

Chapter. 03
ARIMA

| 예측모델링 Summary

M T W T F S S

FASTCAMPUS
ONLINE

금융공학/퀀트 I

강사. 김경륜

지수이동평균법(Exponential Moving Average)

과거일정기간 데이터의 가중평균으로 미래예측

: 최근 데이터는 높은 가중치를 주어 많이 반영하고 오래전 데이터는 가중치를 낮게 줌

→ 직관적인 방법으로 널리 사용되는편이며 바로 다음시점에 대한 예측은 예측력이 비교적 높음

HoltWinters함수를 사용

1. 가장최근데이터에 대한 가중치(alpha=0.2)를 지정하는 경우

모델링: `expma1 <- HoltWinters(data, alpha = 0.2, beta=FALSE, gamma=FALSE)`

예측: `forecast(expma1, h=12)`

2. 가장최근데이터에 대한 가중치를 정하지 않고 자동으로 산출

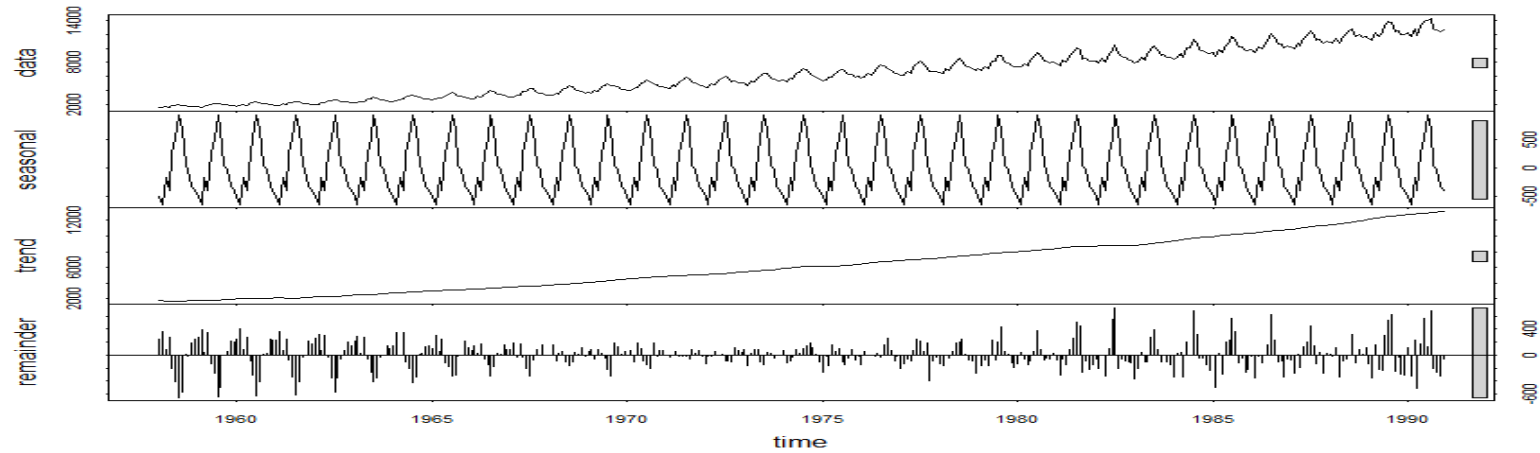
모델링: `expma2 <- HoltWinters(data, beta=FALSE, gamma=FALSE)`

예측: `forecast(expma2, h=12)`

Trend 및 Seasonality를 고려하지 않으므로 2개이상을 예측하더라도 같은 예측치가 반복

시계열분해법(Holt-Winters, STL)

시계열을 Trend/Seasonal로 분해하여 각각을 예측하여 결합하여 예측



1. Holt-Winters 방법

모델링: `example1 <- hw(data)`

예측: `forecast(example1, h=12)`

2. STL 방법

모델링: `example2 <- stl(data, s.window = "per")`

예측: `forecast(example2, h=12)`

ARIMA모형

시계열을 차분하여 정상성을 가지게 한 다음,
자기상관(AR)변수 및 백색잡음변수(MA)를 사용하여 시계열을 설명할 수 있는 모델을 산출
→ $AR(p) + MA(q) + \text{differencing}(d) = \text{ARIMA}(\text{AR integrated MA})$

1. x변수(covariates)가 없을때

모델링: `example1 <- auto.arima(data)`

예측: `forecast(example1, h=12)`

2. x변수를 사용할때

모델링: `example2 <- example1 <- auto.arima(data, xreg = x변수)`

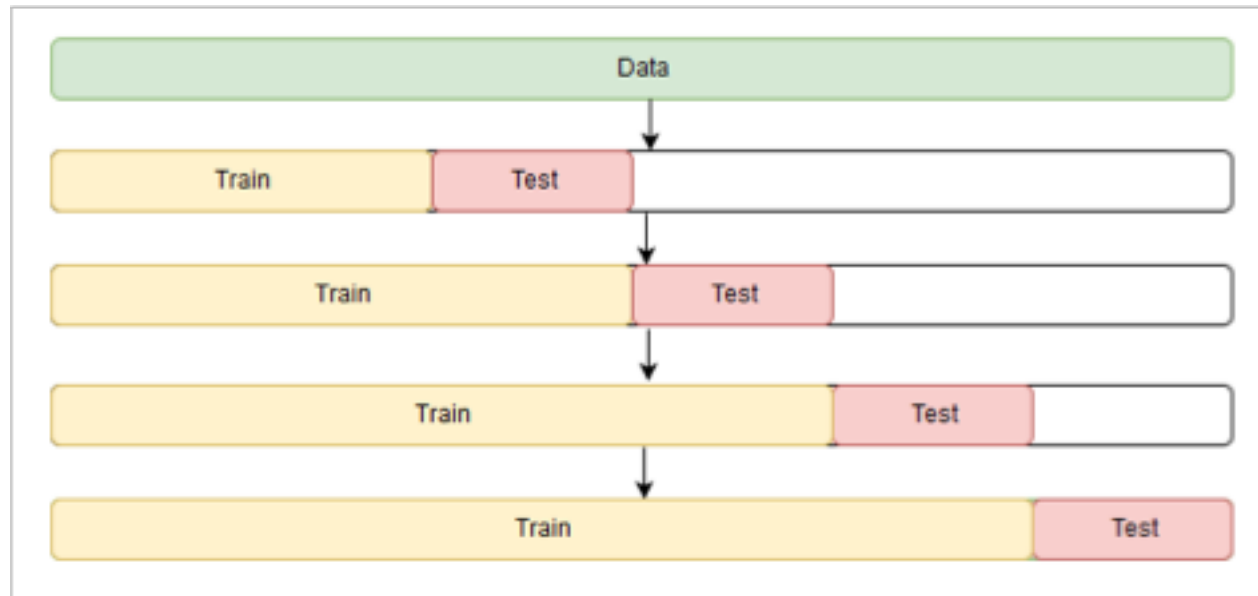
* xreg에 입력되는 데이터는 data와 길이가 같은 data.frame 형식이어야 함

예측: `forecast(example2, xreg)`

예측모형의 성능 : Cross Validation for Forecasting (시계열예측)

시계열 cross validation 방법

1. 데이터에서 정해진 크기의 test set을 추출
2. Test set 이전 데이터를 training set으로 정하고 예측모델링 진행
3. 예측결과를 test set과 비교, 검증
4. 1,2,3을 다른 시간대별로 수행



I 예측모형의 성능

모형성능 및 예측정확도 측정

$$\text{MSE} = \frac{1}{n} \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$$

MSE가 낮은 모형이 설명력이 높음(fitting)

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^n \frac{|Y_i - \hat{Y}_i|}{Y_i}$$

MAPE가 낮은 모형이 정확도가 우수(forecasting)

→ 1-MAPE = 정확도

분모: 실측치

분자: 실측치와 예측치의 차이

정확도 = 1- MAPE

mean(100 - 100 * abs(("실제값" - "예측값") / "실제값"))