

지수평화법

지수 평활법 exponential smoothing

- 모든 관측치에 동일한 가중치를 부여하는 이동평균법은 최근 관측치나 오래된 관측치나 동일한 가중치를 사용하므로 정보를 동일하게 이용하는 단점이 있고 과거 추세 패턴을 인지함,
- 최근 관측치에 높은 가중치, 멀어질수록 지수적으로 가중치 값 감소시키는 방법인 지수평활법은 미래 예측에 사용

단순지수평활법 Simple ES

- 시계열 데이터 $\{Y_t\}$ 에 경향(추세)과 계절성이 없는 경우 사용
- 시계열 모형 : $Y_t = \mu_t + e_t, e_t \sim N(0, \sigma^2)$

평활값

- $S_t = wY_t + (1 - w)S_{t-1}$
- 예측치로 사용 $\hat{Y}_{t+1} = S_t$
- 이전 관측치의 선형결합
$$S_t = wY_t + w(1 - w)Y_{t-1} + w(1 - w)^2Y_{t-2} + \dots$$
- 초기값 설정 $S_0 = \sum_{i=1}^T Y_i / T, T = 6 \text{ or } n/2$

가중치 설정

- 가중치 w 가 클수록 최근 값에 영향이 큼
- Brown : 0.05~0.3
- $M = \frac{2 - w}{w} (m=7, w=0.25)$

- Montgomery & Johnson : $1 - 0.8^{(1/\text{추세 기울기})}$
- 절편 없는 $ARMA(0, 1, 1)$ 에서 $MA(1)$ 의 추정계수 = $(1 - w)$
- MSE (혹은 SSE) 가장 적게하는 가중치

In R : 전력수요량(2011.1.~2013.12.)

```
fit0=HoltWinters(yt.pd, gamma=F, beta=F)
fit1=HoltWinters(yt.pd, gamma=F, beta=F,alpha=0.1)
fit2=HoltWinters(yt.pd, gamma=F, beta=F,alpha=0.2)
fit3=HoltWinters(yt.pd, gamma=F, beta=F,alpha=0.3)
fit0$SSE;fit1$SSE;fit2$SSE;fit3$SSE
fit9=HoltWinters(yt.pd, gamma=F, beta=F,alpha=0.05)
```

- 함수 HoltWinters() : alpha를 지정하지 않으면 , M&J 방법으로 최적 알파를 구한다.
- 계절성 주기 - 단순지수평활법 : $b = 0; S = 0$

$$\hat{Y}_{t+h} = a_t + h \times b_t + S_{t-s+1+(h-1)\text{mod}(s)}$$

```
> fit0$SSE;fit1$SSE;fit2$SSE;fit3$SSE
[1] 32481366754
[1] 33408050288
[1] 32502919755
[1] 32952336031
```

```
> fit0
Holt-Winters exponential smoothing without trend and without seasonal component.

Call:
HoltWinters(x = yt.pd, beta = F, gamma = F)

Smoothing parameters:
alpha: 0.1810032
beta : FALSE
gamma: FALSE

Coefficients:
      [,1]
a 68855.72
```

추정치 $\hat{\mu}_t = \hat{a}_t = 0.69$ in Best fir of "weight"

```
predict(fit0,n.ahead=7) #next t+1, t+2, ..
```

- 함수 predict() : 향후 7개(n.ahead=7) 관측치를 예측하라는 함수이다.
- 단순지수 평활법은 차기 1기만 예측 가능 - 그 이후는 동일함

```
> predict(fit0,n.ahead=7) #next t+1, t+2, ..
Time Series:
Start = c(2014, 2)
End = c(2014, 8)
Frequency = 365
      fit
[1,] 68855.72
[2,] 68855.72
[3,] 68855.72
[4,] 68855.72
```

```
fitted(fit0)[,1] #smoothing value
```

- 함수 fitted() : (단순)지수평활 추정 결과 (추정값 \hat{a}) 이용하여 매 시점 지수평활 값(smoothing value) 계산
- [,1] - 계산 결과 첫 열이 평활값임 - 단순지수평활값의 경우 1열과 2열 결과는 동일함

계절성, 직선 경향이 존재하는 시계열이므로 단순지수평활법 적절하지 않음

In R 사망률(2001.~2015.12.)

```
ds0=read.csv("population_vital.csv")
yt.ma=ts(ds0[13:192,4],frequency=12,start=c(2001,1)) #2000년 12개
데이터 삭제, 4번째 열 데이터 marriage - 시계열 데이터 만들기, 2001년 1월
~2015년 12월
fit.ma=HoltWinters(yt.ma, gamma=F, beta=F, alpha=0.3)
fit.ma; fit.ma$SSE
predict(fit.ma,n.ahead=12) #next t+1, t+2, ..
plot(yt.ma,main="Marriage(2001.1~2015.12.)",ylab="(gun)")
lines(fitted(fit.ma)[,1],col="red", lty="dashed")
lines(predict(fit.ma,n.ahead=6),col="blue")
```

```
> fit.ma
Holt-Winters exponential smoothing

Call:
HoltWinters(x = yt.ma, alpha = 0.3,

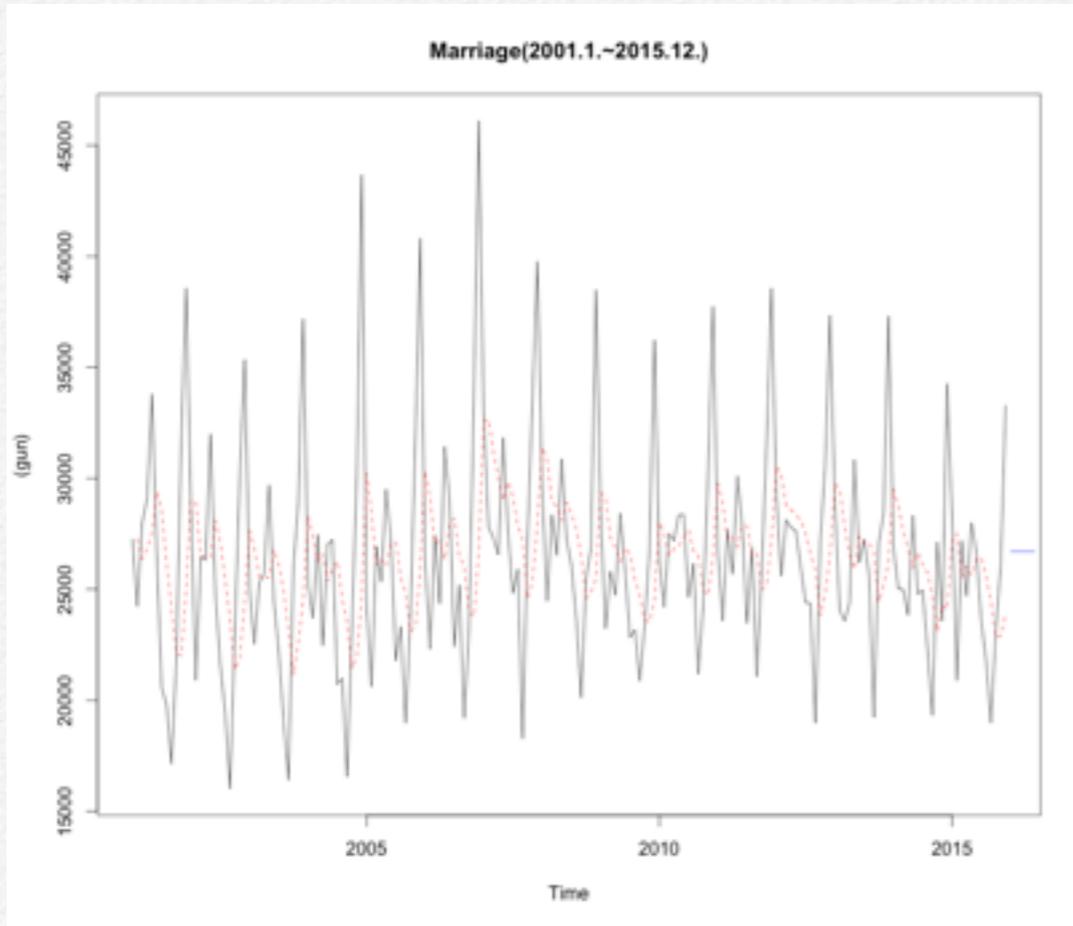
Smoothing parameters:
alpha: 0.3
beta : FALSE
gamma: FALSE

Coefficients:
      [,1]
a 26708.64
```

```

> fit.ma$SSE
[1] 5330177252
> predict(fit.ma,n.ahead=12) #next t+1, t+2, ..
      Jan      Feb      Mar      Apr      May
2016 26708.64 26708.64 26708.64 26708.64 26708.64 26708.64

```

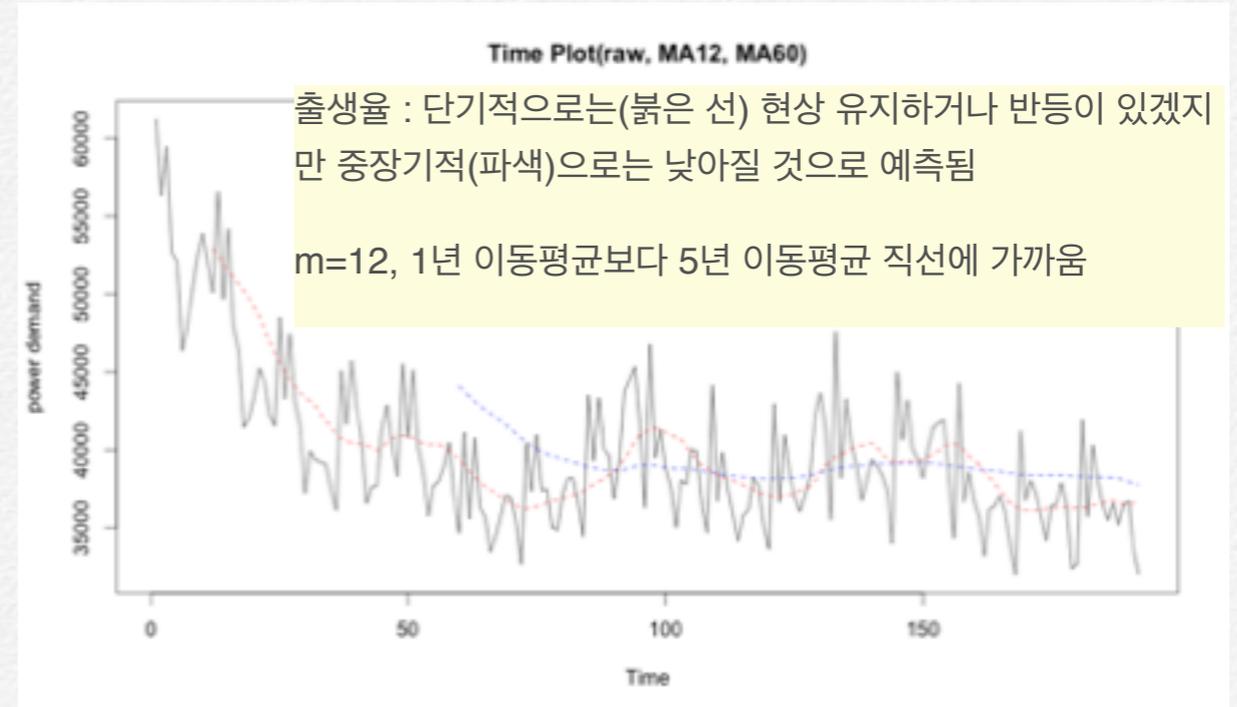
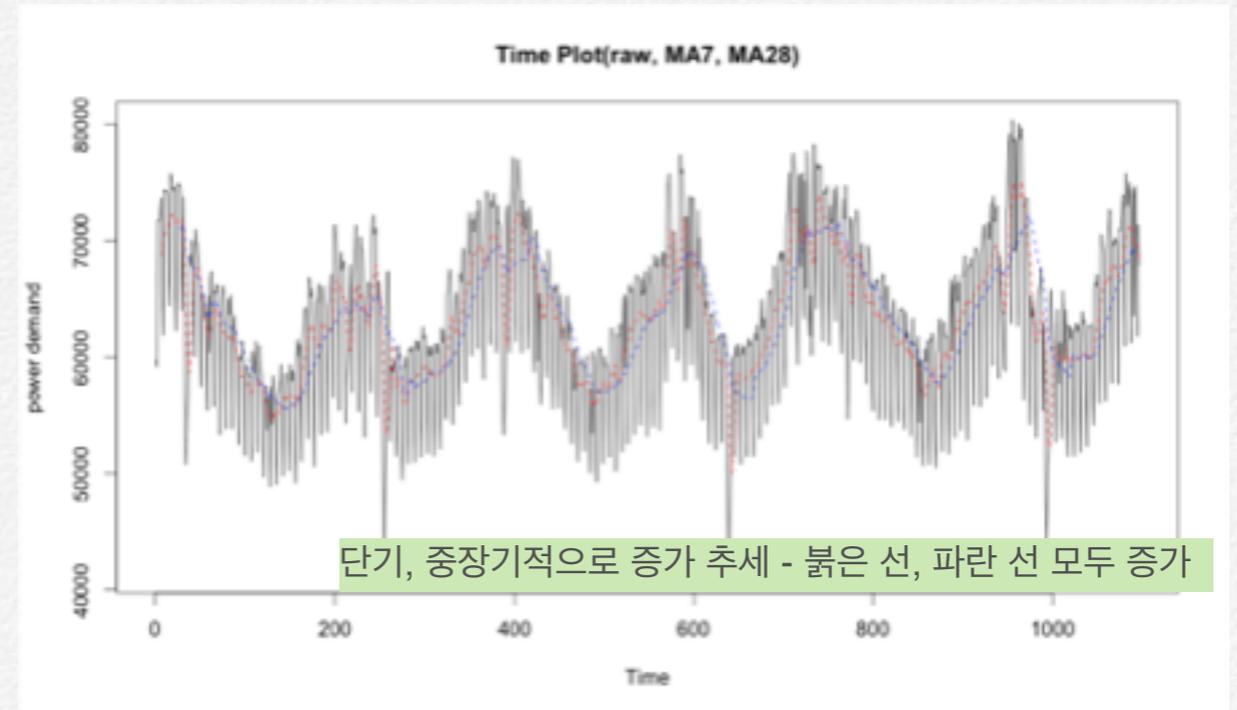


default

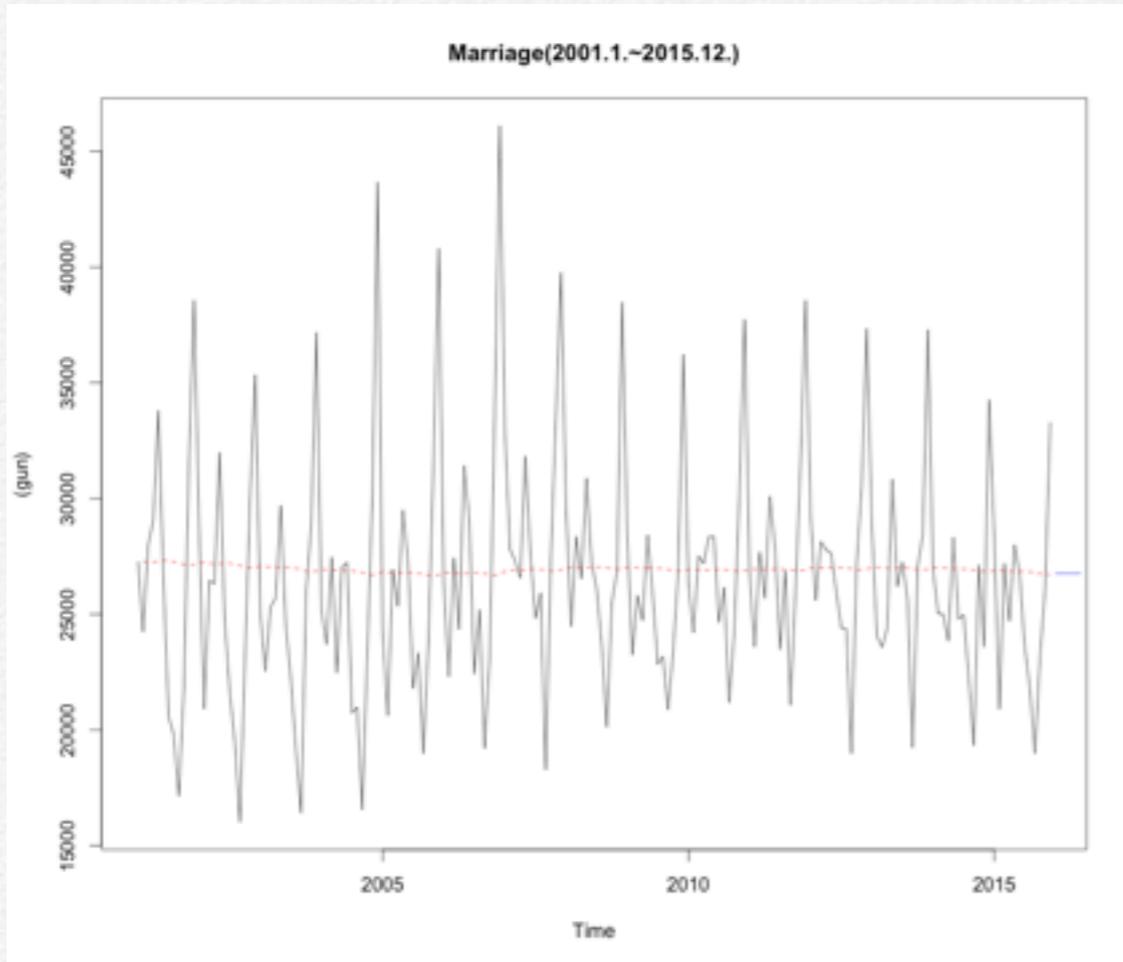
alpha 사용하면 (SSE는 가장 낮음)

최근 값에 가중치가 적고 이전 값들에 가중치가 높아 평활값이 평평함

단순지수평활법에 의한 예측은 다음 1기만 가능하여 12개 모두 동일하다. 다음 달(2016년 1월) 관측값이 관측되면 2월 예측이 가능하다.



계절성이 존재하는 시계열이므로 단순지수평활법 적절하지 않음



시계열 모형 적합도

MAPE 평균절대 퍼센트 Mean Absolute Percentage Error

$$MAPE = \frac{\sum_{i=1}^T \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right|}{T} \times 100(\%) - \text{서로 다른 데이터 모형 비교}$$

MSE 평균제곱오차 Mean Squared prediction Error

$$MSE = \frac{\sum_{i=1}^T (Y_t - \hat{Y}_t)^2}{T} - \text{동일 데이터 모형 비교-분모에 T가 없다면 SSE}$$

이중 지수평활법 Double Exponential Smoothing

시간 추세가 직선형태를 가진 시계열 데이터 예측에 적합한 평활법

$$Y_t = \mu_t + \beta_t T + e_t \quad \text{--- (1)}$$

계산 :

$$S_t^{[1]} = wY_t + (1-w)S_{t-1}^{[1]}, S_t^{[2]} = wS_t^{[1]} + (1-w)S_{t-1}^{[2]} \text{ (일모수)}$$

$$S_t^{[1]} = w_1Y_t + (1-w_1)S_{t-1}^{[1]}, S_t^{[2]} = w_2S_t^{[1]} + (1-w_2)S_{t-1}^{[2]} \text{ (이모수)}$$

$$\text{추정치 } \hat{Y}_{t+L} = \left(2 + \frac{w}{1-w}L\right)S_t^{[1]} - \left(1 + \frac{w}{1-w}L\right)S_t^{[2]}$$

초기값 $S_0^{[1]}, S_0^{[2]}$

식 (1)을 이용하여 OLS 추정치

가중치

- Brown : 0.03~0.16 제안
- ARMA(0,2,2) - ma(1), ma(2) 추정계수 활용
- In R, best - HoltWinters(yt.ma, gamma=F) - no alpha=, beta=

삼중 지수평활법 Double Exponential Smoothing

추세와 계절성이 있는 모형에 적용

계산 :

$$S_t^{[1]} = wY_t + (1 - w)S_{t-1}^{[1]}$$

$$S_t^{[2]} = wS_t^{[1]} + (1 - w)S_{t-1}^{[2]}$$

$$S_t^{[3]} = wS_t^{[2]} + (1 - w)S_{t-1}^{[3]}$$

추정치 - complicated

초기값 $S_0^{[1]}, S_0^{[2]}, S_0^{[3]}$ - 식 (2)을 이용하여 OLS 추정치

가중치

- Brown : 0.02~0.11 제안
- In R - HoltWinters() 함수 이용 GAMMA 모수

In R : 전력수요량(2011.1.1.~2013.12.31.) - 직선 추세, 계절성

```
fit.pd.s=HoltWinters(yt.pd, gamma=F, beta=F); fit.pd.s$SSE
fit.pd.d=HoltWinters(yt.pd, gamma=F); fit.pd.d$SSE
```

```
> fit.pd.s=HoltWinters(yt.pd, gamma=F, beta=F); fit.pd.s$SSE
[1] 32481366754
> fit.pd.d=HoltWinters(yt.pd, gamma=F); fit.pd.d$SSE
[1] 33121272811
```

단순지수평활법이 이중지수평활법보다 예측력이 높음 - 직선추세 없음, 계절성만 고려하는 것이 적절함

In R : 결혼건수(2001.1.~2015.12.) - 계절성, 추세 없음

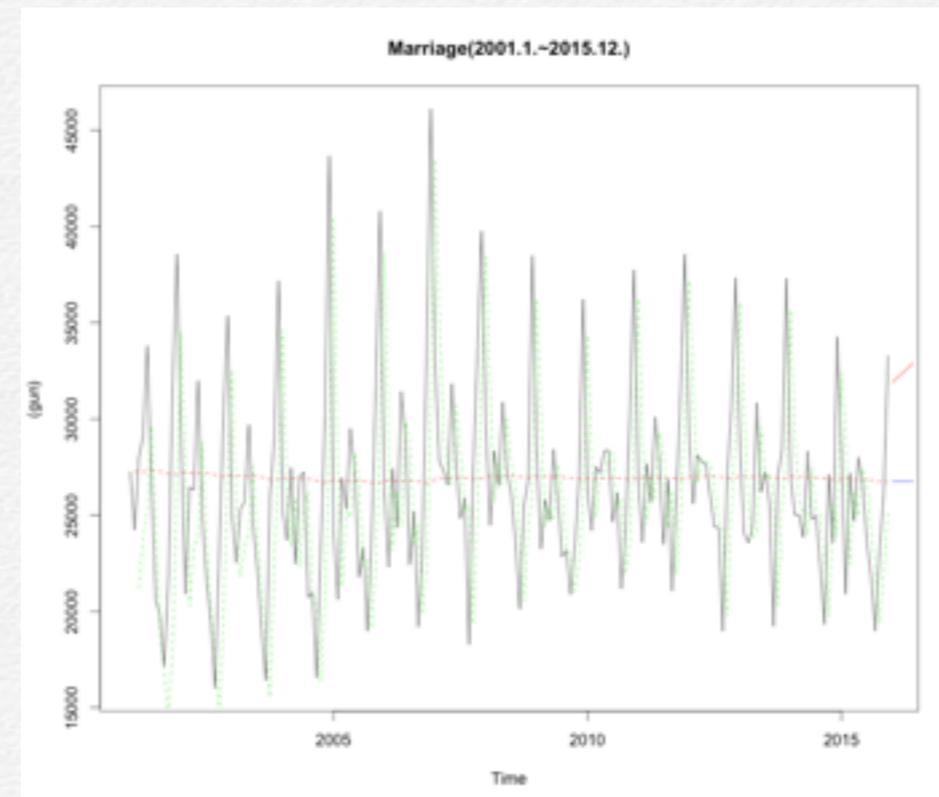
```
pv.ma=ts(ds0[13:192,4],frequency=12,start=c(2001,1))
fit.pv.ma.s=HoltWinters(pv.ma, gamma=F, beta=F); fit.pv.ma.s$SSE
fit.pv.ma.d=HoltWinters(pv.ma, gamma=F); fit.pv.ma.d$SSE
```

```
> fit.pv.ma.s=HoltWinters(pv.ma, gamma=F, beta=F); fit.pv.ma.s$SSE
[1] 4623616757
> fit.pv.ma.d=HoltWinters(pv.ma, gamma=F); fit.pv.ma.d$SSE
[1] 6146497218
```

단순지수평활법이 이중지수평활법보다 예측력이 높음 - 직선추세 없음, 계절성만 고려하는 것이 적절함

지수평활 값 그래프

```
plot(pv.ma,main="Marriage(2001.1.~2015.12.)",ylab="(gun)")
lines(fitted(fit.pv.ma.s)[,1],col="red",lty="dashed")
lines(fitted(fit.pv.ma.d)[,1],col="green",lty="dashed")
lines(predict(fit.pv.ma.s,n.ahead=6),col="blue")
lines(predict(fit.pv.ma.d,n.ahead=6),col="red")
```



단순, 이중 지수평활값을 시간도표에 그린 것임

이중지수평활값은 단순지수평활법과 달리 향후 1차기 아닌 계속적 차기까지 예측가능하다.

```
> predict(fit.pv.ma.d,n.ahead=6)
      Jan      Feb      Mar      Apr      May      Jun
2016 31889.94 32093.65 32297.36 32501.06 32704.77 32908.48
```

Winters 계절지수평활법

- 이전 설명한 지수평활법은 계절성분이 없는 경우 사용된다. 그러므로 계절성이 있는 시계열 데이터에는 적합하지 않다. 사실 대부분의 데이터는 계절성을 가지고 있음
- Winters 모형에는 가법모형과 승법모형이 있고, 변동분해법과 관련이 있음
- Winters 방법은 각 성분을 평활법으로 추정하고 이를 이용하여 시계열 값을 예측한다.
- 이런 점에서 이모수 가중치(이중지수평활법)과 동일

1) 가법모형 $Y_t = \alpha + \beta T + S_t + e_t$

- 계절성이 존재하고 시계열 변동 폭이 시간의 흐름에 따라 변동이 없는 경우 사용

2) 승법모형 $Y_t = (\alpha + \beta T)S_t + e_t$

- 시계열 변동 폭이 시간의 흐름에 따라 커지는 없는 경우 사용 (이분산 모형)
- 이분산 시계열 데이터는 (G)ARCH(General Auto-Regressive Heteroskadicity) 모형을 적용하여 예측한다.

In R : 전력수요량(2011.1.1.~2013.12.31.) - 계절성

```
fit.pd.s=HoltWinters(yt.pd, gamma=F, beta=F); fit.pd.s$SSE
fit.pd.w=HoltWinters(yt.pd, beta=F); fit.pd.w$SSE
fit.pd.w2=HoltWinters(yt.pd, beta=F, seasonal=c("multiplicative"))
fit.pd.w2$SSE
```

7일 계절 성분이 있어 gamma를 추정하였고 추세는 없으므로 beta=F를 사용하였다.

시간도표 분석결과 주기의 변동 폭은 일정하여 가법모형이 적절해 보임 - 실제 가법모형의 SSE=244xxx 가 승법의 249xxx 보다 적음

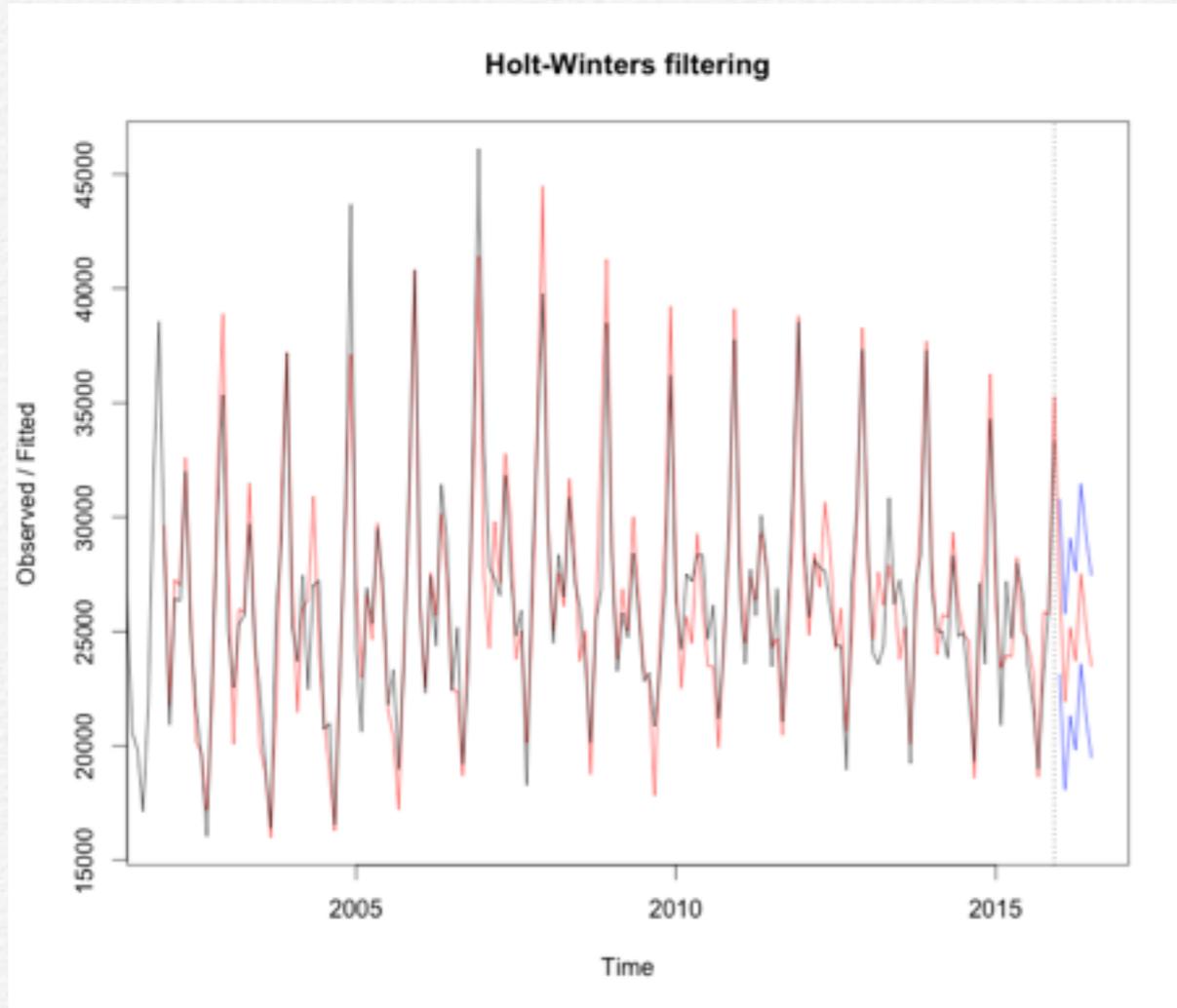
```
> fit.pd.s=HoltWinters(yt.pd, gamma=F, beta=F); fit.pd.s$SSE
[1] 32481366754
> fit.pd.w=HoltWinters(yt.pd, beta=F); fit.pd.w$SSE
[1] 24430039260
> fit.pd.w2=HoltWinters(yt.pd, beta=F, seasonal=c("multiplicative"))
> fit.pd.w2$SSE
[1] 24978480927
```

전력수요 예측의 최적 지수평활법은 “가법 계절 단순 지수평활법”

In R : 결혼건수(2001.1.~2015.12.) - 계절성, 추세 없음

```
pv.ma=ts(ds0[13:192,4],frequency=12,start=c(2001,1))
fit.pv.ma.w2=HoltWinters(pv.ma, beta=F, seasonal=c("multiplicative"))
fit.pv.ma.w2$SSE
fit.pv.ma.w=HoltWinters(pv.ma, beta=F); fit.pv.ma.w$SSE
plot(pv.ma); plot(fit.pv.ma.w)
plot(fit.pv.ma.w,predict(fit.pv.ma.w,n.ahead=7,prediction.interval =T))
```

plot() 함수를 이용하여 지수 평활값, 예측치; 그래프 가능



추세없는 계절 원가법 터스 지수평활법 적용

```
> fit.pv.ma.w2=HoltWinters(pv.ma, beta=F, seasonal=c("multiplicative"))
> fit.pv.ma.w2$SSE
[1] 646304324
> fit.pv.ma.w=HoltWinters(pv.ma, beta=F); fit.pv.ma.w$SSE
[1] 635419724
```

```
> predict(fit.pv.ma.w,n.ahead=7,prediction.interval=T)
              fit      upr      lwr
Jan 2016 26975.47 30796.57 23154.36
Feb 2016 21936.84 25785.90 18087.78
Mar 2016 25186.58 29063.39 21309.76
Apr 2016 23716.50 27620.88 19812.13
May 2016 27532.00 31463.74 23600.25
Jun 2016 25138.18 29097.10 21179.26
Jul 2016 23448.51 27434.42 19462.60
```

```
> fit.pv.ma.w
Holt-Winters exponential smoothing without trend and with additive seasonal
.

Call:
HoltWinters(x = pv.ma, beta = F)

Smoothing parameters:
alpha: 0.1211897
beta : FALSE
gamma: 0.4726588
```