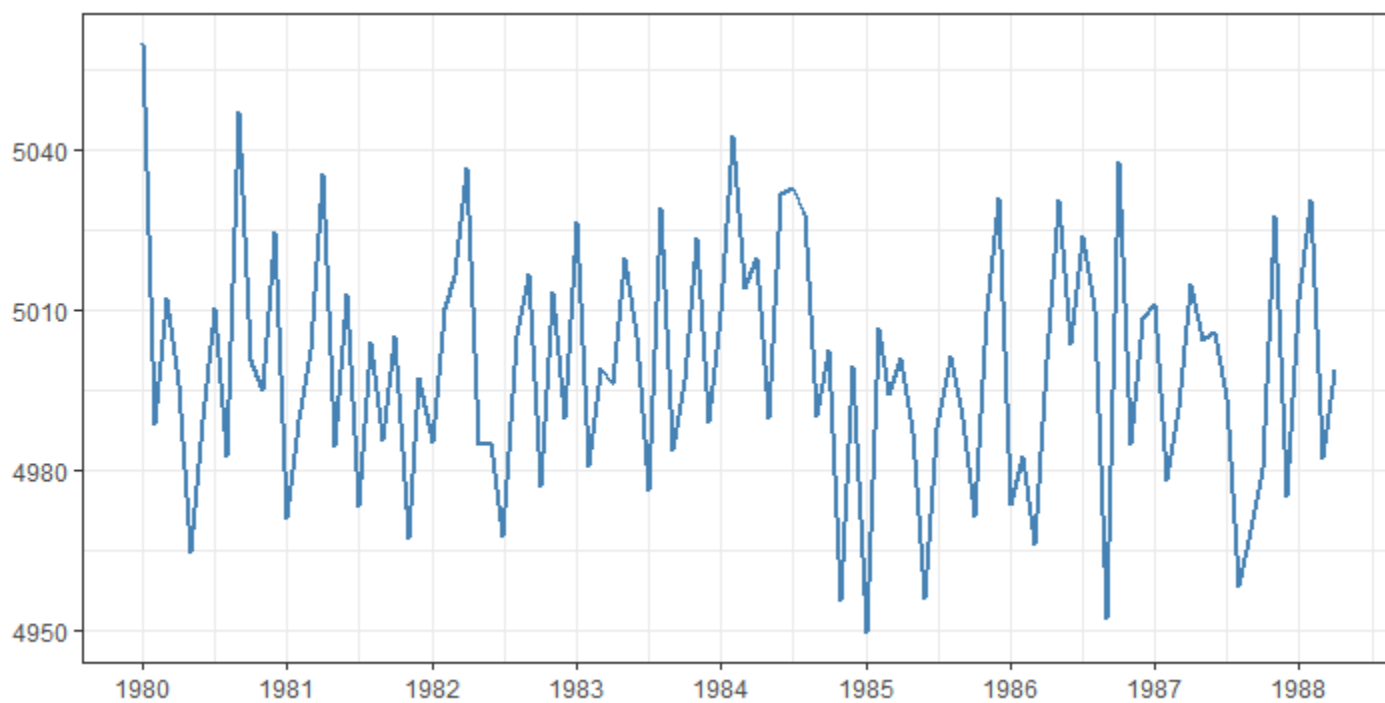
: 시간이 경과함에 따라 관측값이 지속적으로 증가하거나 감소하는 흐름
 - 2) 계절성분 (seasonal component)

: 주별, 월별, 계절별과 같이 주기적인 성분에 의한 변동
 - 3) 순환성분 (cyclical component)

: 계절성분에 비해 주기가 긴 경우의 변동

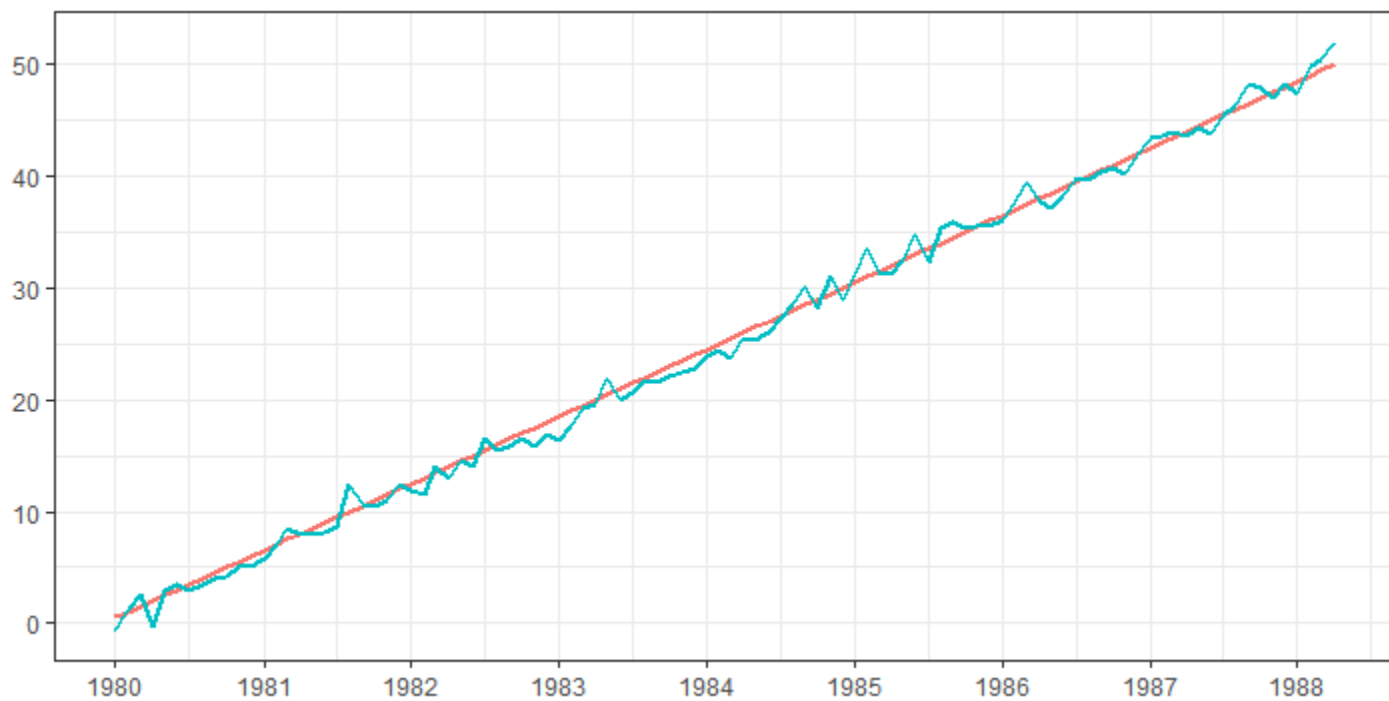
시계열 자료의 특성 - 분해법

- 불규칙 성분



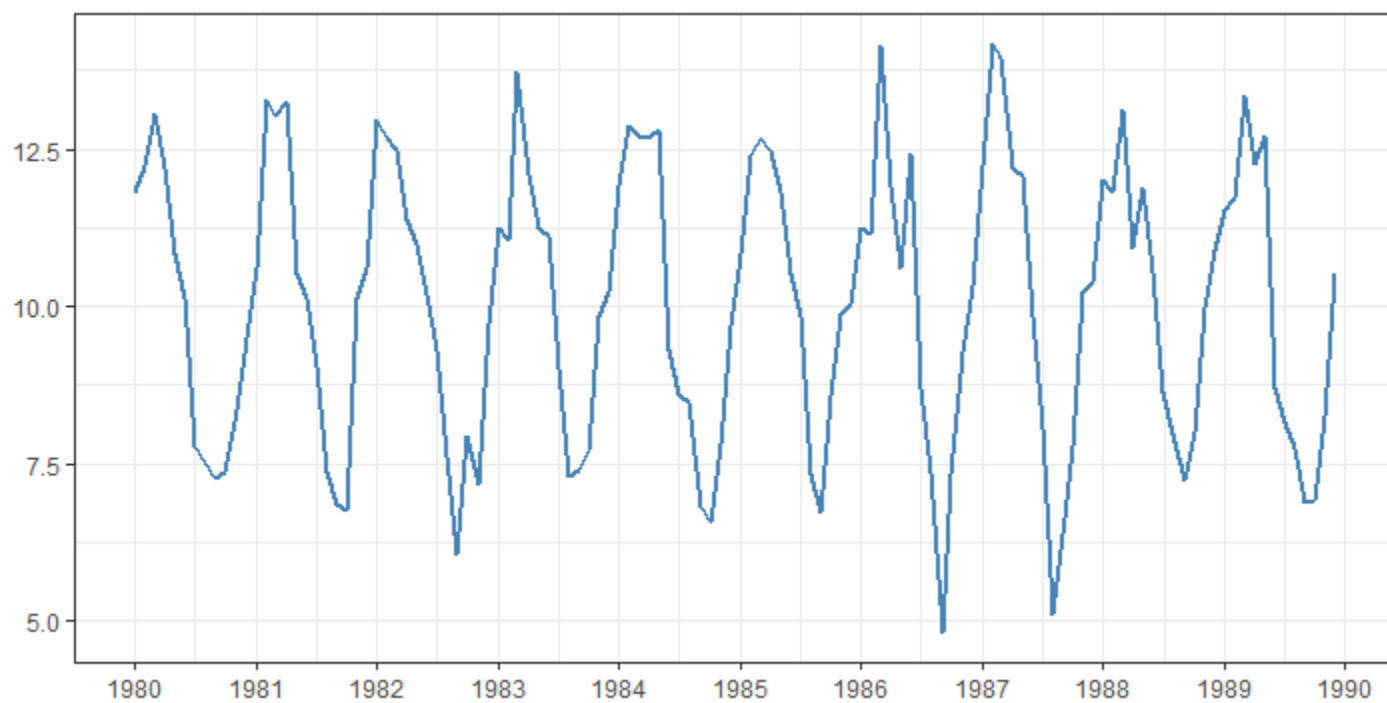
시계열 자료의 특성 - 분해법

■ 추세 성분



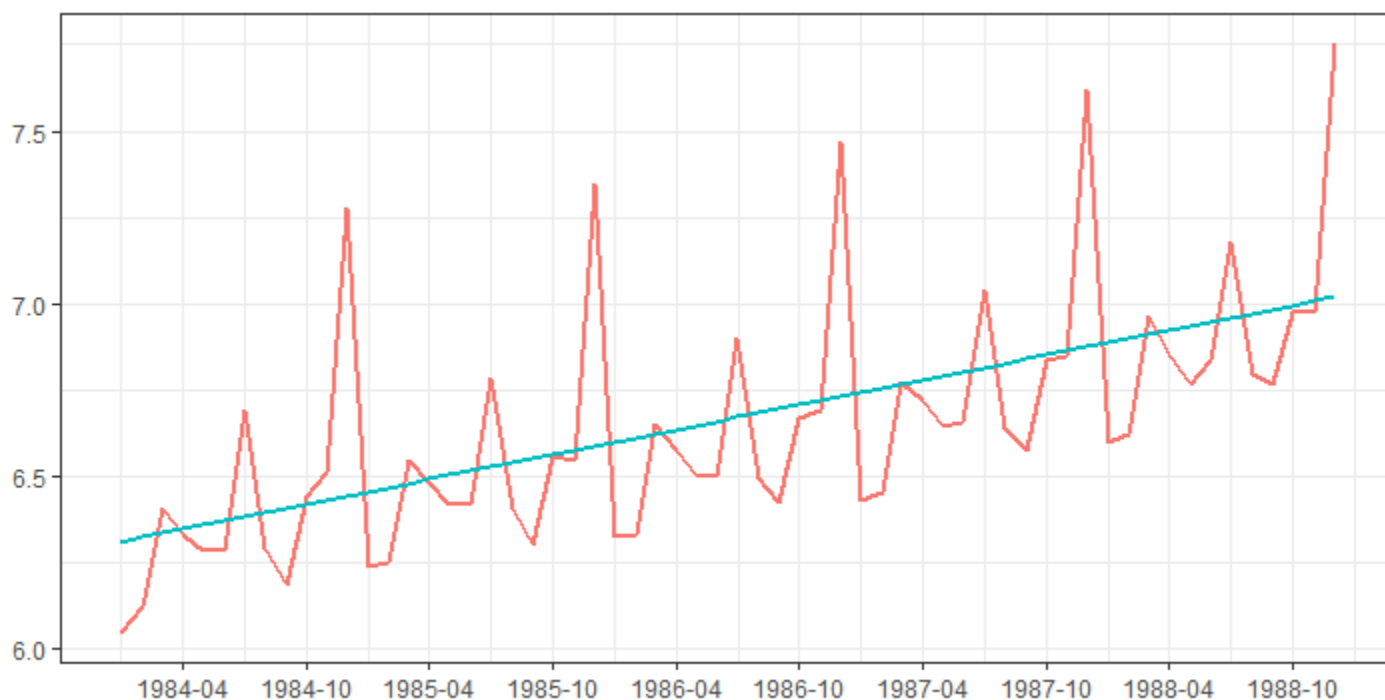
시계열 자료의 특성 - 분해법

■ 계절 성분



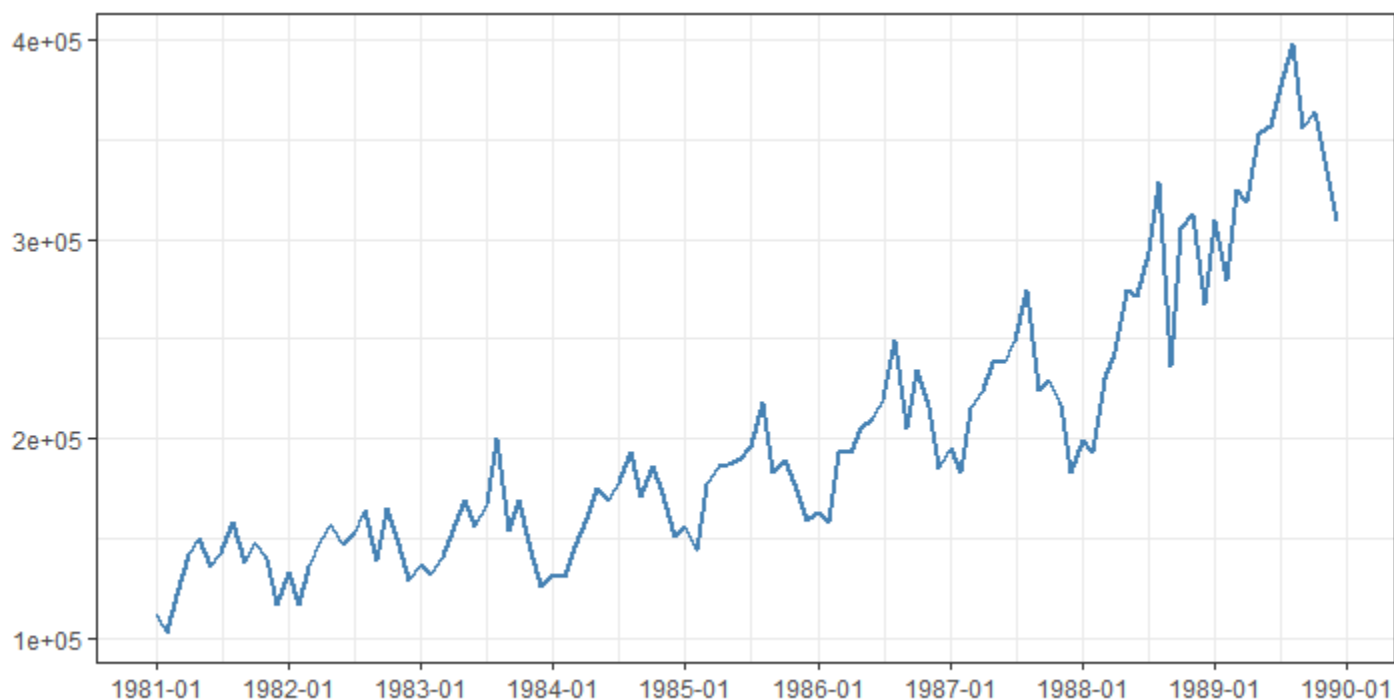
시계열 자료의 특성 - 분해법

- 추세 성분 + 계절 성분



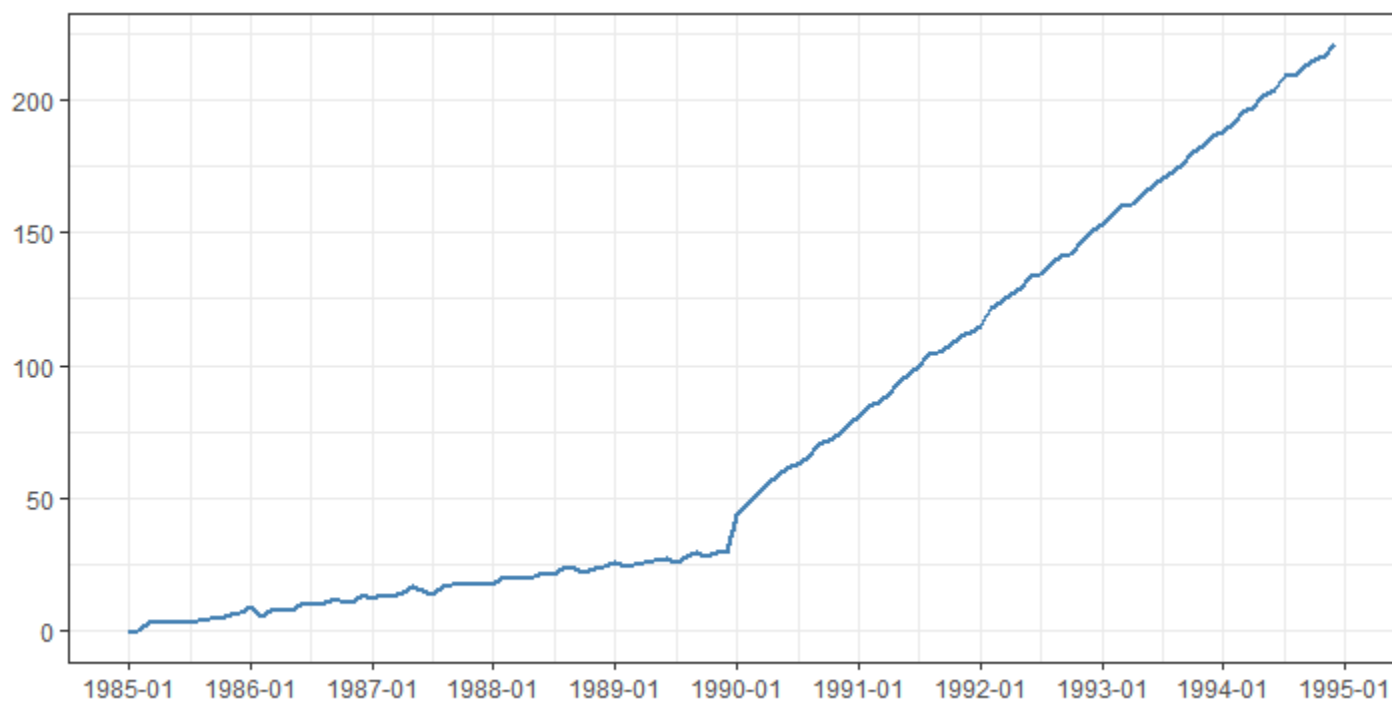
시계열 자료의 특성 - 분해법

- 추세 성분 + 계절 성분(이분산)



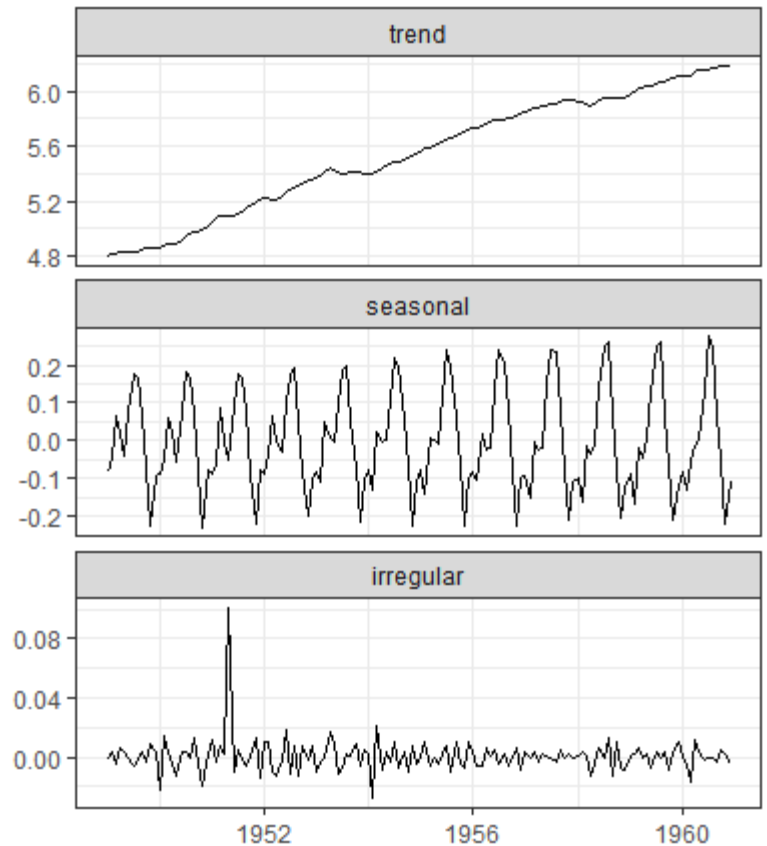
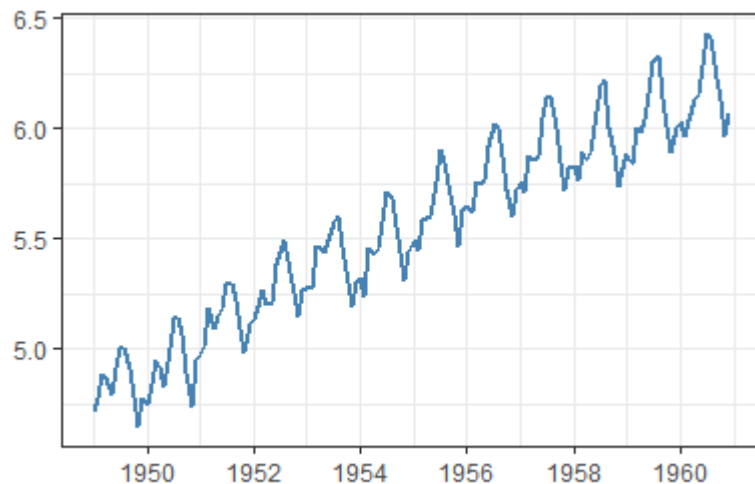
시계열 자료의 특성 - 분해법

- 추세선이 두 개인 시계열



시계열 자료의 특성 - 분해법

■ 시계열 성분 분해 - 예) AirPassengers



모형 적합의 3단계

- **모형의 식별** (model identification)

- 시계열그림, SACF, SPACF 등을 이용하여 적합시킬 모형을 잠정적으로 결정

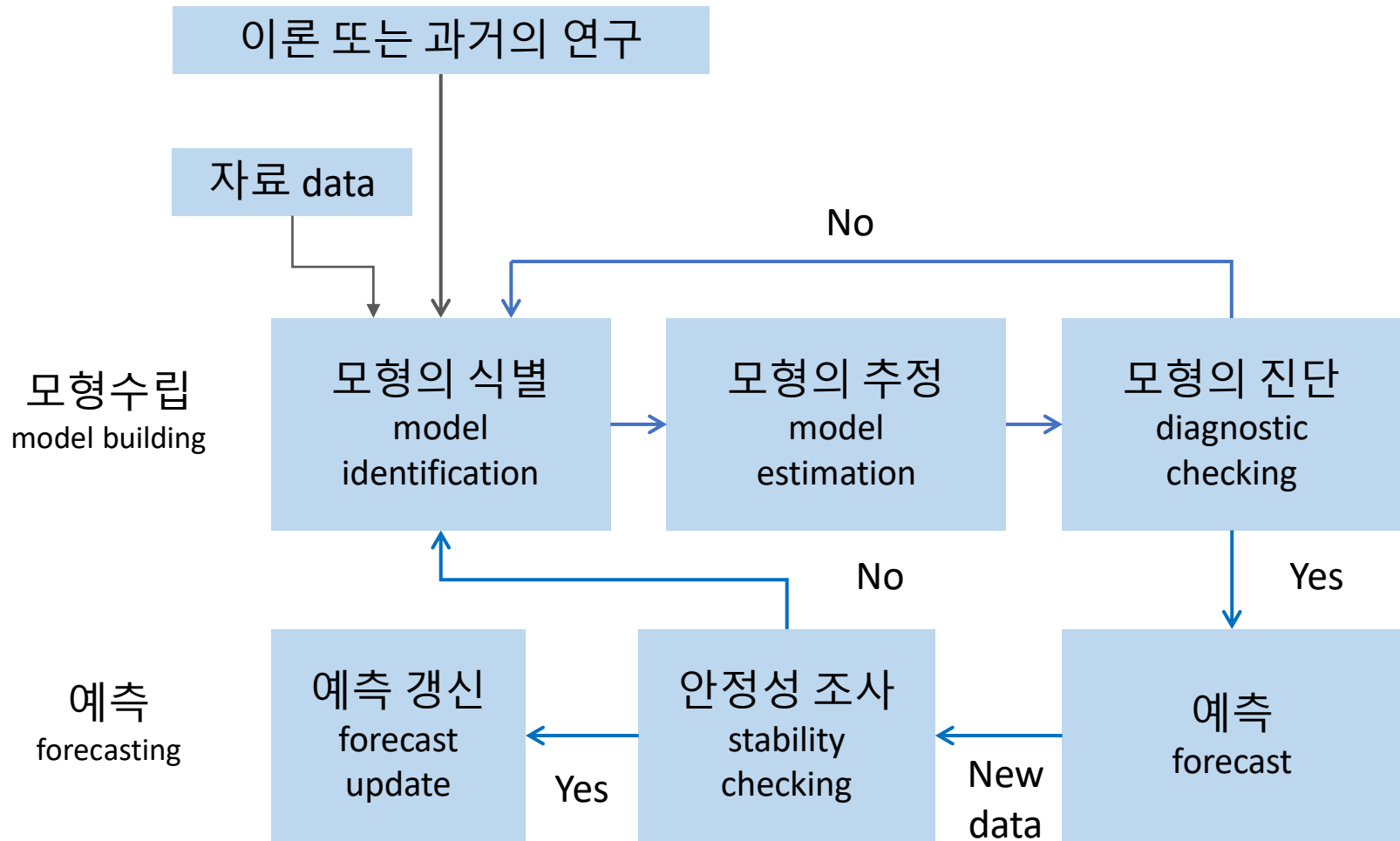
- **모형의 추정** (model estimation)

- LS method, ML method 등을 이용하여 잠정적으로 결정한 모형의 모수 추정

- **모형의 진단** (model diagnostic checking)

- 잔차 분석을 통해 모형의 적합 정도를 진단
- AIC, SBC 등을 이용하여 설명력이 높은 모형 선택

모형 적합의 3단계



예측방법의 선택 기준

■ Factors

- 정확도
- 예측 시간에 소요되는 시간
- 예측 대상의 기간 폭 (Forecast Horizon)
 - 단기 예측 (Short-term) vs. 장기 예측 (Long-term)
- 모형의 단순성 :: The Principle of Parsimony (간결의 원칙)
- 자료의 종류, 크기

예측력을 판단하기 위한 척도

■ 예측 오차

- 관측된 자료 : X_1, X_2, \dots, X_n
- 새로운 자료 : $X_{n+1}, X_{n+2}, \dots, X_{n+m}$
- 시점 t 에서 1-시차 후의 예측값

$$\hat{X}_t(1), \quad t = n, n+1, \dots, n+m-1$$

- 1-시차 후의 예측 오차

$$e_t(1) = X_{t+1} - \hat{X}_t(1), \quad t = n, n+1, \dots, n+m-1$$

예측력을 판단하기 위한 척도

- 평균제곱오차 (mean square prediction error) :

$$MSE = \frac{1}{m} \sum_{l=1}^m e_t(1)^2$$

- 평균절대백분위예측오차 (mean absolute percentage prediction error):

$$MAPE = \frac{100}{m} \sum_{l=1}^m \left| \frac{e_t(1)}{X_{t+1}} \right|$$

- 평균절대예측오차 (mean absolute prediction error) :

$$MAE = \frac{100}{m} \sum_{l=1}^m |e_t(1)|$$

End of Document