STAT4005 - Time Series

Tutorial 1 - Decomposition of Time Series

Ma Ting Tin, Martin

Week 3 & 4

1 Decomposition of Time Series

In general, we can decompose a time series $\{X_t\}$ into three component, i.e. trend (T_t), seasonal effect (S_t) and noise (N_t). Explicitly, we have

$$\overbrace{X_t}^{\text{Observable}} = \overbrace{\underbrace{T_t + S_t}_{\text{(Non-random)}} + \underbrace{N_t}_{\text{(Random)}}}^{\text{Unobservable}}.$$

- X_t is the observation.
- T_t describes the general trend (non-seasonal).
- S_t (with period d) describes the seasonal effect, which is characterized by:
  1. S_i takes value from $\{s_1, \ldots, s_d\}$ with $\sum_{i=1}^{d} s_i = 0$.
  2. $S_i = s_i$ for $i = 1, \ldots, d$ and $S_{t+d} = S_t$ for all t.
- N_t is the noise, i.e., the random component, which satisfies $\mathrm{E}(N_t) = 0$.

Remark 1. *See a further discussion of the decomposition in Section 1.4.*

(★★☆) Identifying the decomposed component

Exercise 1. *Suppose that $X_t = T_t + S_t + N_t$, identify the explicit form of T_t, S_t, and N_t if*

(a) $X_t = \exp(t) + 2\mathbb{1}(t \text{ is odd}) - 2\mathbb{1}(t \text{ is even}) + a_t$, *where* $\mathrm{E}(a_t) = 0$.

(b) $X_t = 26 + 9t + 20\sin(\pi t/2) + a_t$, *where* $\mathrm{E}(a_t) = 0$.

Solution

1. $T_t = e^t$, $N_t = a_t$, and

$$S_t = \begin{cases} 2 & \text{, if } t \text{ is odd} \\ -2 & \text{, if } t \text{ is even} \end{cases}$$

2. $T_t = 26 + 9t$, $N_t = a_t$, and

$$S_t = \begin{cases} 0 & \text{, if } t = 4d \text{ or } t = 4d+2 \\ 20 & \text{, if } t = 4d+1 \\ -20 & \text{, if } t = 4d+3 \end{cases}$$

where d is an integer.

Remark 2. $\mathbb{1}(\cdot)$ *is known as the indicator (or characteristic) function, which is defined as*

$$\mathbb{1}(A) = \begin{cases} 1 & \text{, if } A \text{ is true;} \\ 0 & \text{, if } A \text{ is false.} \end{cases}$$

In particular, we could use

1. Least-square method,
2. Filtering method,
3. Differencing method

to deal with it. Furthermore, we may assume the existence (or absence) of seasonal effect and it would affect our inference procedure, i.e., there are totally **6 cases** to handle.

1.1 Least-square Method

The least-square method assumed a polynomial trend, i.e., $T_t = \beta_0 + \sum_{k=1}^{p} \beta_k t^k$, and estimate it by minimizing the sum of squared error.

1. **(L − S)** Assume that there is no seasonal effect.
   (a) Define the design matrix as
   $$\boldsymbol{X}_{-\boldsymbol{s}} = \begin{pmatrix} 1 & 1 & \cdots & 1^p \\ 1 & 2 & \cdots & 2^p \\ \vdots & \vdots & \ddots & \vdots \\ 1 & n & \cdots & n^p \end{pmatrix}, \quad \text{and} \quad \boldsymbol{Y} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}.$$
   (b) Estimate $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)$ by $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}_{-\boldsymbol{s}}^{\boldsymbol{T}} \boldsymbol{X}_{-\boldsymbol{s}})^{-1} \boldsymbol{X}_{-\boldsymbol{s}}^{\boldsymbol{T}} \boldsymbol{Y}$.
   (c) The **estimated trend**: $\hat{T}_t = \hat{\beta}_0 + \sum_{k=1}^{p} \hat{\beta}_k t^k$.
   (d) The **estimated noise**: $\hat{N}_t = X_t - \hat{T}_t$.
2. **(L + S)** Assume that there is a seasonal effect with period d.
   (a) Define (assuming $d = 3$ here for demonstration purposes) the design matrix as
   $$\boldsymbol{X}_{+\boldsymbol{s}} = \begin{pmatrix} 1 & 0 & 0 & 1 & \cdots & 1^p \\ 0 & 1 & 0 & 2 & \cdots & 2^p \\ 0 & 0 & 1 & 3 & \cdots & 3^p \\ 1 & 0 & 0 & 4 & \cdots & 4^p \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 1 & n & \cdots & n^p \end{pmatrix}, \quad \text{and} \quad \boldsymbol{Y} = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}.$$
   (b) Let $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_d)$ and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)$ (excluding β_0). Then $\boldsymbol{\alpha}$ and $\boldsymbol{\beta}$ could be estimated through
   $$(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}) = (\boldsymbol{X}_{+\boldsymbol{s}}^{\boldsymbol{T}} \boldsymbol{X}_{+\boldsymbol{s}})^{-1} \boldsymbol{X}_{+\boldsymbol{s}}^{\boldsymbol{T}} \boldsymbol{Y}.$$
   (c) Define $\bar{\alpha} = \sum_{i=1}^{d} \hat{\alpha}_i / d$. The **estimated seasonal effect** is $\hat{S}_i = \hat{\alpha}_i - \bar{\alpha}$.
   (d) The **estimated trend**: $\hat{T}_t = \bar{\alpha} + \sum_{k=1}^{p} \hat{\beta}_k t^k$.
   (e) The **estimated noise**: $\hat{N}_t = X_t - \hat{S}_t - \hat{T}_t$.

Remark 3. *There might be a slight confusion in the notation. Note that $\boldsymbol{X}$ ($\boldsymbol{X}_{-\boldsymbol{s}}$ or $\boldsymbol{X}_{+\boldsymbol{s}}$) refer to the* ***design/covariate matrix*** *instead of the time series vector $(X_1, \ldots, X_n)$. However, the statistical meaning of $\boldsymbol{X}$ would be clear according to the context.*

(★☆☆) Least Square Method

✎ **Exercise 2.** *Given observations* $\boldsymbol{x} = (10, 18, 32, 48)$.

(a) (L − S) *Suppose that there is no seasonal effect and that a* ***linear trend*** *is suitable. Estimate the trend using the least squares method.*

(b) (L + S) *Suppose that there are seasonal effects with period 2 and* ***a quadratic trend*** *is suitable. Let* $\boldsymbol{X}_{+s}$ *and* $\boldsymbol{Y}$ *be the design matrix and the response vector for applying the least squares method. We have*

$$(\boldsymbol{X}_{+s}^{\boldsymbol{T}}\boldsymbol{X}_{+s})^{-1}\boldsymbol{X}_{+s}^{\boldsymbol{T}}\boldsymbol{Y} = (5, 4, 3, 2)^T.$$

(i) Write down the explicit form of $\boldsymbol{X}_{+s}$ *and the estimated trend* $\hat{T}_t$.

(ii) What is the estimated seasonal effect $\hat{S}_t$*?*

(iii) Evaluate the value of $\hat{N}_3$.

Solution

(a) The design matrix and response vector are given by

$$\boldsymbol{X}_{-s} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} \quad \text{and} \quad \boldsymbol{Y} = \begin{pmatrix} 10 \\ 18 \\ 32 \\ 48 \end{pmatrix}$$

Therefore, $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}_{-s}^{\boldsymbol{T}}\boldsymbol{X}_{-s})^{-1}\boldsymbol{X}_{-s}^{\boldsymbol{T}}\boldsymbol{Y} = (-5, 64/5)^T$. The estimated trend is thus given by

$$\hat{T}_t = \hat{\beta}_0 + \hat{\beta}_1 t = -5 + \frac{64}{5}t.$$

(b) (i) The explicit form of $\boldsymbol{X}_{+s}$ is given by

$$\boldsymbol{X}_{+s} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 0 & 1 & 2 & 4 \\ 1 & 0 & 3 & 9 \\ 0 & 1 & 4 & 16 \end{pmatrix}.$$

Also, the estimated coefficients are given by $(\hat{\boldsymbol{\alpha}}, \hat{\boldsymbol{\beta}}) = (\boldsymbol{X}_{+s}^{\boldsymbol{T}}\boldsymbol{X}_{+s})^{-1}\boldsymbol{X}_{+s}^{\boldsymbol{T}}\boldsymbol{Y} = (5, 4, 3, 2)^T$. Therefore, $\hat{\alpha}_1 = 5,\ \hat{\alpha}_2 = 4,\ \hat{\beta}_1 = 3$ and $\hat{\beta}_2 = 2$. As $\bar{\alpha} = (\hat{\alpha}_1 + \hat{\alpha}_2)/2 = 9/2$, we have the estimated trend

$$\hat{T}_t = \bar{\alpha} + \hat{\beta}_1 t + \hat{\beta}_2 t^2 = \frac{9}{2} + 3t + 2t^2.$$

(ii) $\hat{S}_1 = \hat{\alpha}_1 - \bar{\alpha} = 1/2$ and $\hat{S}_2 = \hat{\alpha}_2 - \bar{\alpha} = -1/2$. Therefore,

$$\hat{S}_{2k-1} = \frac{1}{2} \quad \text{and} \quad \hat{S}_{2k} = -\frac{1}{2}$$

for $k \in \mathbb{N}$.

(iii) $\hat{N}_3 = X_3 - \hat{T}_3 - \hat{S}_3 = X_3 - \hat{T}_3 - \hat{S}_1 = 0$.

1.2 Filtering Method

Filtering is a non-parametric method that does not require us to specify the explicit form of T_t. First, we define a filter as follows.

Filter

Definition 1. $\{a_r\}_{r=-s}^{s}$ *is said to be a* ***filter*** *if it satisfies*

1. *(Symmetry):* $a_r = a_{-r}$ *for all* r.
2. *(Normalized):* $\sum_{r=-s}^{s} a_r = 1$.

The smoothed-series of $\{X_t\}$ *is defined by*

$$S_m(X_t) = \sum_{r=-s}^{s} a_r X_{t+r}.$$

The following is an important theorem justifying whether a filter can get rid of a polynomial trend.

Filtering Effect

Theorem 1. *A* p*-th order polynomial passes through a filter, i.e.*

$$S_m(P_t) = P_t \quad \textit{whenever } P_t = \sum_{k=0}^{p} c_k t^k$$

if and only if

$$\sum_{r=-s}^{s} a_r r^j = 0 \quad \textit{for } j = 1, \dots, p.$$

1. **(F − S)** When there is no seasonal effect, the **estimated trend** and **estimated noise** are given by

$$\hat{T}_t = S_m(X_t) \quad \text{and} \quad \hat{N}_t = X_t - \hat{T}_t$$

for $t = s+1, s+2, \dots, n-s$.

2. **(F + S)** When there is a seasonal effect with period d. In this case, we set the filter as

$$\{a_r\}_{r=-q}^{q} = \begin{cases} (1/d, 1/d, \dots, 1/d, 1/d) & , \text{ if } d = 2q+1 \text{ (odd)} \\ (1/2d, 1/d, \dots, 1/d, 1/2d) & , \text{ if } d = 2q \text{ (even)} \end{cases}$$

Then the **estimated trend** is given by $\hat{T}_t = S_m(X_t)$. Also, let n_i be the number of observations that belong to season i for $i = 1, \dots, d$. Then the **estimated seasonal effect** equals

$$\hat{S}_i = \frac{\sum_{t \text{ belongs to season } i}(D_t - \bar{D})}{n_i}, \quad D_t = X_t - \hat{T}_t, \quad \bar{D} = \frac{1}{n-2q}\sum_{t=q+1}^{n-q} D_t.$$

And finally, the **estimated noise** is $\hat{N}_t = X_t - \hat{S}_t - \hat{T}_t$.

(★★☆) Filtering Method (F − S)

✎ **Exercise 3.** *In addition to Spencer-type filters,* ***Henderson's weighted moving average*** *is also a popular class of filters. We denote Henderson's 5-point filter as*

$$(a_{-2}, a_{-1}, a_0, a_1, a_2).$$

It is known that $a_0 = 80/143$ and quadratic trend passes through $\{a_r\}_{r=-2}^{2}$.

(a) Find the value of a_{-2}, a_{-1}, a_1 and a_2.

(b) Find the maximum order of polynomial that can pass through $\{a_r\}_{r=-2}^{2}$.

Solution

(a) By symmetricity,

$$a_{-1} = a_1 \quad \text{and} \quad a_{-2} = a_2.$$

By the normalized property,

$$1 = \sum_{r=-2}^{2} a_r = a_0 + 2a_1 + 2a_2 = \frac{80}{143} + 2a_1 + 2a_2.$$

By Theorem 1, as a quadratic trend can pass through $\{a_r\}_{r=-2}^{2}$,

$$0 = \sum_{r=-2}^{2} r^2 a_r = 0^2 \times a_0 + 2 \times (1)^2 \times a_1 + 2 \times 2^2 \times a_2 = 2a_1 + 8a_2.$$

On solving, $a_1 = a_{-1} = 42/143$ and $a_2 = a_{-2} = -21/286$.

(b) By assumption, $\sum_{r=-2}^{2} a_r r^j = 0$ for $j = 1, 2$. We then check whether $\sum_{r=-2}^{2} r^3 a_r = 0$. While

$$\sum_{r=-2}^{2} r^3 a_r = (-2)^3 \times \left(-\frac{21}{286}\right) + (-1)^3 \times \left(\frac{42}{143}\right) + (1)^3 \times \left(\frac{42}{143}\right) + (2)^3 \times \left(-\frac{21}{286}\right) = 0.$$

Therefore, a cubic trend can also pass through $\{a_r\}_{r=-2}^{2}$. However,

$$\begin{aligned}&\sum_{r=-2}^{2} r^4 a_r \\ =&(-2)^4 \times \left(-\frac{21}{286}\right) + (-1)^4 \times \left(\frac{42}{143}\right) + (1)^4 \times \left(\frac{42}{143}\right) + (2)^4 \times \left(-\frac{21}{286}\right) = -\frac{252}{143} \neq 0.\end{aligned}$$

Therefore, the quartic trend cannot pass through $\{a_r\}_{r=-2}^{2}$ and the maximum order of polynomial that can pass through $\{a_r\}_{r=-2}^{2}$ equals 3.

❗ **Remark 4.** *Each filter has its own limitations and assumptions. A filter being able to get rid of higher-order polynomials does not necessarily imply that it is a "good" filter in statistical meaning.*

Email contact: martinmtt@link.cuhk.edu.hk ; Personal website: https://sites.google.com/view/martinmtt/home

The following shows the trade-off of filter between three factors:

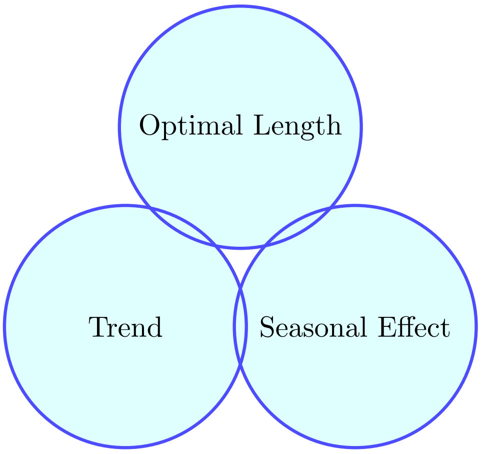
1. retain polynomial trend, i.e., $S_m(T_t) = T_t$.
   $\Rightarrow$ have to satisfy $\sum_{r=-s}^{s} a_r r^j = 0$ for $j = 0, \ldots, p$. (See Exercise 3)
   but it could not handle seasonal effect in general.

2. eliminate seasonal effect, i.e., $S_m(S_t) = 0$.
   $\Rightarrow$ have to satisfy $\sum_{r=-s}^{s} a_r S_{t+r} = 0$ for all t.
   $\Rightarrow$ Solution: evenly weighted filter (try to verify it)
   but it could not handle polynomial trend in general.

3. optimal filter length (need to satisfy additional equations for handling trend)

   - (Under absence of seasonal effect) Solve $\sum_{r=-s}^{s} a_r r^j = 0$ for is a $s+1$ unknown and $p+1$ equation problem, i.e., can pick filter with $s = p$.
   - (Under presence of seasonal effect) have to satisfy $\sum_{r=-s}^{s} a_r S_{t+r} = 0$ in addition.

   $\Rightarrow$ solution: see Exercise 5. It could handle both trend and seasonal effect by sacrificing length.



(★☆☆) Filtering Method (Implementation)

✎ **Exercise 4.** *Given observations* $\boldsymbol{x} = (-8, 6, 0, -2, 0, -2)$ *with seasonal effects of period* 2. *Suppose there is no trend, suggest a sensible filter and find* $\hat{T}_t, \hat{S}_t, \hat{N}_t$.

Solution

As the period $d = 2$ is even (then $q = 1$), we should set $(a_{-1}, a_0, a_1) = (1/4, 1/2, 1/4)$. Then we have

X_4	-8	6	0	-2	0	-2
$\hat{T}_t$	NA	1	1	-1	-1	NA
D_t	NA	5	-1	-1	1	NA
	$\bar{\boldsymbol{D}} = (5-1-1+1)/4 = 1$					
$D_t - \bar{\boldsymbol{D}}$	NA	4	-2	-2	0	NA
	$\hat{S}_1 = \frac{-2+0}{2} = -1$			$\hat{S}_2 = \frac{4-2}{2} = 1$		
$\hat{S}_t$	-1	1	-1	1	-1	1
$\hat{N}_t$	NA	4	0	-2	2	NA

(★★★) Simultaneously handling polynomial trend and seasonal effect through filtering

✎ **Exercise 5.** *Suppose that $X_t = T_t + S_t + N_t$, where $T_t = at^2 + b$ ($a \neq 0$) and S_t is of period $d = 2$. Consider the filter*

$$(a_{-2}, a_{-1}, a_0, a_1, a_2)$$

that satisfies $S_m(T_t) = T_t$ and $S_m(S_t) = 0$. Find the value of all unknowns.

Solution

Recall through symmetricity, we have $a_{-1} = a_1$ and $a_{-2} = a_2$. By the normalized property,

$$\sum_{r=-2}^{2} a_r = a_0 + 2a_1 + 2a_2 = 1. \tag{1}$$

As $S_m(T_t) = T_t$, i.e., quadratic trend would pass through the filter, we can write

$$\sum_{r=-2}^{2} a_r r^2 = 2a_1 + 8a_2 = 0. \tag{2}$$

Notice that as $S_{t-2} = S_t = S_{t-2}$ and $S_{t-1} = S_{t+1}$, we thus obtain

$$S_m(S_t) = (a_0 + 2a_2)S_t + (2a_1)S_{t+1}$$

In order to force $S_m(S_t) = 0$, recall that $S_t + S_{t+1} = 0$, as long as

$$a_0 + 2a_2 = 2a_1, \tag{3}$$

we have

$$S_m(S_t) = (a_0 + 2a_2)S_t + (2a_1)S_{t+1} = 2a_1(S_t + S_{t+1}) = 0.$$

Solving (1), (2) and (3) gives

$$a_0 = \frac{5}{8}, \qquad a_{-1} = a_1 = \frac{1}{4}, \quad \text{and} \quad a_{-2} = a_2 = -\frac{1}{16}.$$

1.3 Differencing Method

Recall the backshifting operator B is defined through $B^k X_t = X_{t-k}$. Then

- First-order differencing: $\Delta X_t = (1-B)X_t = X_t - X_{t-1}$.
- Second-order differencing:

$$\Delta^2 X_t = \Delta(\Delta X_t) = (1-B)(X_t - X_{t-1}) = (X_t - X_{t-1}) - (X_{t-1} - X_{t-2}) = X_t - 2X_{t-1} + X_{t-2}.$$

- k-th order differencing: $\Delta^k X_t = (1-B)^k X_t$.

Differencing can be used to remove trend and seasonal effects. In particular:

1. ***k*-th order differencing**: $\Delta^k X_t$ reduce the degree of polynomial trend by k.
2. **Seasonal differencing**: $\Delta_d X_t = (1-B^d)X_t = X_t - X_{t-d}$ will simultaneously
   - remove seasonal effect with period d.
   - reduce the degree of the polynomial trend by 1.

(★☆☆) Differencing Method (D − S)

✎ **Exercise 6.** *Given* $\boldsymbol{x} = (2, 0, 2, 3, 9, 3)$. *Assume* $T_t = \sum_{k=0}^{3} \beta_0 t^k$ *with* $\beta_3 \neq 0$ *and that there is no seasonal effect. Apply minimal order differencing to remove the trend and write down the differenced series.*

Solution

Third-order differencing is required to remove a cubic trend:

X_t	2	0	2	3	9	3
ΔX_t	NA	-2	2	1	6	-6
$\Delta^2 X_t$	NA	NA	4	-1	5	-12
$\Delta^3 X_t$	NA	NA	NA	-5	6	-17

(★☆☆) Differencing Method (D + S)

✎ **Exercise 7.** *Given* $\boldsymbol{x} = (x_1, \ldots, x_{45})$. *Assuming quadratic trend and seasonal effect with period 9.*

(a) How many times of differencing are needed to remove both the trend and seasonal effects?

(b) How many data points are left in the differenced series?

(c) Express $\Delta(\Delta_9 X_{20})$ *in terms of* $x_1, \ldots, x_{45}$.

Solution

(a) 2 times (1 seasonal differencing + 1 first-order differencing).

(b) $45 - 9 - 1 = 35$ data points left. (9 are dropped due to seasonal differencing, and one more is dropped due to first-order differencing).

(c) $\Delta(\Delta_9 X_{20}) = (x_{20} - x_{11}) - (x_{19} - x_{10})$.

1.4 ✄ Philosophical Discussion on Decomposition

In this section, we discuss the reason for decomposing the time series as $X_t = T_t + S_t + N_t$. There are two reasons to do so,

- In simple cases, the trend and seasonal effect can be identified easily and estimated accurately. Therefore, we do not want to build a too complicated model for model $\{X_t\}$, but instead only model $\{N_t\}$.

- The components T_t and S_t are **nonrandom** and the only randomness comes from N_t. Notice that once we obtained $\hat{T}_t$ and $\hat{S}_t$, we get the estimated noise $\hat{N}_t$ by

$$\hat{N}_t = X_t - \hat{T}_t - \hat{S}_t \quad \Rightarrow \quad \mathrm{Var}(X_t) = \mathrm{Var}(T_t + S_t + N_t) = \mathrm{Var}(N_t) \overset{(*)}{\approx} \mathrm{Var}(\hat{N}_t).$$

  Similarly, most inference concerning $\{X_t\}$ can be reduced to the study of $\{N_t\}$.

❶ **Remark 5.** *It is just a rough argument to understand the reason for performing the decomposition. The approximation* $(*)$ *does not hold in general for different reasons, e.g.:*

1. *The fluctuation of trend masks the noise.*
2. *The performance of the filtering method depends heavily on the choice of filter.*
3. *The least-square method relies on the parametric assumption, i.e., the trend is restricted to the class of p-th order polynomial.*

Therefore, using different methods to remove (or estimate) the trend and seasonal effect will lead to a totally different estimated noise sequence $\{\hat{N}_t\}$.

1.5 Additional Exercises

These additional exercises are designed to help to build a better understanding of the methodologies and clarify some common misconceptions.

(★☆☆) Filtering Method (F + S)

✎ **Exercise 8.** *Given observations* $\boldsymbol{x} = (-7, 9, 1, 5, 0, 4, 2, 3, -2)$ *with seasonal effects with period* 3, *suggest a sensible filter and find* $\hat{T}_t, \hat{S}_t, \hat{N}_t$.

Solution

As the period $d = 3$ is odd, we should set $(a_{-1}, a_0, a_1) = (1/3, 1/3, 1/3)$. Then we have

X_t	-7	9	1	5	0	4	2	3	-2
$\hat{T}_t$	NA	1	5	2	3	2	3	1	NA
D_t	NA	8	-4	3	-3	2	-1	2	NA
	$\bar{\boldsymbol{D}} = (8 - 4 + 3 - 3 + 2 - 1 + 2)/7 = 1$								
$D_t - \bar{\boldsymbol{D}}$	NA	7	-5	2	-4	1	-2	1	NA
	$\hat{S}_1 = \frac{2-2}{2} = 0$			$\hat{S}_2 = \frac{7-4+1}{3} = \frac{4}{3}$			$\hat{S}_3 = \frac{-5+1}{2} = -2$		
$\hat{S}_t$	0	$4/3$	-2	0	$4/3$	-2	0	$4/3$	-2
$\hat{N}_t$	NA	$20/3$	-2	3	$-13/3$	4	-1	$2/3$	NA

(★★☆) Comparison among Methods I

✎ **Exercise 9.** *Let* $\boldsymbol{x} = (x_1, \ldots, x_n)$ *be the observation and assume that there is no seasonal effect. Recall that we learned the least-square, differencing, and filtering methods in this tutorial.*

(a) *Suppose that you are not sure whether the trend* T_t *is a polynomial, then which method should NOT be applied?*

(b) *Suppose that you are not sure whether the trend* T_t *is a polynomial and you wish to obtain an estimated value of* N_t*, which method should be used?*

(c) *Suppose that you are confident that* $T_t = \sum_{k=0}^{p} \beta_k t^k$. *However, if you do not want to estimate the trend, but are only interested in the investigation of* $\{N_t\}$*, which method should be used?*

(d) *Suppose that you are confident that* $T_t = \sum_{k=0}^{p} \beta_k t^k$. *Say* n *is a small number and, hence, you want to retain as much data as possible for analysis, which method should be used?*

Solution

(a) Least-square method.

(b) Filtering method.

(c) Differencing method.

(d) Least-square method.

(★☆☆) Comparison among Methods II

✎ **Exercise 10.** *Let $\boldsymbol{x} = (x_1, \ldots, x_n)$ be the observation. Suppose that $T_t = \sum_{k=0}^{p} \beta_k t^k$ with $\beta_p \neq 0$ and the seasonal effect of period d exists. Compare the method mentioned in the following aspect.*

(a) Can we estimate trend and seasonal effects? If so, are those estimators well defined for all t?

(b) Is there any loss of data points? If yes, how many data points are lost?

Solution

(a) Both the least-square and filtering (but NOT differencing) methods can provide an estimate of the trend ($\hat{T}_t$) and seasonal effect ($\hat{S}_t$).

- For the least-square method, those estimators are well defined for $t = 1, \ldots, n$.
- For filtering method with filter $\{a_r\}_{r=-s}^{s}$, these estimators are well defined for $t = s+1, \ldots, n-s$.

(b)
- For the least-square method, there is no loss of data points.
- For the filtering method with filter $\{a_r\}_{r=-s}^{s}$, $2s$ data points are lost.
- In order to remove the seasonal effect, we first apply seasonal differencing (d data points are lost). The differenced series now has the $(p-1)$-order trend. Therefore, $(p-1)$-order differencing has to be applied. In total, $d + p - 1$ data points are lost.

(★★★) True/False Question

✎ **Exercise 11.** *Let $\boldsymbol{x} = (x_1, \ldots, x_n)$ be the observation. Suppose that $T_t = \sum_{k=0}^{p} \beta_k t^k$ and the seasonal effect of period d exists.*

(a) Andrew claimed that: "In order to remove the seasonal effect, we can apply seasonal differencing. And in addition, we have to apply p-th order differencing so that we can remove the degree p polynomial trend." Do you agree with his claim?

(b) Brian claimed that: "Using the least squares method, we estimate $(\boldsymbol{\alpha}, \boldsymbol{\beta})$, where $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)$ does not include the intercept term β_0. Therefore, it would be more general to define $\boldsymbol{\beta}$ as $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)$." Do you agree with his claim?

(c) Winky claimed that: "As the filter $(a_{-1}, a_0, a_1) = (-1/3, 5/3, -1/3)$ satisfies $\sum_{r=-1}^{1} a_r = 1$ and $\sum_{r=-1}^{1} a_r r^3 = 0$, a cubic polynomial would pass through it." Do you agree with her claim?

(d) Mathan claimed that: "Even the $(2p-1)$-order polynomial can pass through a filter $\{a_r\}_{r=-s}^{s}$, the $2p$-order polynomial may not pass through $\{a_r\}_{r=-s}^{s}$." Do you agree with his claim?

Solution

(a) No. (b) No. (c) No. (d) Yes.