 2022.03.02.
김은영

제1장 시계열 예측을 위한 평활 기법

1.1 시계열분석의 목적

■ 시계열(time series): 하나의 변수에 대한 시간에 따른 관측치
■ 시계열분석 목적
- 시계열의 패턴을 요약하고 시간에 따른 상관관계, 추세, 계절성 등의 특성을 파악
- 미래 시점에 대한 예측

1.2 이동평균법

■ **평활법(smoothing method)**: 시계열의 시간에 따른 패턴 파악을 용이하게
➤ **이동평균법(moving average)** $cf)$ **지수평활법(exponential smoothing)**
매 시점에서 직전 N개 데이터의 평균을 산출하여 평활치로 사용하는 방법
- **단순이동평균법(simple moving average)**
시계열 데이터 $\{X_1, X_2, \cdots\}$가 수평적인 패턴인 경우 사용하는 방법
- **이중이동평균법(double moving average)**
시계열 데이터 $\{X_1, X_2, \cdots\}$가 추세 패턴을 따르는 경우 사용하는 방법

1.2.1 수평적 패턴과 단순이동평균

■ 시계열값 X_t가 수평적 패턴을 갖는다고 가정하면

$$X_t = c + a_t$$

여기서, c: 상수, $a_t \sim i.i.d.\left(0,\ \sigma_a^2\right)$: 오차항, 백색잡음(white noise)
■ 만약 시점 T에서 최근 N개의 시계열 데이터만 사용하여 상수 c를 추정한다면

$$\min_c Q = \sum_{t=T-N+1}^{T} (X_t - c)^2 \rightarrow \frac{\partial Q}{\partial c} = 0$$

$\Rightarrow \hat{c} = \frac{1}{N}\sum_{t=T-N+1}^{T} X_t$: 시점 T에서 시계열 X_t에 대한 기간 N의 단순이동평균

$$M_T = \frac{1}{N}\sum_{t=T-N+1}^{T} X_t$$

■ 단순이동평균 M_T의 평균과 분산

- 평균: $E(M_T) = E(X_t) = c$ → M_T : 상수 c의 불편추정량

- 분산: $Var(M_T) = \frac{1}{N^2}[NVar(X_t)] = \frac{\sigma_a^2}{N}$ → 시계열 분산의 $\frac{1}{N}$

⇒ 이동평균을 사용할 때 N값을 어떻게 정해야 하는가?

: 최근 추세를 보다 많이 반영하려면 작은 N값을 사용할 수 있고

평활효과를 높이려면 큰 N값을 사용할 수 있다.

■ 단순이동평균법

- 시계열 데이터 $\{X_1,\ X_2,\ \cdots\}$가 수평적인 패턴인 경우 사용하는 방법

시점	시계열	이동평균
$t-N+1$	X_{t-N+1}	
$t-N+2$	X_{t-N+2}	
$\vdots$	$\vdots$	
$t-1$	X_{t-1}	
t	X_t	M_t
$t+1$	X_{t+1}	M_{t+1}

- 시점 t에서의 단순이동평균: $M_t = \frac{1}{N}(X_{t-N+1} + \cdots + X_t)$

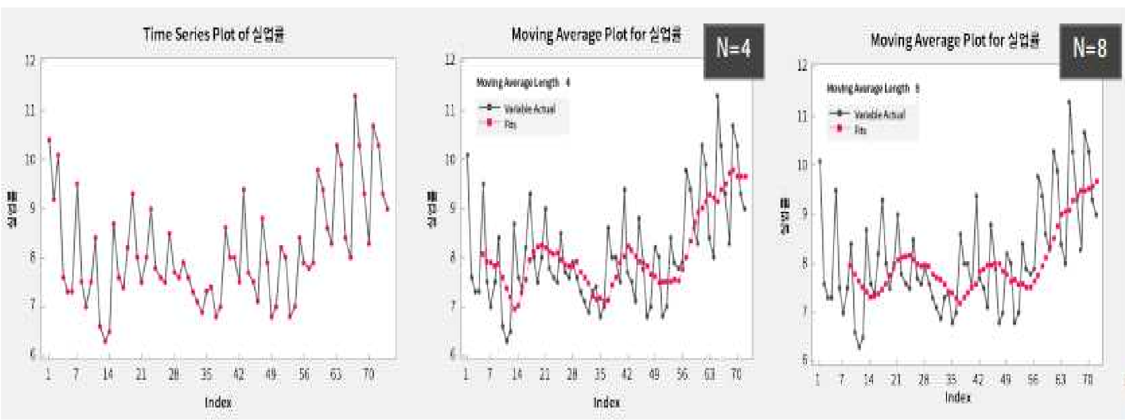
- 시점 $t+1$에서의 단순이동평균: $M_{t+1} = \frac{1}{N}(X_{t-N+2} + \cdots + X_{t+1})$

- 시점 T에서 시점 $T+1$의 값 예측(한 단계 이후 예측값): $f_{T,1} = M_T$

- N이 클수록 평활효과가 커진다.

■ 예: 2000년~2017년 분기별 청년(15~29세) 실업률

$N=4$와 $N=8$일 때의 단순이동평균



■ 일단계 이후 예측값에 대한 예측오차

$$e_{T,1} = X_{T+1} - \hat{f}_{T,1} = X_{T+1} - M_T$$

→ 예측오차의 분산

$$Var[e_{T,1}] = Var[X_{T+1} - M_T] = Var[X_{T+1}] + Var[M_T] = (1 + 1/N)\sigma_a^2$$

■ 시점 T에서 가장 나이브한 일단계 이후 예측값으로 현시점 값 X_T를 사용한다면 예측오차의 분산

$$Var[X_{T+1} - X_T] = Var[X_{T+1}] + Var[X_T] = 2\sigma_a^2 \geq (1 + 1/N)\sigma_a^2$$

⇒ 예측값으로 단순이동평균을 사용한 것이 나이브한 예측값보다 예측오차분산이 작다.

※ k단계 이후 예측오차에 대해서도 동일한 결과가 성립한다.

■ N값을 크게 할수록 예측오차분산이 작아지지만, 단기적 변동을 반영하지 못하는 단점이 있다.

1.2.2 선형추세와 이중이동평균

■ 시계열값 X_t가 다음과 같은 선형추세(linear trend)를 갖는다고 가정하면

$$X_t = c + bt + a_t$$

여기서, c, b: 상수, $a_t \sim i.i.d.\left(0,\ \sigma_a^2\right)$: 오차항, 백색잡음(white noise)

■ 단순이동평균 $M_T = \frac{1}{N}\sum_{t=T-N+1}^{T} X_t$

$$\begin{aligned} E[M_T] &= \frac{1}{N}\sum_{t=T-N+1}^{T} E[X_t] = \frac{1}{N}\sum_{t=T-N+1}^{T} E[c + bt + a_t] \\ &= c + bT - \frac{N-1}{2}b \end{aligned}$$

$$E\left[M_T\right] + \frac{N-1}{2}b = c + bT$$

→ 단순이동평균 M_T는 추세를 늦게 따라가므로, 이를 보정하기 위해 이중이동평균을 활용

■ 이중이동평균(double moving average)

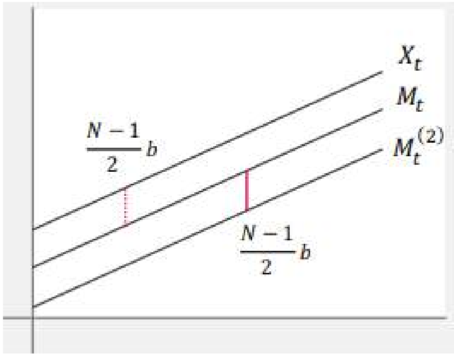
$$M_T^{(2)} = \frac{1}{N}\sum_{i=T-N+1}^{T} M_i$$

$$E[M_T]=\frac{1}{N}\sum_{t=T-N+1}^{T}E[X_t]=\frac{1}{N}\sum_{t=T-N+1}^{T}E[c+bt+a_t]$$
$$=c+bT-\frac{N-1}{2}b$$

$$E[M_T^{(2)}]=\frac{1}{N}\sum_{i=T-N+1}^{T}E[M_i]=\frac{1}{N}\sum_{i=T-N+1}^{T}E[c+bi-\frac{N-1}{2}b]$$
$$=c+bT-(N-1)b$$

$$\rightarrow\ E[M_T]-E\left[M_T^{(2)}\right]=\frac{N-1}{2}b$$

$$\Rightarrow\ \hat{b}=\frac{2}{N-1}\left[E(M_T)-E\left(M_T^{(2)}\right)\right],\quad \hat{c}=2M_T-M_T^{(2)}-\hat{b}\,T$$



■ 예측

- 시점 T에서 다음 시점의 예측치(한 단계 이후 예측)

$$f_{T,1}=E[X_{T+1}|X_T,\ X_{T-1},\ \cdots]=c+b(T+1)$$

$$\hat{f}_{T,1}=\hat{c}+\hat{b}(T+1)=(2+\frac{2}{N-1})M_T-(1+\frac{2}{N-1})M_T^{(2)}$$
$$=2M_T-M_T^{(2)}+\hat{b}$$

- k-단계 이후의 예측치

$$\hat{f}_{T,k}=\hat{c}+\hat{b}(T+k)=2M_T-M_T^{(2)}+k\hat{b},\ \ k=1,2,\cdots$$

■ 예: 1993년~2016년 특허건수

이동평균법과 이중이동평균법을 사용하여 시간에 따라 한 단계 이후 예측치 계산($N=4$)