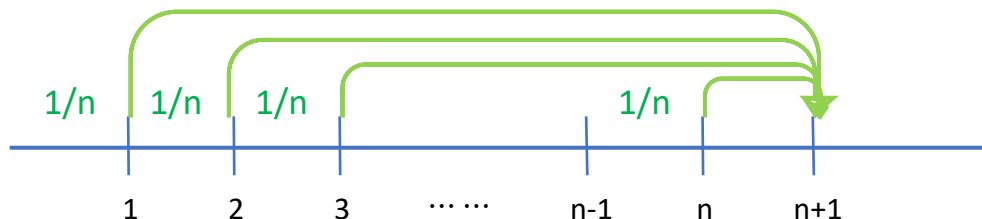


CH3. 평활법

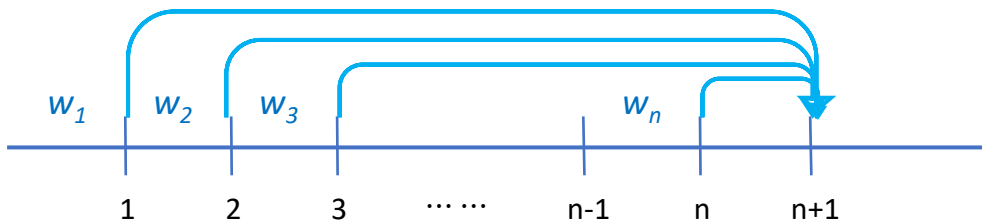
■ 상수평균모형 ($Z_t = \beta_0 + \varepsilon_t$) 에서의 예측

- $\hat{X}_n(1) = \bar{X}_n = \frac{1}{n}X_1 + \frac{1}{n}X_2 + \cdots + \frac{1}{n}X_n$



■ 평활법 (smoothing method)

- 과거의 관측값들을 가중평균 (weighted average)하여 예측값으로 이용

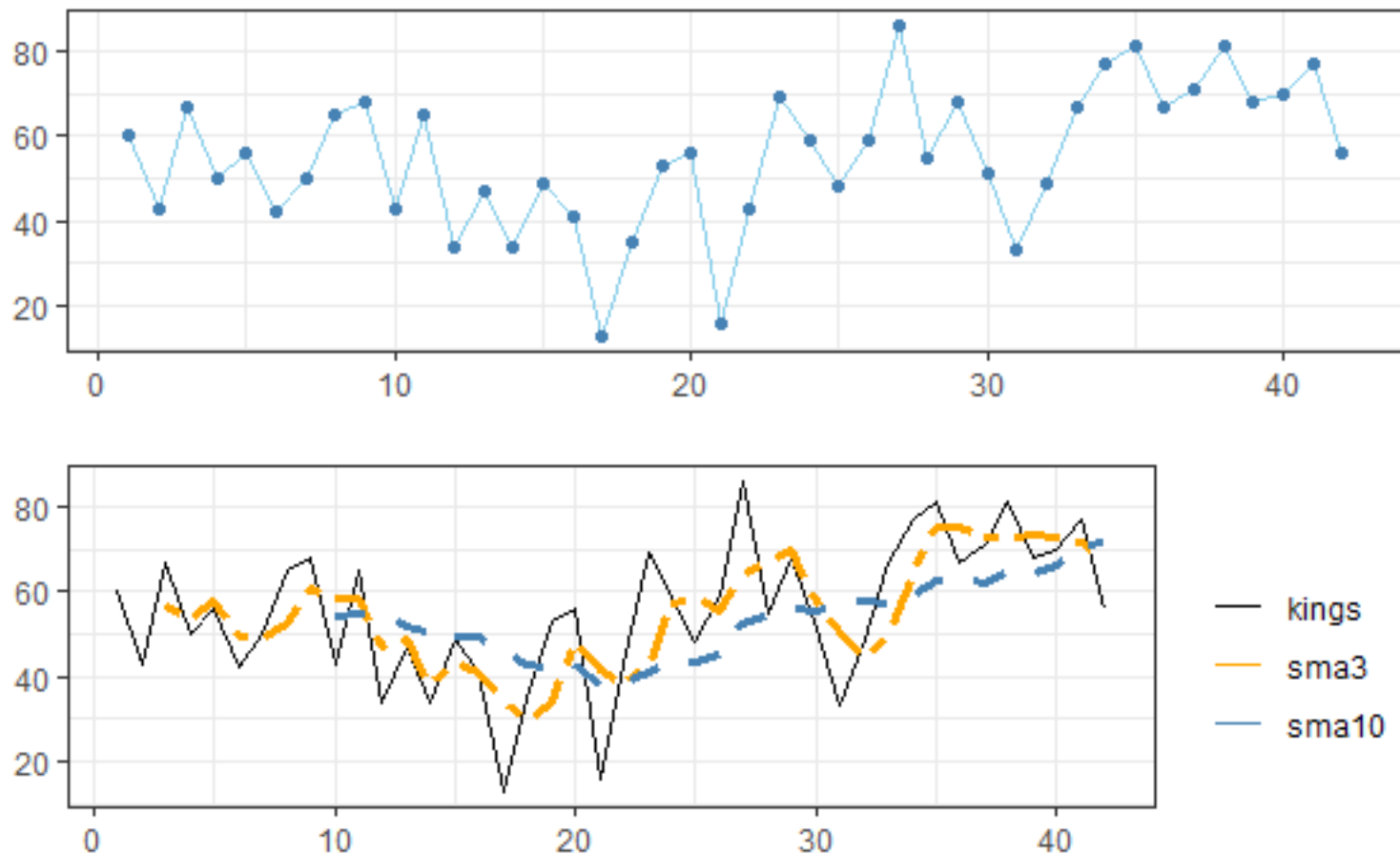


■ 이동평균법(moving average method)

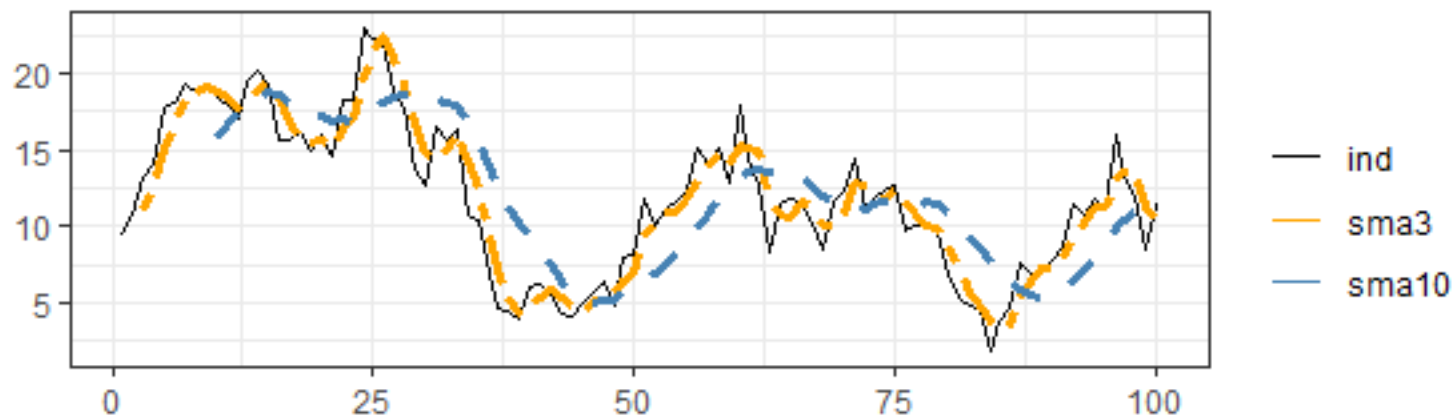
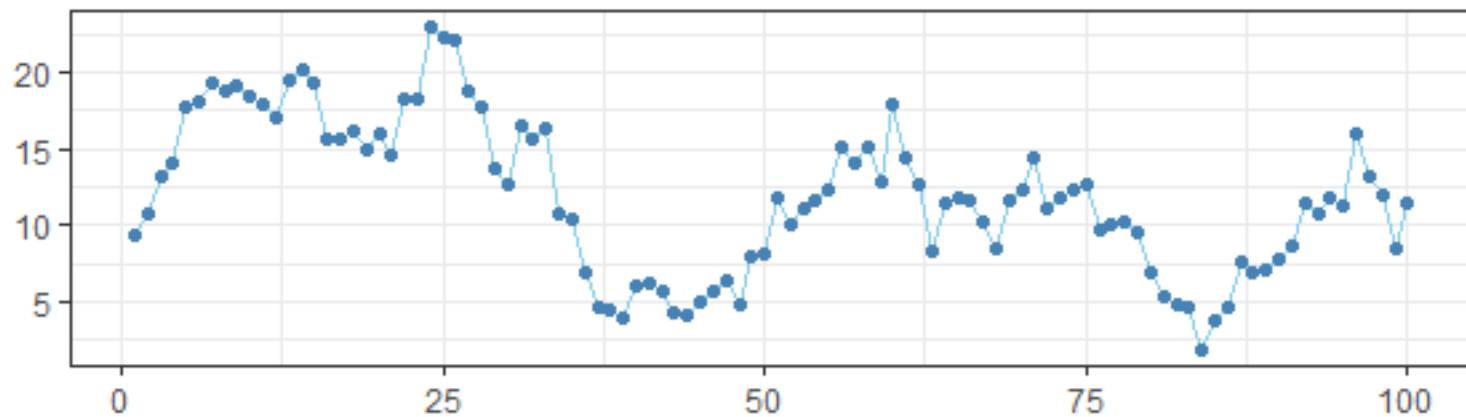
- 표본평균처럼 관측값 전부에 동일한 가중치를 주는 대신에 최근 m 개의 값들만을 이용하여 평균을 구하는 방법
- 장점 :
 - 지엽적인 변동을 제거하여 장기적인 추세 파악 가능
 - 시계열이 생성되는 시스템에 변화에 쉽게 대처 가능
 - 쉬운 계산법
- 예측의 목적보다는 주로 분해법에서 계절조정을 하는데 사용

이동평균법 - 예제

- 이동평균법 - 예제) 42명의 영국 왕들의 수명에 대한 자료



■ 이동평균법 - 예제) 중간재 출하 지수



단순지수 평활법 (simple exponential smoothing)

■ 모형

$$Z_t = \beta_{0,t} + \varepsilon_t, \quad t = 1, 2, \dots, n$$

- ε_t : 서로 독립, 평균 0, 분산 σ^2
- $\beta_{0,t}$: 시간에 따라 변하는 미지의 모수
- 상수평균 모형과의 차이점
 - 상수평균 모형 : $Z_t = \beta_0 + \varepsilon_t \Rightarrow$ 평균 수준이 시간에 상관없이 일정
 - 국지적으로는(locally) 동일한 평균 수준을 갖지만 전체적으로는(globally) 시간대별로 평균수준이 변함

■ 상수평균모형에서의 예측

- 시점 $n - 1$ 에서의 1-시차 후의 예측값

$$\hat{Z}_{n-1}(1) = \frac{1}{n-1} \sum_{t=1}^{n-1} Z_t$$

- 예측오차 : $\hat{e}_{n-1}(1) = Z_n - \hat{Z}_{n-1}(1)$
- (예측 갱신) 시점 n 에서 1-시차 후의 예측값

$$\hat{Z}_n(1) = \frac{1}{n} \sum_{t=1}^n Z_t = \hat{Z}_{n-1}(1) + \frac{1}{n} \hat{e}_{n-1}(1)$$

■ 예측

- 시점 $n - 1$ 에서의 1-시차 후의 예측값 : $\hat{Z}_{n-1}(1) = \hat{\beta}_{0,n-1}$
- 예측오차 : $\hat{e}_{n-1}(1) = Z_n - \hat{Z}_{n-1}(1) = Z_n - \hat{\beta}_{0,n-1}$
- (예측 갱신) 시점 n 에서 1-시차 후의 예측값

$$\hat{Z}_n(1) = \hat{Z}_{n-1}(1) + \omega \hat{e}_{n-1}(1)$$

$$\hat{\beta}_{0,n} = \hat{\beta}_{0,n-1} + \omega (Z_n - \hat{\beta}_{0,n-1}), \quad 0 < \omega < 1$$

■ 예측

- $\hat{\beta}_{0,n} \equiv S_n^{(1)}$ 이라고 하면

$$S_n^{(1)} = S_{n-1}^{(1)} + \omega (Z_n - S_{n-1}^{(1)}) = \omega Z_n + (1 - \omega) S_{n-1}^{(1)}$$

$$\Rightarrow S_n^{(1)} = \omega Z_n + (1 - \omega) \{ \omega Z_{n-1} + (1 - \omega) S_{n-2}^{(1)} \}$$

$$= \omega Z_n + \omega(1 - \omega) Z_{n-1} + (1 - \omega)^2 S_{n-2}^{(1)}$$

$$= \omega Z_n + \omega(1 - \omega) Z_{n-1} + (1 - \omega)^2 \{ \omega Z_{n-2} + (1 - \omega) S_{n-3}^{(1)} \}$$

...

$$= \omega Z_n + \omega(1 - \omega) Z_{n-1} + \cdots + \omega(1 - \omega)^{n-1} Z_1 + (1 - \omega)^n S_0^{(1)}$$

$$= \omega \sum_{j=0}^{n-1} (1 - \omega)^j Z_{n-j} + (1 - \omega)^n S_0^{(1)}$$

■ 단순지수 평활 통계량 (simple exponential smoothing statistic)

- $S_n^{(1)} = \omega \sum_{j=0}^{n-1} (1 - \omega)^j Z_{n-j} + (1 - \omega)^n S_0^{(1)}$

- ω : 평활상수 (smoothing constant)

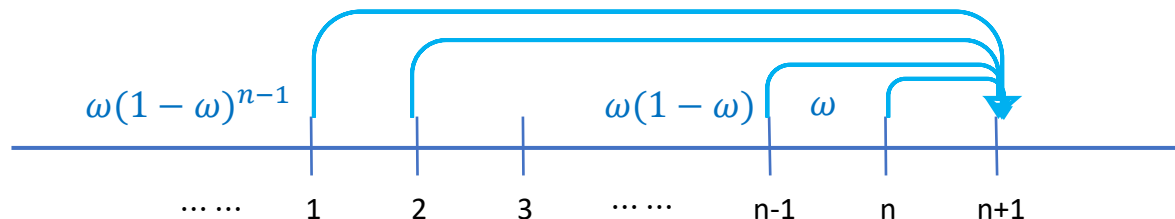
- $S_0^{(1)}$: 초기 평활값

■ $n \rightarrow \infty$ 에 따라,

- $\lim_{n \rightarrow \infty} \omega \sum_{j=0}^{n-1} (1 - \omega)^j = \frac{\omega}{1 - (1 - \omega)} = 1$

- $\lim_{n \rightarrow \infty} (1 - \omega)^n S_0^{(1)} = 0$

- $S_n^{(1)} = \omega \sum_{j=0}^{n-1} (1 - \omega)^j Z_{n-j}$



단순지수 평활 통계량 (성질)

- **지수평활 통계량** : $S_n^{(1)} = \omega \sum_{j=0}^{n-1} (1 - \omega)^j Z_{n-j}$

- 상수평균모형 ($Z_t = \beta_0 + \varepsilon_t$) 의 경우,

- $E \left(S_n^{(1)} \right) = E \left(\omega \sum_{j=0}^{n-1} (1 - \omega)^j Z_{n-j} \right)$

$$= \omega \sum_{j=0}^{n-1} (1 - \omega)^j E(Z_{n-j})$$

$$= \omega \sum_{j=0}^{n-1} (1 - \omega)^j E(\beta_0 + \varepsilon_{n-j})$$

$$= \omega \sum_{j=0}^{n-1} (1 - \omega)^j \beta_0 = \beta_0$$

- $S_n^{(1)}$ 은 평균수준 β_0 의 불편추정량

- 시점 n 에서 l -시차 후 예측값

$$\hat{Z}_n(l) = S_n^{(1)} = \hat{\beta}_{0,n}$$

- 예측 갱신

$$S_n^{(1)} = \omega Z_n + (1 - \omega) S_{n-1}^{(1)}$$

■ 초기 평활값 $S_0^{(1)}$ 의 선택

- 표본 평균 사용

- Brown (1962), Montgomery and Johnson (1976)

- $S_{0,1} = \bar{Z} = \frac{1}{n} \sum_{t=1}^n Z_t$

- 최초 관측값 사용

- Makridakis and Wheelwright (1978)

- $S_{0,1} = Z_1$

- 후향예측값 사용 (back forecsat)

- Abraham and Ledolter (1984)

- $S_{n+1}^* = Z_n, S_j^* = \omega Z_j + (1 - \omega)S_{j+1}^*, S_1^* = \omega Z_1 + (1 - \omega)S_2^*$

■ n 이 큰 경우 초기 평활값은 중요한 역할을 하지 않음

■ 평활상수의 값에 따른 평활의 양상

- ω 값이 작은 경우 : 평활 효과 지엽적인 변화에 대해 둔감하게 서서히 반응
- ω 값이 큰 경우 : 평활 효과 감소
 - 예측값이 최근 관측값에 의해 크게 영향을 받으므로, 시계열의 지엽적인 변화에 대해 민감하게 반응

■ 평활상수 ω 의 선택

- Brown (1962) : $0.05 < \omega < 0.3$
- Montgomery and Johnson (1976) : $\omega = 1 - 0.8^{(1/trend)}$

trend=평활의 차수, 단순지수평활의 경우 trend=1
- 1시차 후의 예측오차의 제곱합 $SSE(\omega) = \sum_{t=2}^n \hat{e}_{t-1}^2(1)$ 가 최소가 되도록 평활상수 선택

단순지수 평활 통계량 - 장점과 단점

■ 장점

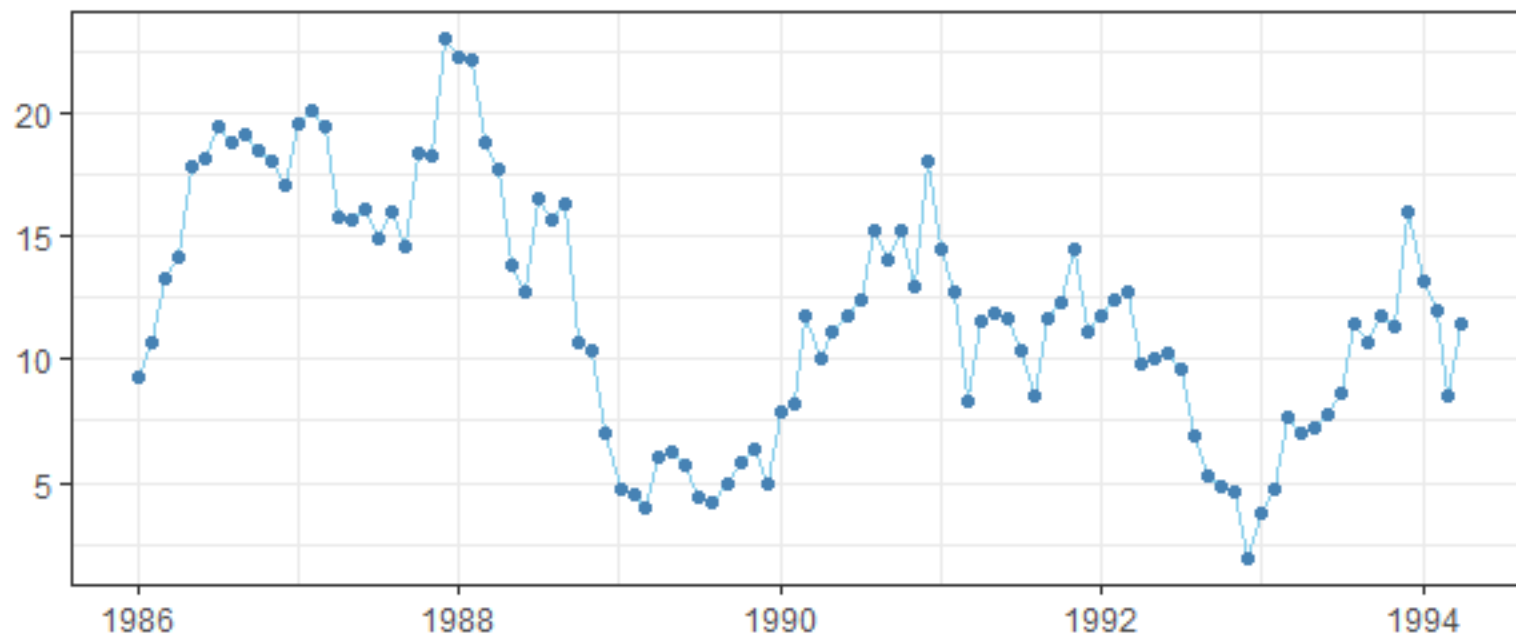
- 직관적이고 사용하기 쉽다.
- 이상점이나 개입이 존재할 경우 ARIMA 모형보다 영향을 적게 받는다(robust).
- 예측갱신이 쉽다

■ 단점

- 평활상수의 선택이 임의다.
- 특정모형 하에서만 최적이며, 예측구간을 구하기 어렵다.
- 이론적으로 미흡하므로 ARIMA 모형을 사용한 예측보다 신뢰도가 덜하다

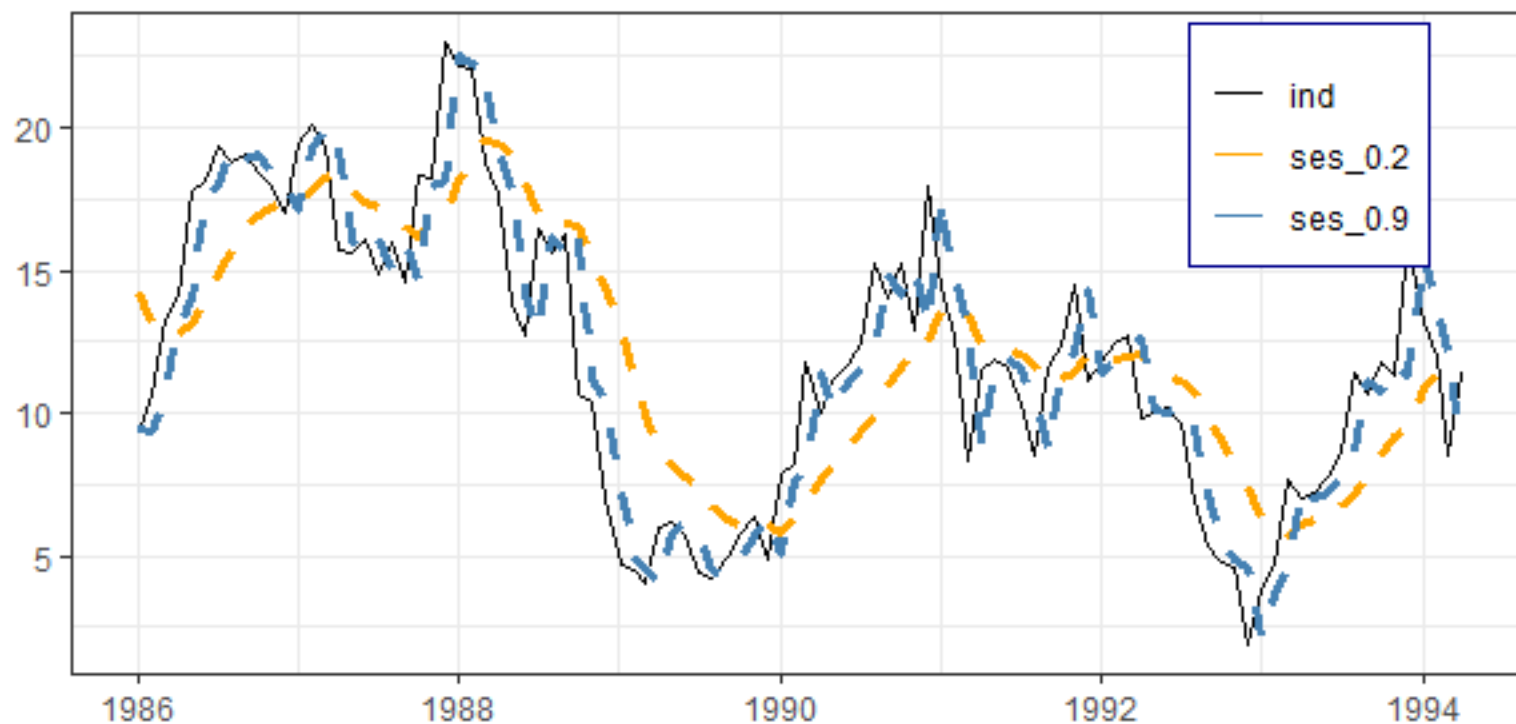
■ 중간재 출하지수 - 시계열 그림

중간재 출하지수



단순지수 평활법 - 예제

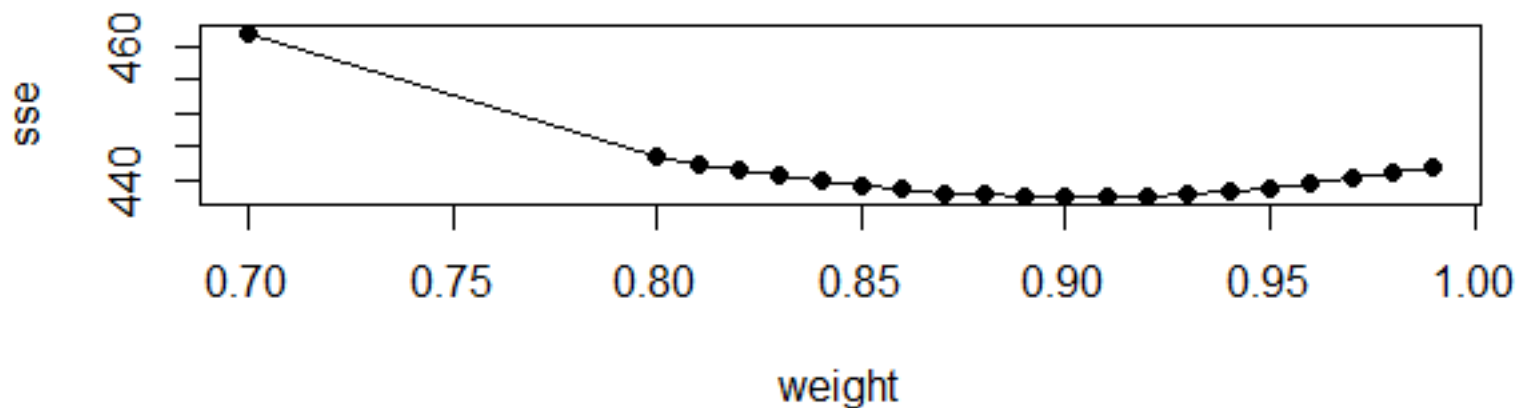
- 중간재 출하지수 – $\alpha=0.2/0.9$ 인 경우의 단순지수평활 예측값



- 중간재 출하지수 : SES = 1시차 후 예측오차의 제곱합

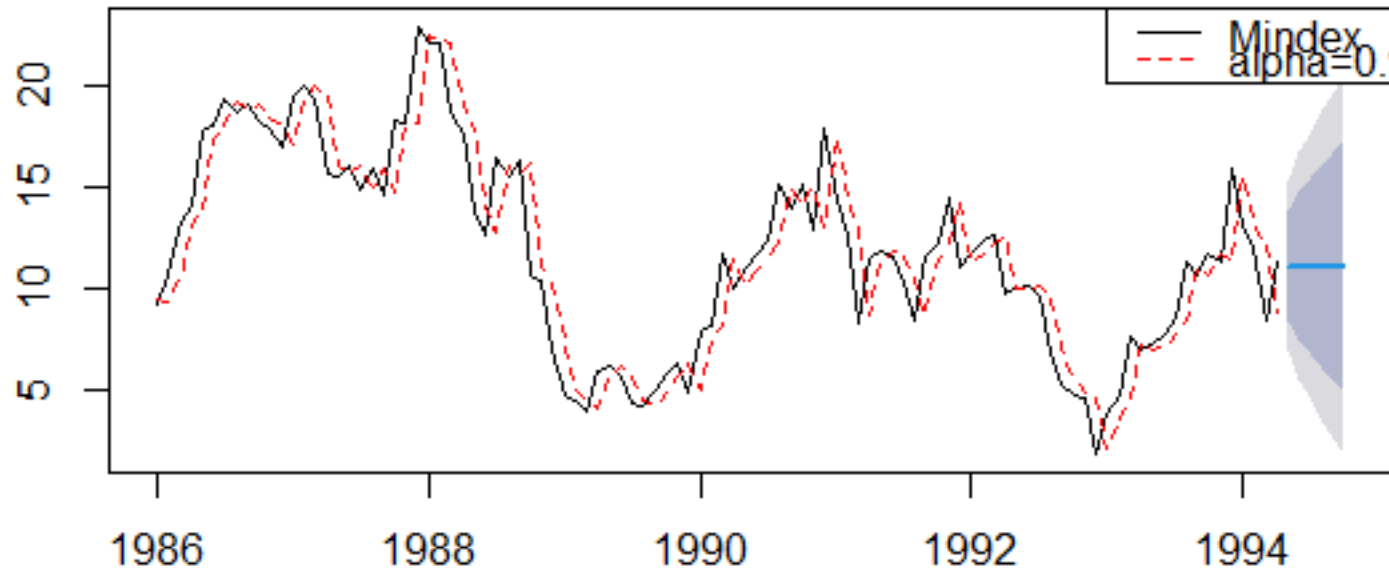
w	ses	w	ses	w	ses
0.1	1459.64	0.82	441.59	0.91	437.63
0.2	1010.03	0.83	440.69	0.92	437.77
0.3	768.28	0.84	439.90	0.93	438.03
0.4	630.73	0.85	439.23	0.94	438.40
0.5	547.18	0.86	438.68	0.95	438.90
0.6	494.62	0.87	438.24	0.96	439.51
0.7	461.95	0.88	437.91	0.97	440.25
0.8	443.75	0.89	437.70	0.98	441.12
0.81	442.61	0.9	437.61	0.99	442.11

1 시차 후 예측오차의 제곱합



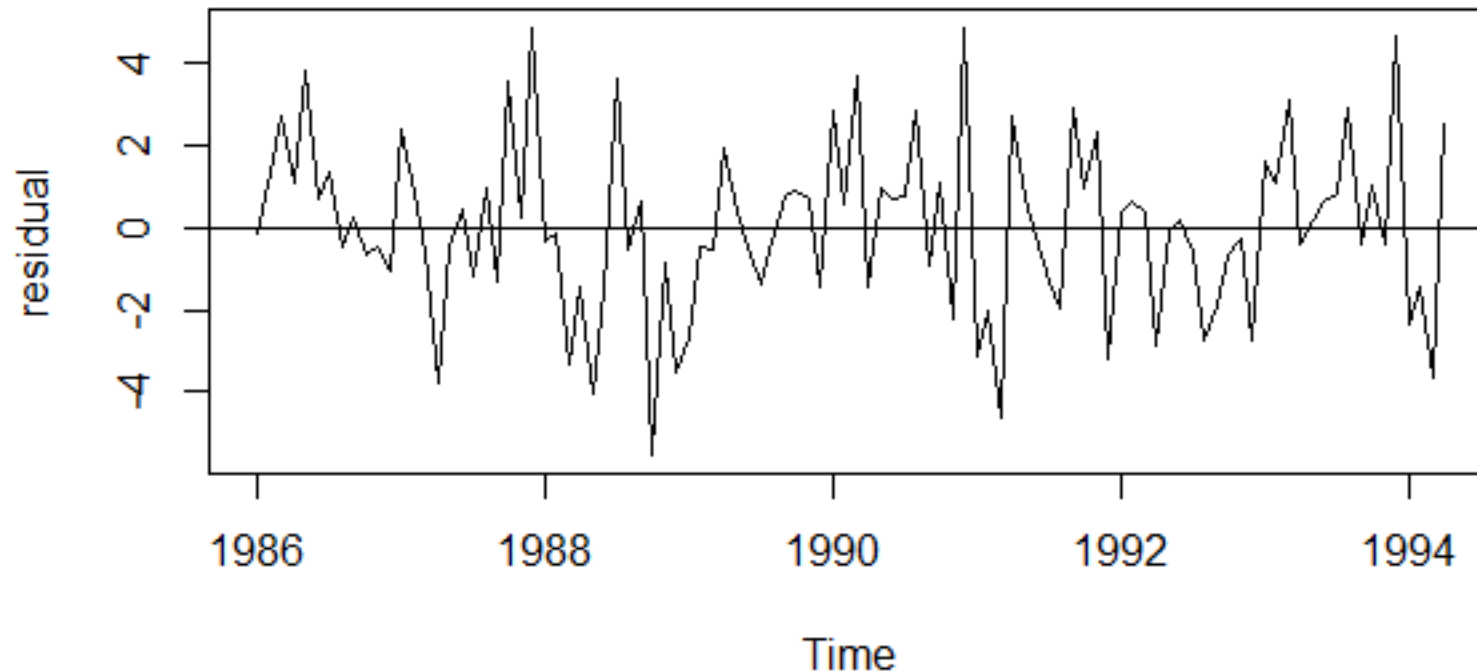
- 중간재 출하지수 - $\alpha=0.9$ 인 경우의 단순지수평활 예측값

중간재 출하지수와 단순지수평활값 $\alpha=0.9$



- 중간재 출하지수 - $\alpha=0.9$ 인 경우의 단순지수평활 예측 오차

예측 오차의 시계열그림 : $\alpha=0.9$



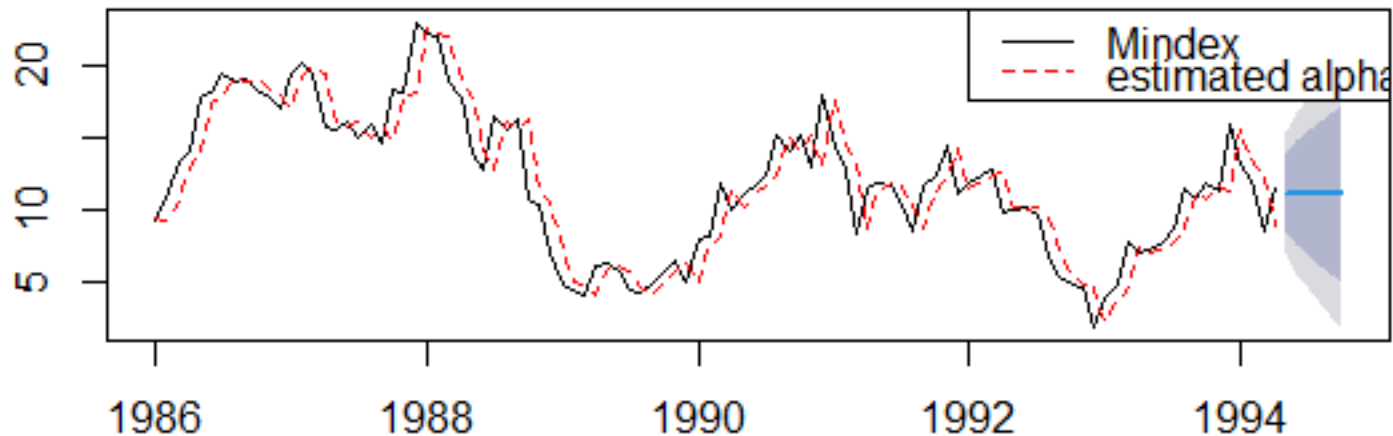
- 중간재 출하지수 - 최적의 α 를 이용한 모형 적합

simple exponential smoothing

call:

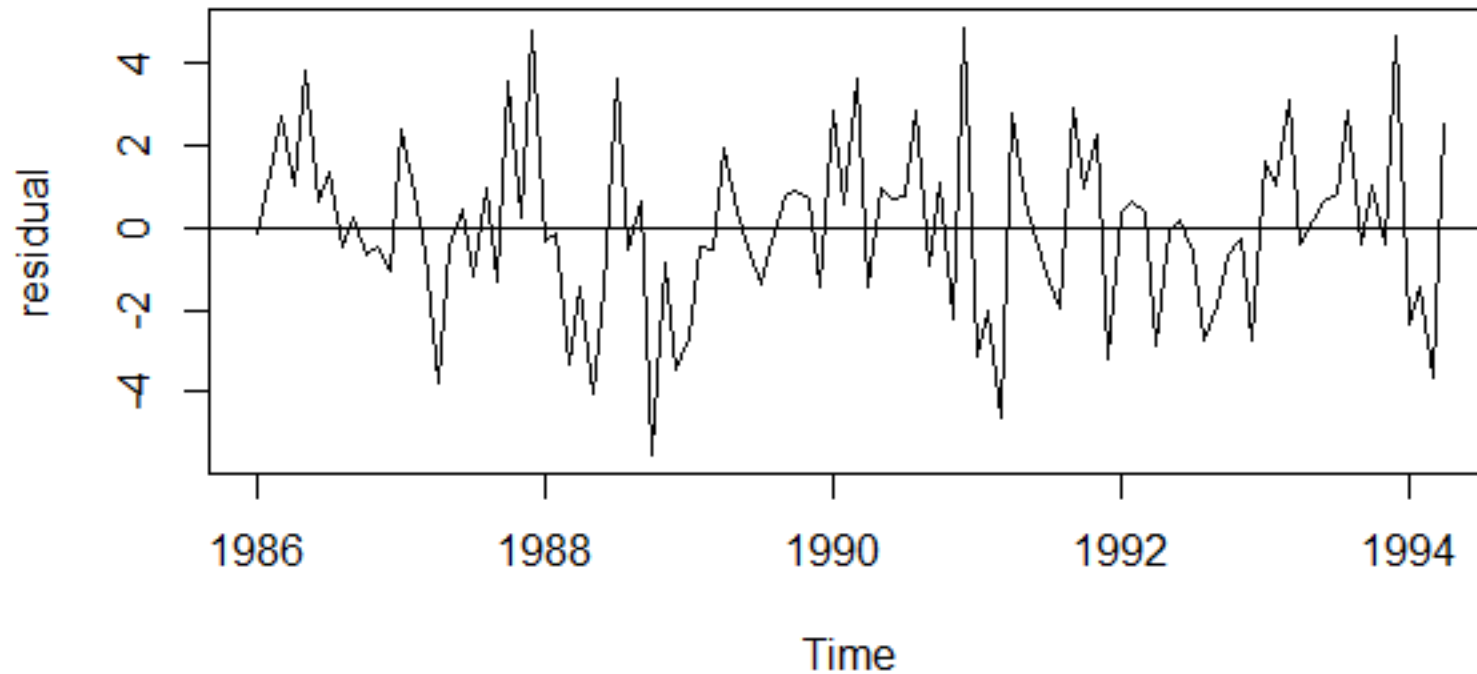
```
ses(y = mindex, h = 6)
```

Smoothing parameters:
 $\alpha = 0.9031$



- 중간재 출하지수 - 최적의 α 를 이용한 단순지수평활 예측 오차

예측 오차의 시계열그림 : 추정된 α



■ 연간 강우량 (1813-1912)

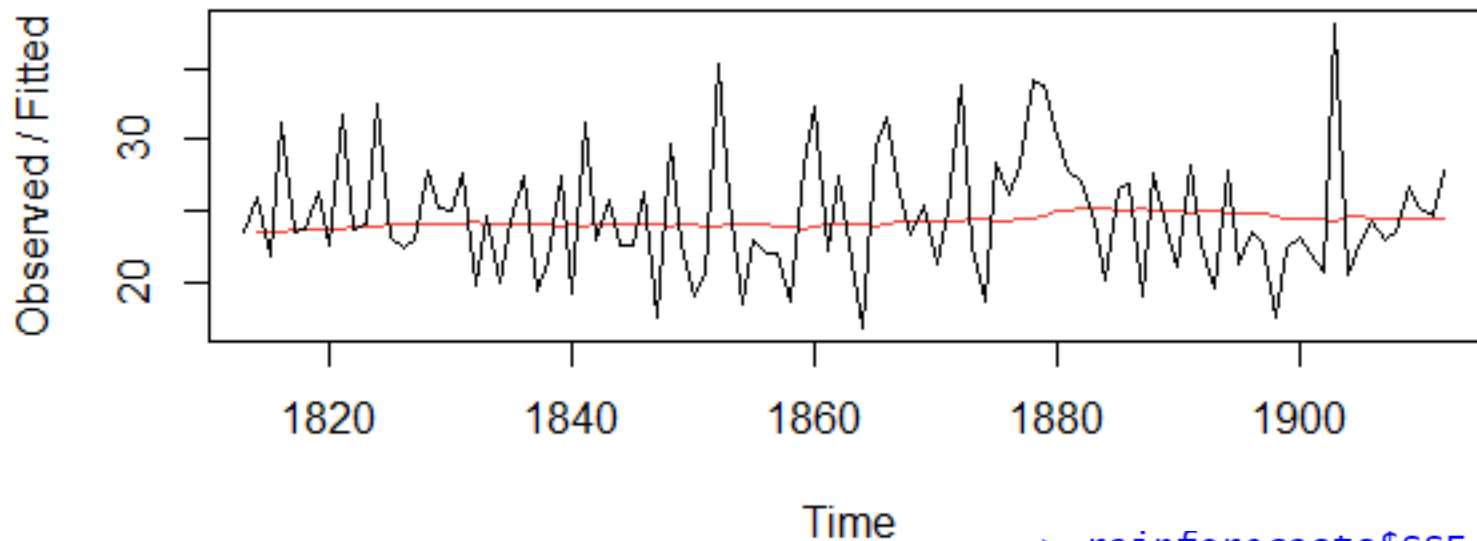
```
Holtwinters(x = rainseries, beta = FALSE, gamma = FALSE)
```

Smoothing parameters:

alpha: 0.02412151

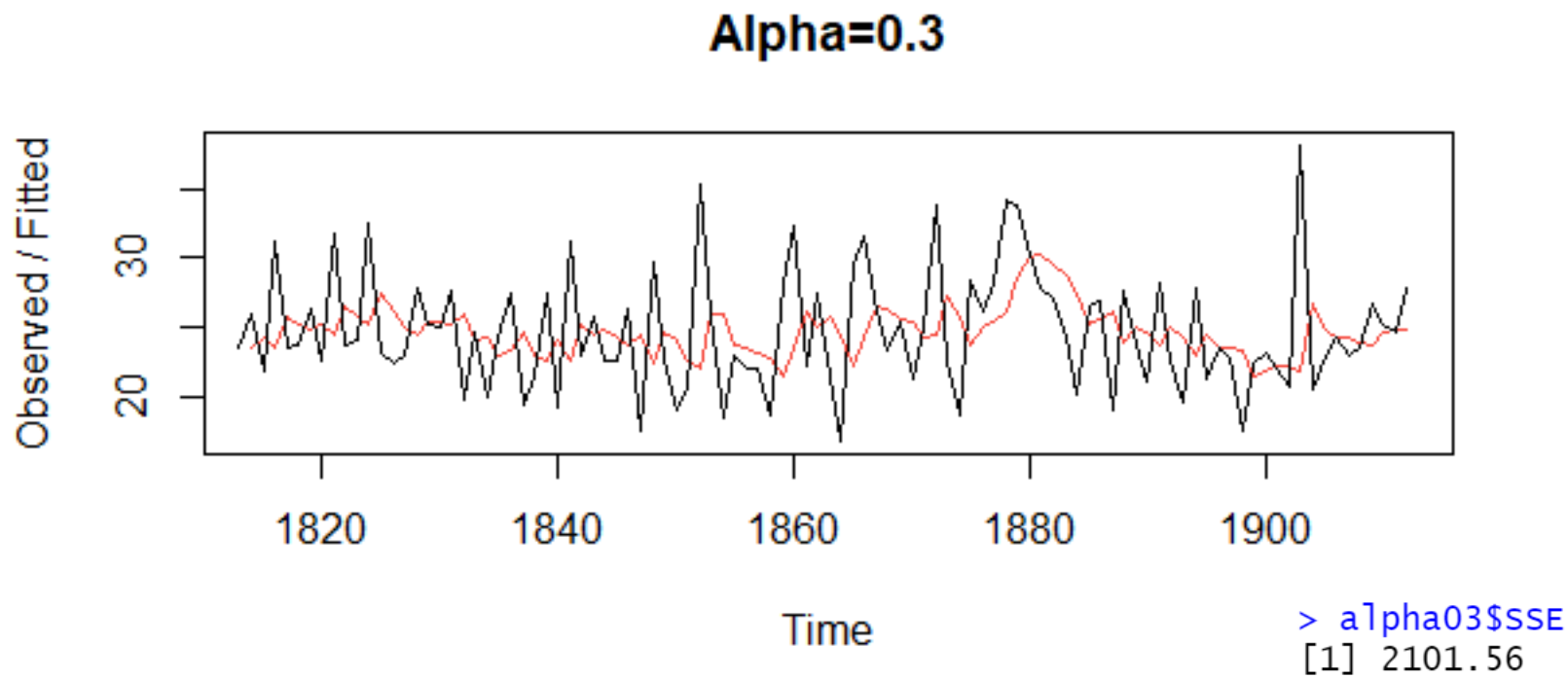
beta : FALSE

gamma: FALSE

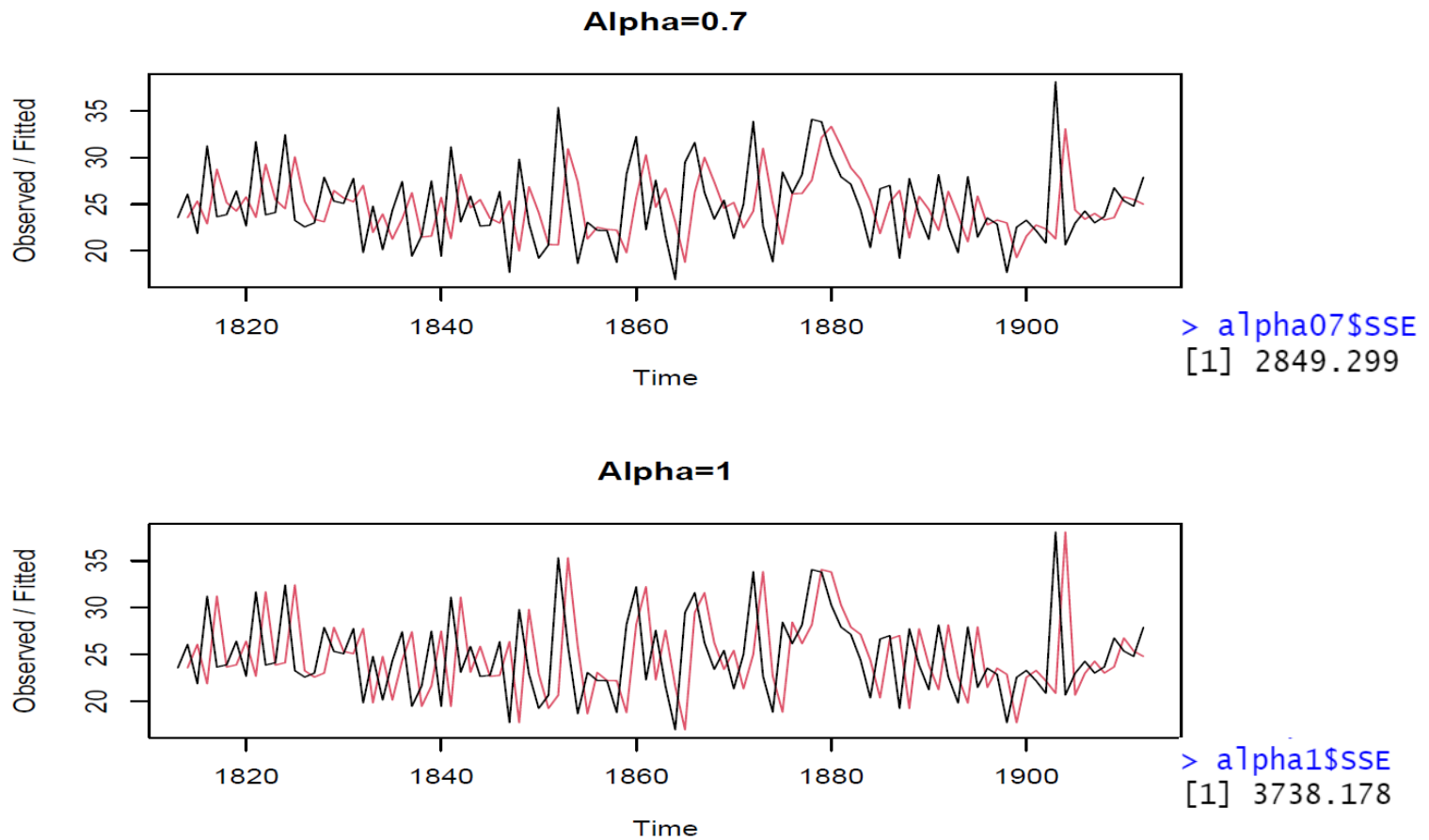


```
> rainforecasts$SSE  
[1] 1828.855
```

- 연간 강우량 (1813-1912)



■ 연간 강우량 (1813-1912)



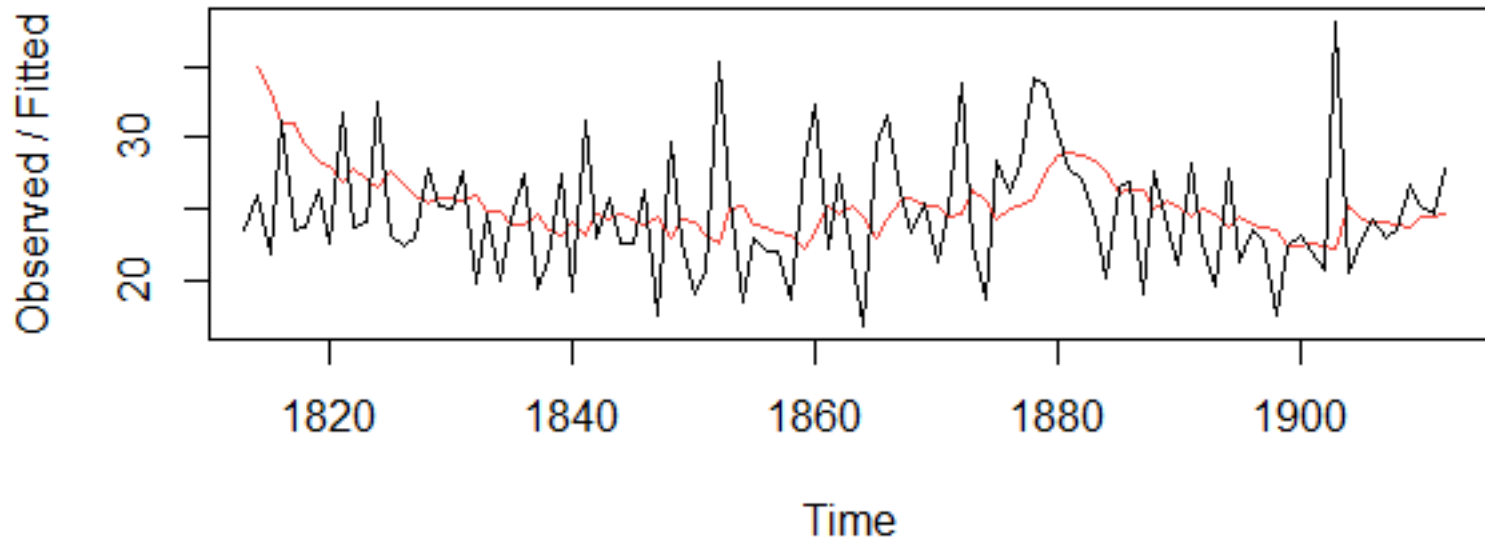
- 연간 강우량 (1813-1912) – 초기값을 바꾸었을 때 (초기값 = 35)

Smoothing parameters:

alpha: 0.1955854

beta : FALSE

gamma: FALSE



■ 모형

$$Z_t = \beta_{0,t} + \beta_{1,t}t + \varepsilon_t$$

- ε_t : 서로 독립, 평균 0, 분산 σ^2
- $\beta_{0,t}, \beta_{1,t}$: 시간에 따라 변하는 미지의 모수
- 선형추세모형 ($Z_t = \beta_0 + \beta_1 t + \varepsilon_t$)과의 차이점
 - 국지적으로는(locally) 동일한 추세를 갖지만 전체적으로는(globally) 시간대별로 추세가 변함

■ 단순지수평활 통계량

$$S_n^{(1)} = \omega \sum_{j=0}^{\infty} (1 - \omega)^j Z_{n-j}$$

■ 선형추세모형의 경우

$$\begin{aligned} - E \left(S_n^{(1)} \right) &= E \left(\omega \sum_{j=0}^{\infty} (1 - \omega)^j Z_{n-j} \right) \\ &= \omega \sum_{j=0}^{\infty} (1 - \omega)^j E(Z_{n-j}) \\ &= \omega \sum_{j=0}^{\infty} (1 - \omega)^j E(\beta_0 + \beta_1(n - j) + \varepsilon_{n-j}) \\ &= \{\beta_0 + \beta_1(n + 1)\} - \beta_1 \omega \sum_{j=0}^{\infty} (1 - \omega)^j (j + 1) \\ &= \beta_0 + \beta_1(n + 1) - \frac{\beta_1}{\omega} \\ - E \left(Z_{n+1} - S_n^{(1)} \right) &= \{\beta_0 + \beta_1(n + 1)\} - \left\{ \beta_0 + \beta_1(n + 1) - \frac{\beta_1}{\omega} \right\} = \frac{\beta_1}{\omega} \end{aligned}$$

■ 단순지수평활 통계량의 단순지수평활

- 단순지수평활 통계량 :

$$\begin{aligned} S_n^{(1)} &= \omega Z_n + (1 - \omega) S_{n-1}^{(1)} \\ &= \omega \sum_{j=0}^{n-1} (1 - \omega)^j Z_{n-j} + (1 - \omega)^n S_0^{(1)} \end{aligned}$$

- 이중지수평활 통계량

$$\begin{aligned} S_n^{(2)} &= \omega S_n^{(1)} + (1 - \omega) S_{n-1}^{(2)} \\ &= \omega \sum_{j=0}^{n-1} (1 - \omega)^j S_{n-j}^{(1)} + (1 - \omega)^n S_0^{(2)} \end{aligned}$$

- 이중지수평활 통계량의 기대값

$$\begin{aligned} E(S_n^{(2)}) &= \omega \sum_{j=0}^{\infty} (1 - \omega)^j E(S_{n-j}^{(1)}) \\ &= \omega \sum_{j=0}^{\infty} (1 - \omega)^j \{E(S_n^{(1)}) - \beta_1 j\} \\ &= E(S_n^{(1)}) - \frac{(1 - \omega)\beta_1}{\omega} \end{aligned}$$

- $S_n^{(1)}, S_n^{(2)}$ 의 기대값

- $E(S_n^{(1)}) = \beta_0 + \beta_1(n+1) - \frac{\beta_1}{\omega}$
- $E(S_n^{(2)}) = E(S_n^{(1)}) - \frac{(1-\omega)\beta_1}{\omega}$

- β_0, β_1 에 대해 정리

- $\beta_1 = \frac{\omega}{1-\omega} \{E(S_n^{(1)}) - E(S_n^{(2)})\}$
- $\beta_0 = 2E(S_n^{(1)}) - E(S_n^{(2)}) - \beta_1 n$

- β_0, β_1 의 불편추정량

- $\hat{\beta}_{1,n} = \frac{\omega}{1-\omega} \{S_n^{(1)} - S_n^{(2)}\}$
- $\hat{\beta}_{0,n} = 2S_n^{(1)} - S_n^{(2)} - \hat{\beta}_{1,n}n$

■ 시점 n 에서 l -시차 후의 예측값

$$- Z_{n+l} = \beta_0 + \beta_1(n+l) + \varepsilon_{n+l}$$

$$- \hat{Z}_n(l) = \hat{\beta}_{0,n} + \hat{\beta}_{1,n}(n+l)$$

$$= (2S_n^{(1)} - S_n^{(2)} - \hat{\beta}_{1,n}n) + \hat{\beta}_{1,n}(n+l)$$

$$= (2S_n^{(1)} - S_n^{(2)}) + \frac{\omega}{1-\omega}(S_n^{(1)} - S_n^{(2)})l$$

$$= \left(2 + \frac{\omega}{1-\omega}l\right)S_n^{(1)} - \left(1 + \frac{\omega}{1-\omega}l\right)S_n^{(2)}$$

$$- E(Z_{n+l} - \hat{Z}_n(l))$$

$$= E\{(\beta_0 + \beta_1(n+l) + \varepsilon_{n+l}) - (\hat{\beta}_{0,n} + \hat{\beta}_{1,n}(n+l))\}$$

$$= \beta_0 + \beta_1(n+l) - E(\hat{\beta}_{0,n}) + E(\hat{\beta}_{1,n})(n+l)$$

$$= 0$$

■ $n = 0$ 인 경우

$$- \hat{\beta}_{1,0} = \frac{\omega}{1-\omega} \{S_0^{(1)} - S_0^{(2)}\}, \hat{\beta}_{0,0} = 2S_0^{(1)} - S_0^{(2)}$$

■ 초기 평활값

$$- S_0^{(1)} = \hat{\beta}_{0,0} - \frac{1-\omega}{\omega} \hat{\beta}_{1,0}, S_0^{(2)} = \hat{\beta}_{0,0} - 2\frac{1-\omega}{\omega} \hat{\beta}_{1,0}$$

$$- \hat{\beta}_{0,0}, \hat{\beta}_{1,0} \text{의 값은 일반적으로}$$

$$Z_t = \beta_0 + \beta_1 t + \varepsilon_t$$

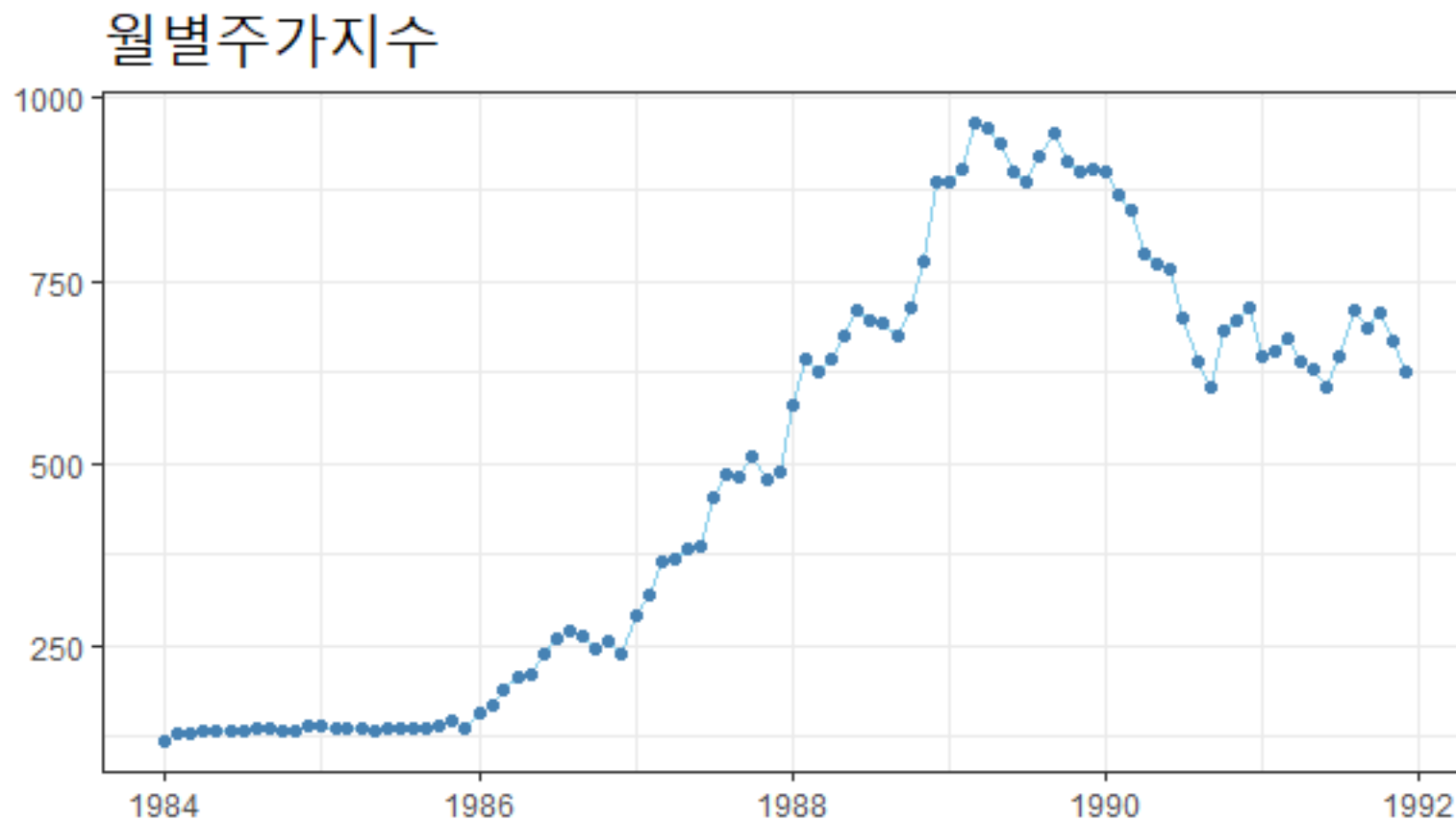
의 최소제곱추정량을 이용

- 평활상수는 단순지수평활법에서와 같은 방법으로 구함

■ 모형

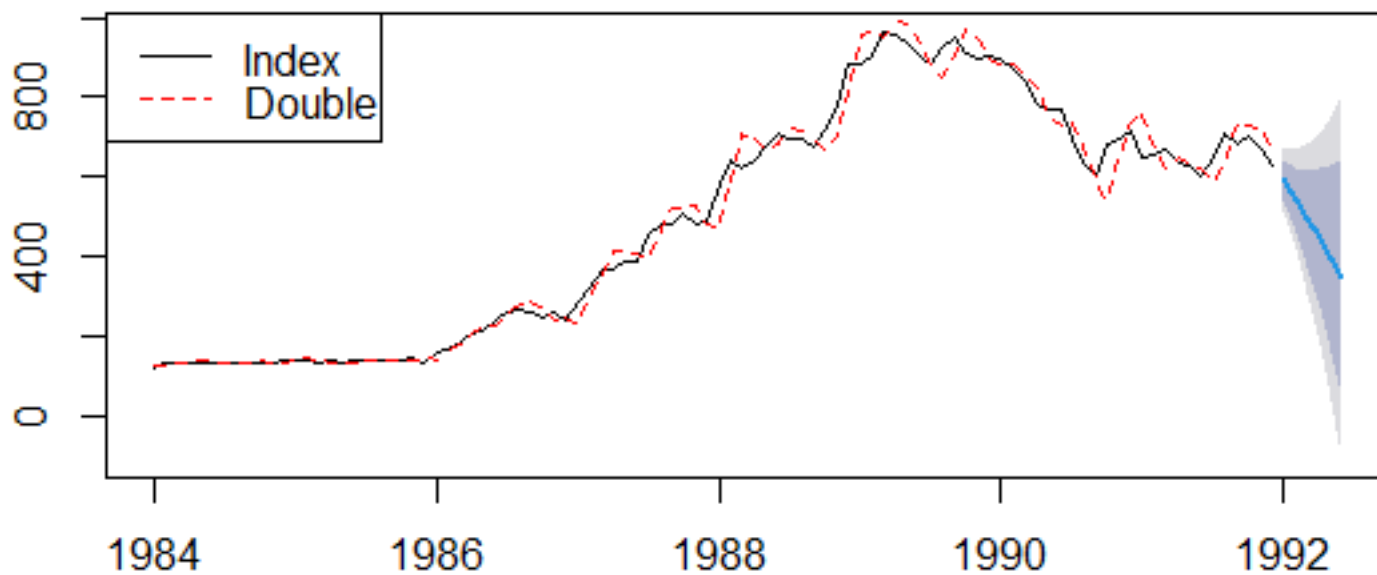
- $S_n^{(1)} = \alpha Z_n + (1 - \alpha) (S_{n-1}^{(1)} + S_{n-1}^{(2)})$
- $S_n^{(2)} = \beta (S_{n-1}^{(1)} - S_{n-1}^{(2)}) + (1 - \beta) S_{n-1}^{(2)}$
- α, β : 평활상수
- $S_n^{(1)}$: 수준(평균)에 대한 추정값
- $S_n^{(2)}$: 추세에 대한 추정값

■ 월별 주가지수 - 시계열그림



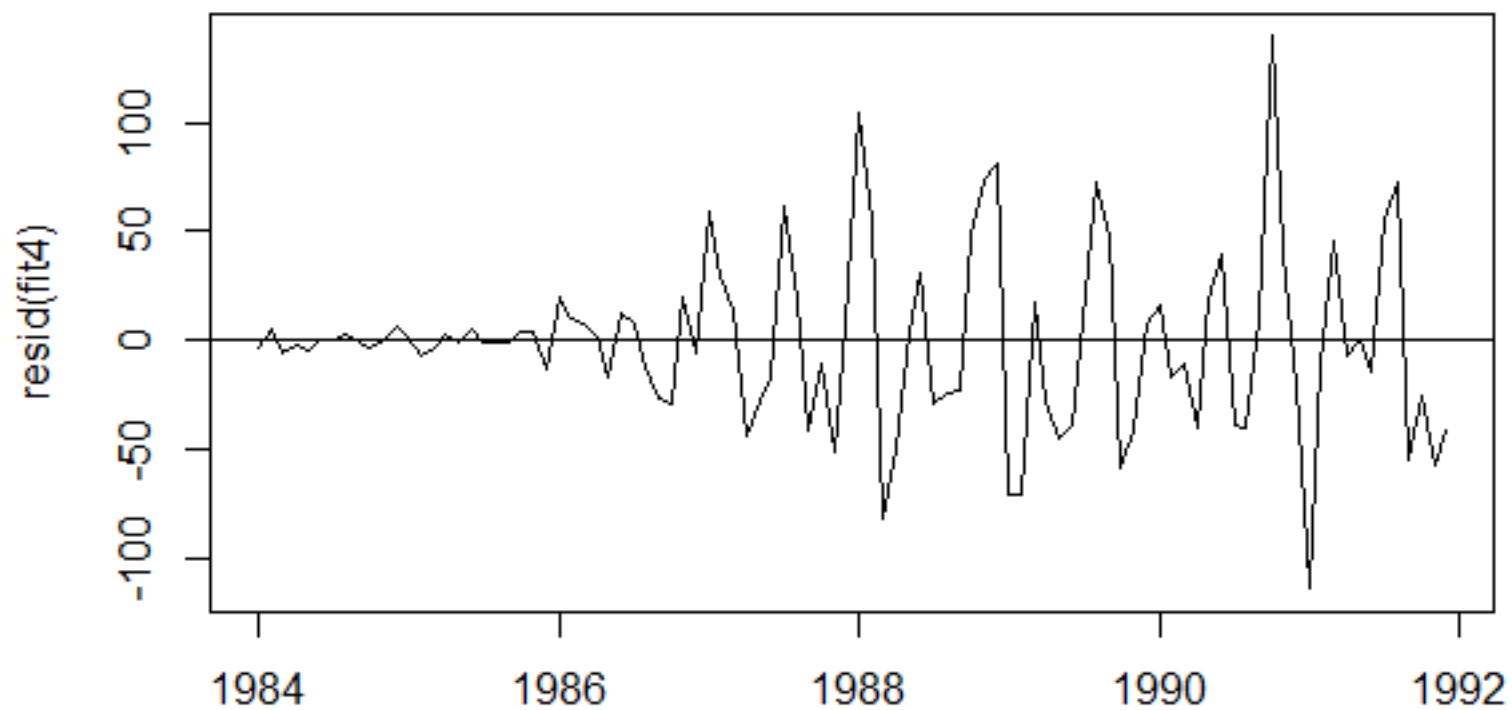
■ 월별 주가지수

주가지수와 이중지수평활값 : $\alpha=\beta=0.6$



■ 월별 주가지수

예측오차의 시계열그림



- Winters의 가법계절모형 (Winters' additive seasonal model) :

$$Z_t = T_t + S_t + I_t$$

- T_t : 추세성분
- S_t : 계절주기 s 를 갖는 계절성분
- I_t : 오차항, 불규칙 성분

- Winters의 가법계절지수평활법 :

- 시계열이 추세성분, 계절성분, 불규칙성분들의 합으로 구성
- 각 성분들을 평활법에 의해 추정한 후 이를 이용하여 예측값을 구하는 방법
- 시계열의 변동 폭이 시간의 흐름에 관계없이 일정한 경우에 사용
- $\hat{Z}_n(l) = \hat{T}_n(l) + \hat{S}_{n+l-ks}, \quad k = \left\lfloor \frac{l}{s} \right\rfloor + 1$

- 선형추세를 가정

$$T_{n+k} = \beta_{0,n} + \beta_{1,n}(n+k) = (\beta_{0,n} + \beta_{1,n}n) + \beta_{1,n}k = T_n + \beta_{1,n}k$$

- 계절성분

$$S_i = S_{i+s} = S_{i+2s} = \cdots (i = 1, 2, \dots, s), \quad \sum_{i=1}^s S_i = 0$$

- 예측

$$\hat{Z}_n(k) = \hat{T}_n + \hat{\beta}_{1,n}k + \hat{S}_{n+k-s}, \quad k = 1, 2, \dots, s$$

$$\hat{Z}_n(k) = \hat{T}_n + \hat{\beta}_{1,n}k + \hat{S}_{n+k-2s}, \quad k = s+1, s+2, \dots, 2s$$

$$\hat{Z}_n(k) = \hat{T}_n + \hat{\beta}_{1,n}k + \hat{S}_{n+k-3s}, \quad k = 2s+1, 2s+2, \dots, 3s$$

■ 예측갱신

$$\hat{T}_{n+1} = \alpha(Z_n - \hat{S}_{n+1-s}) + (1 - \alpha)(\hat{T}_n + \hat{\beta}_{1,n})$$

$$\hat{\beta}_{1,n+1} = \beta(\hat{T}_{n+1} - \hat{T}_n) + (1 - \beta)\hat{\beta}_{1,n}$$

$$\hat{S}_{n+1} = \gamma(Z_n - \hat{T}_{n+1}) + (1 - \gamma)\hat{S}_{n+1-s}$$

- α, β, γ : 평활상수
- \hat{T}_{n+1} : 수준(평균)에 대한 추정값
- $\hat{\beta}_{1,n+1}$: 추세에 대한 추정값
- \hat{S}_{n+1} : 계절성분에 대한 추정값

■ Winters의 승법계절모형 (Winters' additive seasonal model) :

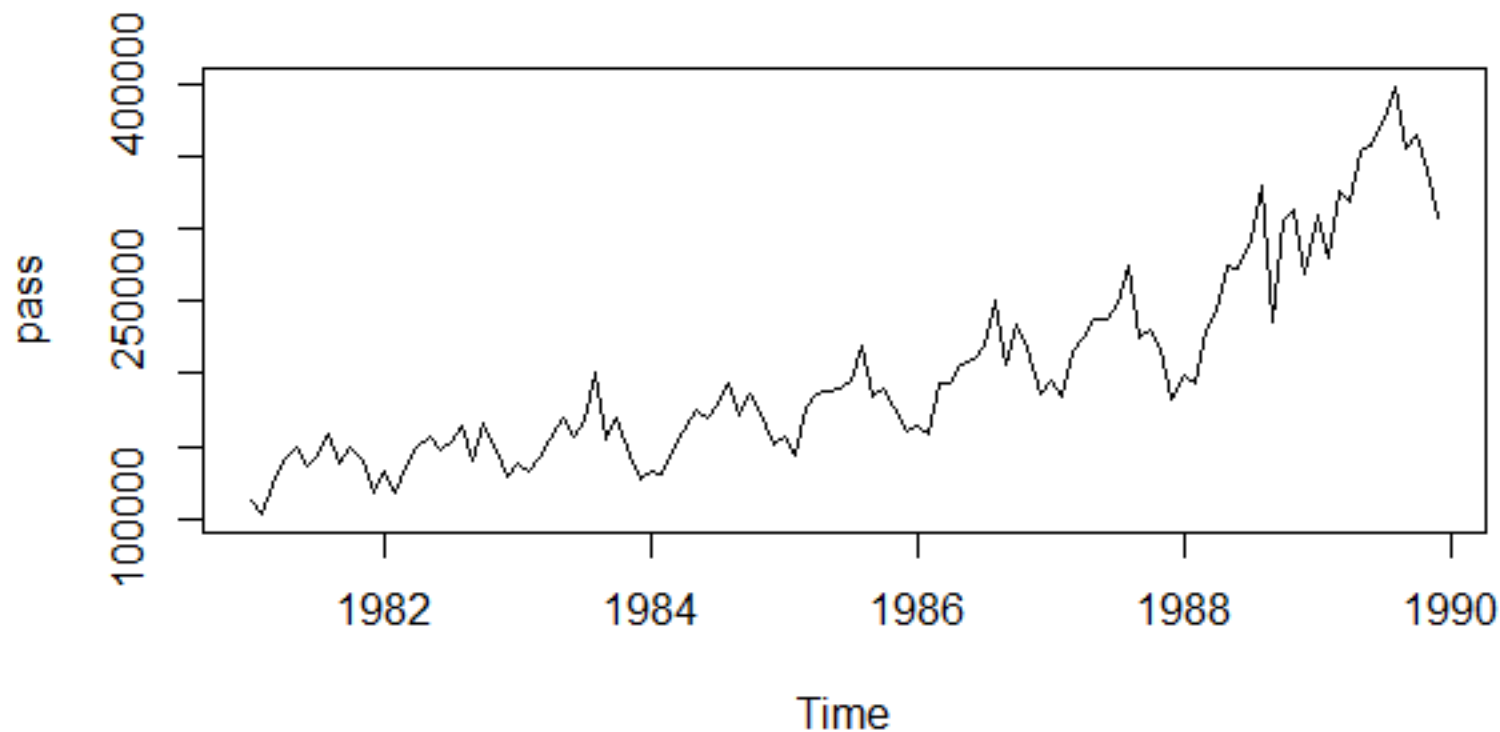
$$Z_t = T_t \times S_t + I_t$$

- T_t : 추세성분
- S_t : 계절주기 s 를 갖는 계절성분
- I_t : 오차항, 불규칙 성분

■ Winters의 승법계절지수평활법 :

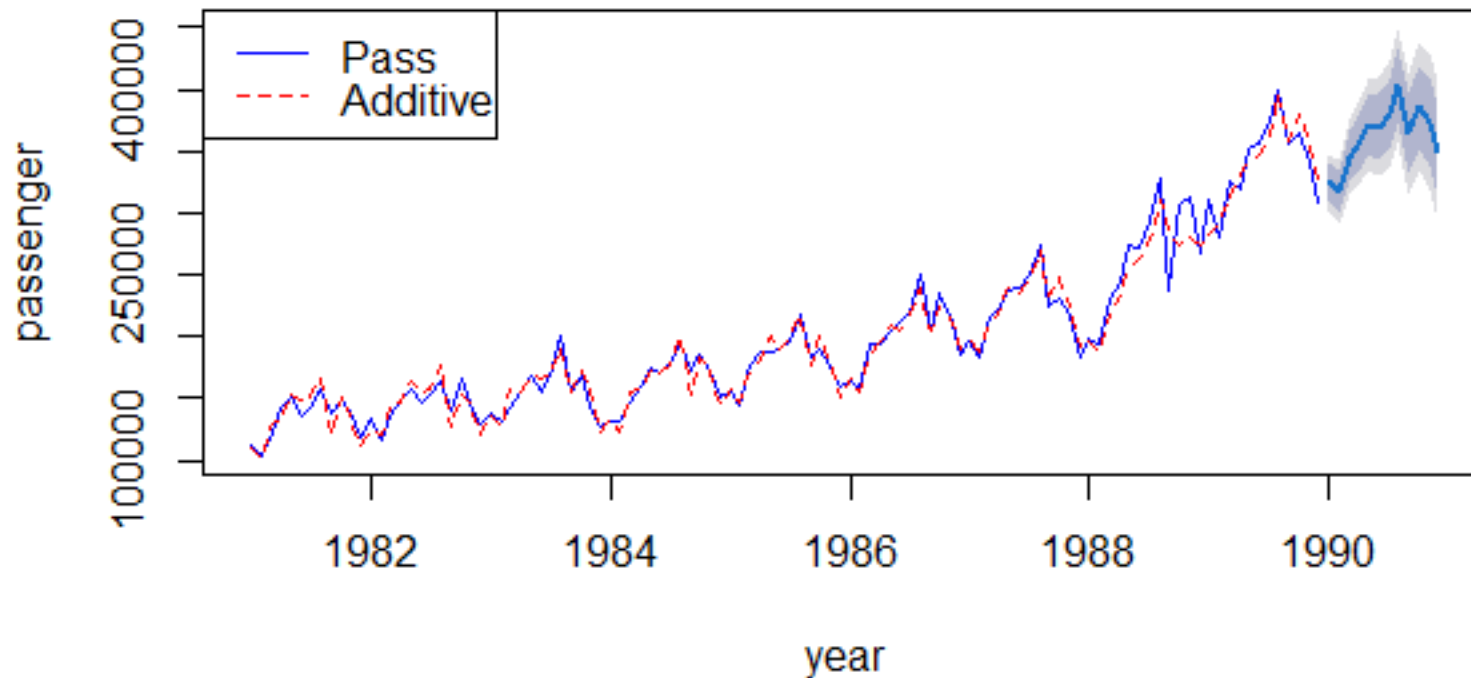
- 시계열이 추세성분과 계절성분의 곱과 불규칙성분의 합으로 구성
- 각 성분들을 평활법에 의해 추정한 후 이를 이용하여 예측값을 구하는 방법
- 시계열 변동의 폭과 계절주기의 폭이 추세에 비례하여 변화할 때 사용
- $\hat{Z}_n(l) = \hat{T}_n(l) \times \hat{S}_{n+l-kS}, \quad k = \left\lfloor \frac{l}{s} \right\rfloor + 1$

- 비행기 승객 수 - 시계열 그림

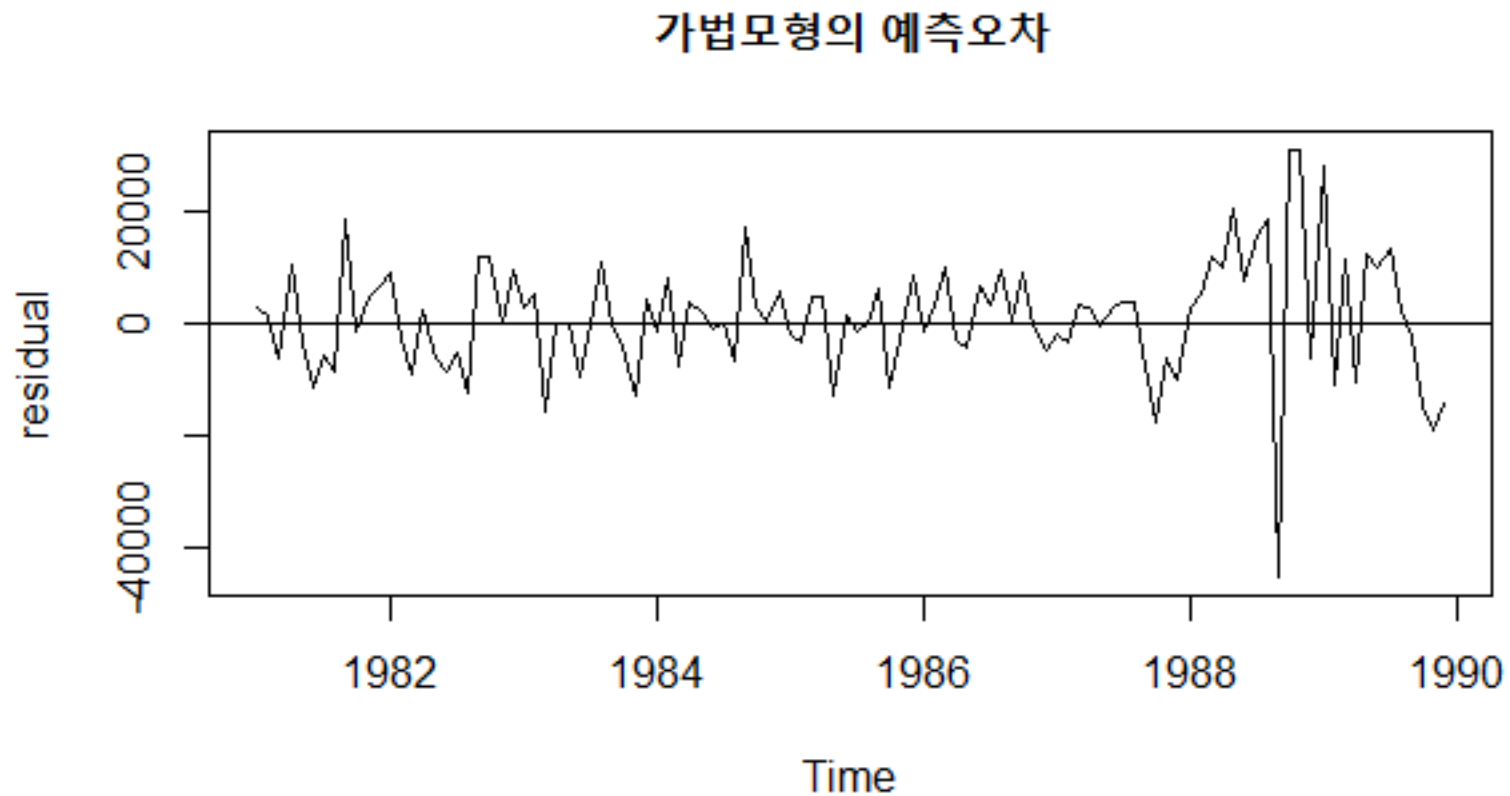


■ 비행기 승객 수 -가법모형

Winters 가법계절지수평활된 자료의 시계열 그림

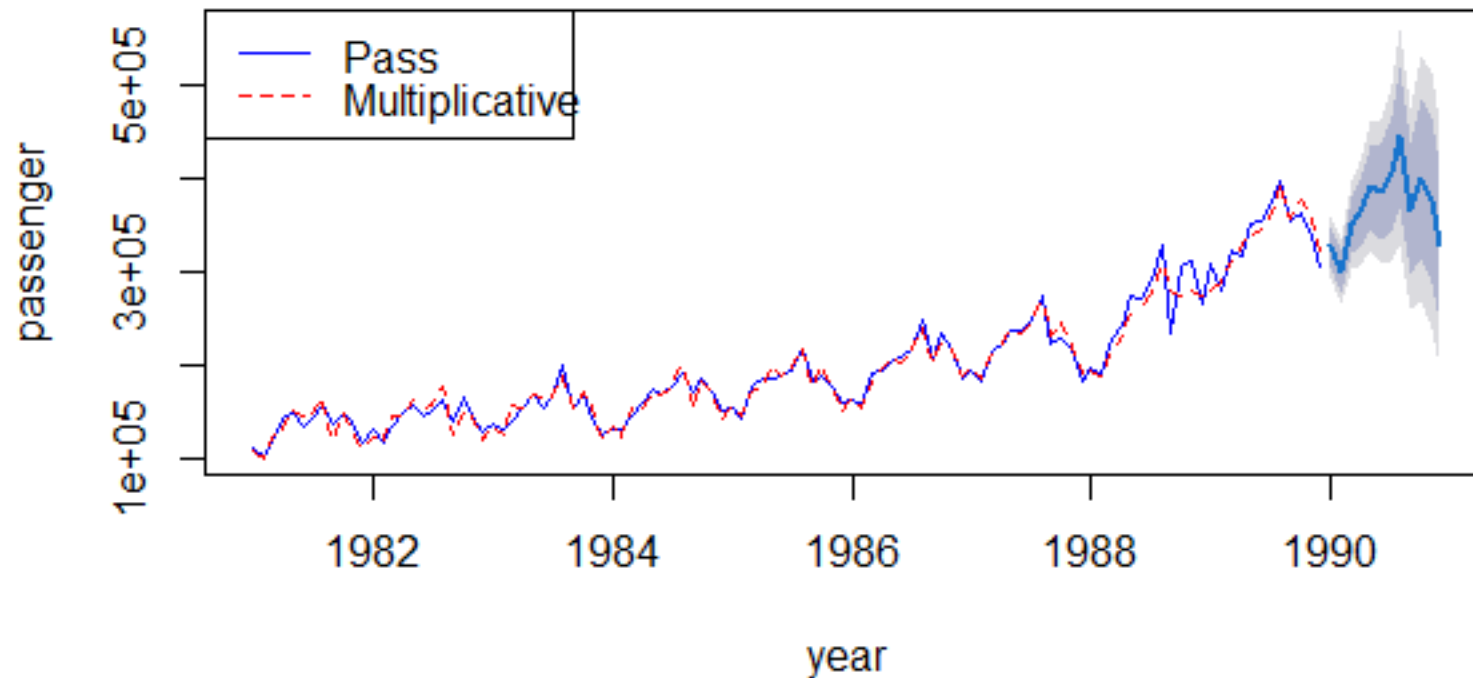


- 비행기 승객 수 -가법모형(예측 오차 잔차 그림)

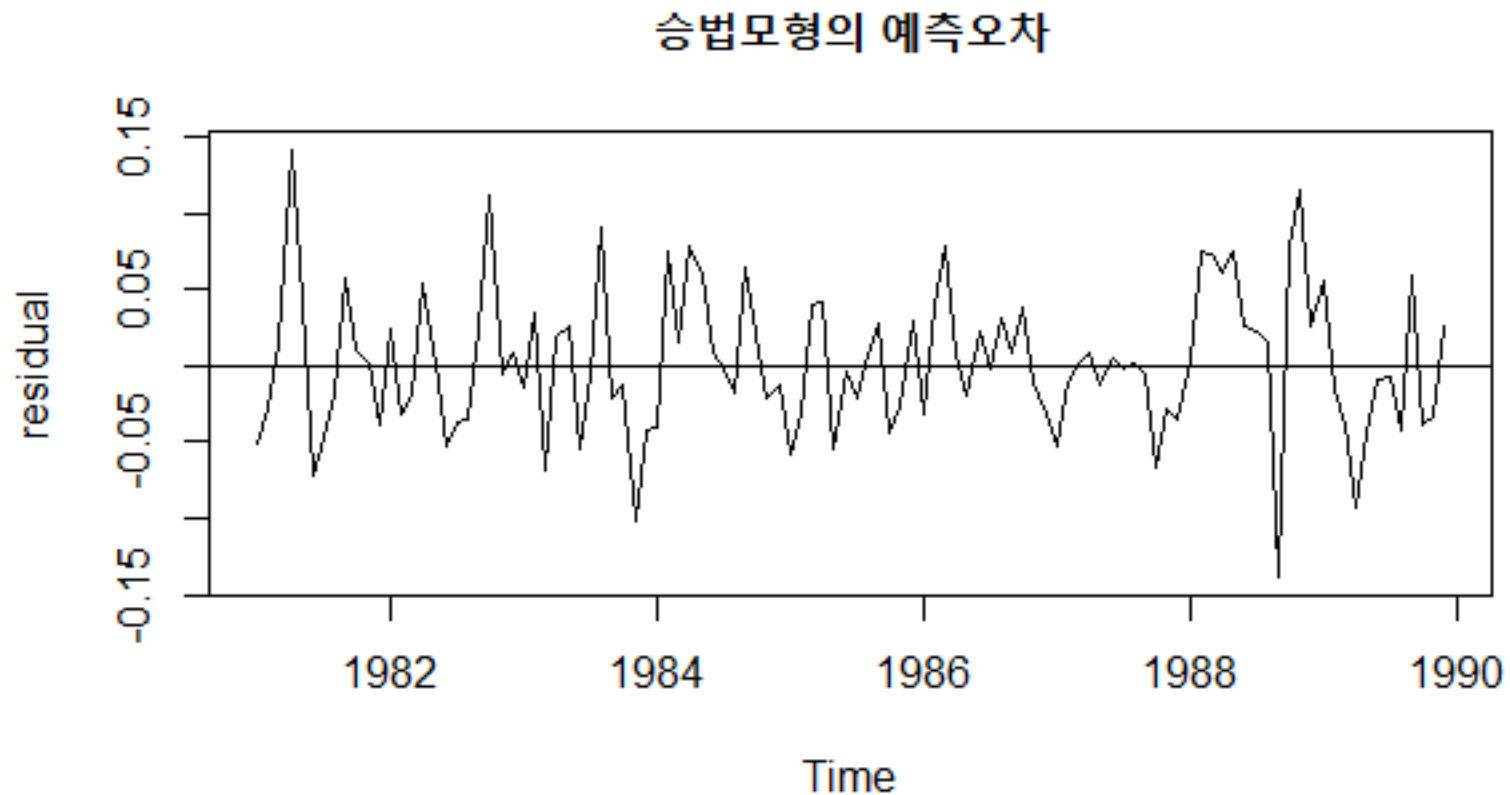


■ 비행기 승객 수 -승법모형

Winters 승법계절지수평활된 자료의 시계열 그림



- 비행기 승객 수 -승법모형(예측 오차 잔차 그림)



End of Document