STAT 4005 Time Series

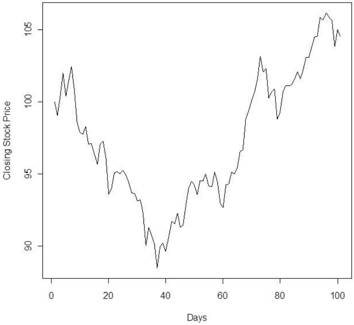
Chapter 1 - Introduction

Agenda

What is Time Series?

- A collection of values $X_t : t = 1, 2, ..., n$
  - t : time index (e.g. days, months, years)
  - X_t : observed data at different time (e.g. temperature,stock price)

Example of Time Series :



Distinctive feature of time series

- In classical setting (e.g. t-test, regression), we assume X_t's are independent
- In time series, we relax the independence assumption
  - Assume X_t's are dependent
  - The dependence in time series is quantified by auto-covariance

$$\boxed{\text{For small } k,\ \mathrm{Cov}(X_t, X_{t+k}) \neq 0}$$

Example 1 - Using correlation to find a correct C.I of μ

Model : $X_t = \mu + a_t - \theta a_{t-1}$ where $a_t \sim N(0,1)$

- $E(X_t) = \mu + E(a_t) - \theta E(a_{t-1}) = \mu$
- There exists time dependence because

$$\mathrm{Cov}(X_t, X_{t-k}) = \begin{cases} -\theta & k = \pm 1 \\ 1+\theta^2 & k = 0 \\ 0 & \text{otherwise} \end{cases}$$

- Approximate 95% Confidence interval for the mean μ

$$\overline{X} \pm 2\sqrt{Var(\overline{X})}$$

Proof.

By Central Limit Theorem (C.L.T.),

$$\frac{\overline{X}-\mu}{\sqrt{Var(\overline{X})}} \to N(0,1)$$

$$\therefore P\left(-2 \leq \frac{\overline{X}-\mu}{\sqrt{Var(\overline{X})}} \leq 2\right) \approx 0.95$$

$$\Rightarrow P\left(\overline{X} - 2\sqrt{Var(\overline{X})} \leq \mu \leq \overline{X} + 2\sqrt{Var(\overline{X})}\right) \approx 0.95 \qquad \square$$

Importance of Time Series

Recall 95% CI for μ is $\overline{X} \pm 2\sqrt{Var(\overline{X})}$. Note that

$$
\begin{aligned}
\mathrm{Var}(\overline{X}) &= \mathrm{Cov}\left(\frac{1}{n}\sum_{t=1}^{n} X_t, \frac{1}{n}\sum_{t=1}^{n} X_t\right) = \frac{1}{n^2}\sum_{t=1}^{n}\sum_{s=1}^{n}\mathrm{Cov}(X_t, X_s) \\
&= \frac{1}{n^2}\left(n\mathrm{Cov}(X_1, X_1) + 2(n-1)\mathrm{Cov}(X_2, X_1)\right) \\
&= \frac{1}{n^2}[n(1+\theta^2) - 2(n-1)\theta] = \frac{1}{n}\left[(1-\theta)^2 + \frac{2\theta}{n}\right].
\end{aligned}
$$

Therefore, a 95% C.I. is

$$
\overline{X} \pm \frac{2}{\sqrt{n}}\left[(1-\theta)^2 + \frac{2\theta}{n}\right]^{1/2}.
$$

- Without dependence ($\theta = 0$), 95% C.I. is $\overline{X} \pm \frac{2}{\sqrt{n}}$
- Need to incorporate correlation structure to produce a correct C.I.

Exercise

Model : $X_t = \mu + a_t - 0.5a_{t-1}$ where $a_t \sim N(0,1)$. Let $(X_1, X_2, X_3, X_4, X_5) = (4, 5, 6, 4, 1)$.

1. Find $\overline{X}$.
2. Find the autocovariance function $\gamma(k) = \mathrm{Cov}(X_t, X_{t-k})$.
3. Find $\mathrm{Var}(\overline{X})$.
4. Find the 95% C.I. for μ.

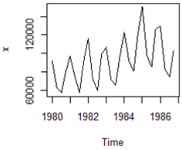
Agenda

Simple Descriptive Techniques

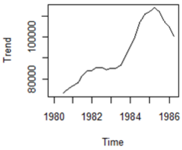
Decomposition of a time series

$$X_t = \underset{\text{(Trend)}}{T_t} + \underset{\text{(Seasonality)}}{S_t} + \underset{\text{(Noise)}}{N_t}$$

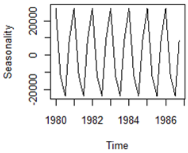
- T_t and S_t are Macroscopic Components
- N_t is Microscopic Component



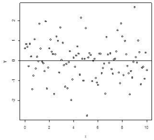
Series



Trend



Seasonality

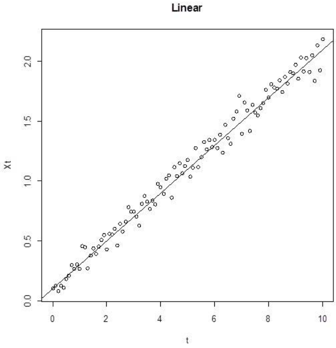


Noise

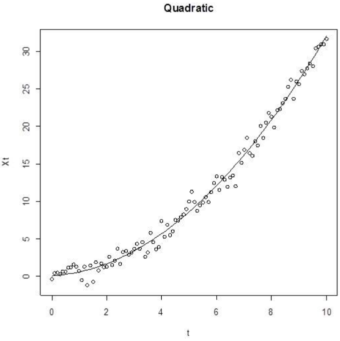
Trends

Examples

- Linear Trend : $T_t = \alpha + \beta t$
- Quadratic Trend : $T_t = \alpha + \beta t + \gamma t^2$



Linear Trend



Quadratic Trend

Outline

Estimation of trends without Seasonality

When $X_t = T_t + N_t$ (no seasonal effect), the trend can be estimated by

1) Least Squares Method
   - minimizes $\sum (X_t - T_t)^2$
   - T_t is a simple model of trend, e.g. $T_t = \alpha + \beta t$
2) Filtering
   - $\widehat{T}_t = S_m(X_t)$
   - S_m is a smoother: an operator that computes weighted average of observations near X_t
3) Differencing
   - $\Delta X_t = X_t - X_{t-1}$
   - Removing trend instead of estimating trend

1. Least Squares Method

- Idea : Find α , β_js in $T_t = \alpha + \beta_1 t + \cdots + \beta_k t^k$ such that

$$RSS = \sum_{t=1}^{n} (X_t - T_t)^2 \text{ is minimized}$$

- Method: Regression $Y = \mathbf{X}\beta + e \Rightarrow \hat{\beta} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Y$
  - Comparing $X_t = \alpha + \beta_1 t + \cdots + \beta_k t^k + N_t$ to $Y = \mathbf{X}\beta + e$

$$\Rightarrow \quad Y = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{pmatrix}; \quad \mathbf{X} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 2^k \\ \vdots & & \vdots & \\ 1 & n & \cdots & n^k \end{pmatrix}$$

$$\Rightarrow \quad \widehat{T} = \mathbf{X}\hat{\beta} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Y$$

  - R-Implementation:
    - `n=100;k=2;Y=3+2*(1:n)+10*rnorm(n)`
    - `X=rep(1,n);for (j in 1:k){ X=cbind(X,(1:n)∧(j))}`
    - `Trend.coef=solve(t(X)%*%X,t(X)%*%Y)`
    - `Trend=X%*%Trend.coef;`
      `ts.plot(Y);points(1:n,Trend,col=2,type='l')`
- Drawbacks
  - Only allow simple form of T_t, otherwise the minimization is difficult

1. Least Squares Method:

Exercise:

- Given the data $\{1.2, 2.1, 2.9, 3.8\}$, estimate the trend in the form $T_t = \alpha + \beta t$.

2. Filtering

- Idea : Smooth the series using local data to estimate the trend

$$\underbrace{\widehat{T}_t = S_m(X_t)}_{\text{“smoothed” series}} = \underbrace{\sum_{r=-s}^{s} a_r X_{t+r}}_{\text{Weighted average of } \{X_{t-s}, X_{t-s+1}, \ldots, X_{t+s}\}}$$

- Local: Window size s is much smaller than sample size n.
- Conditions on the weight $\{a_r\}$:
  1. Symmetric: $a_r = a_{-r}$
  2. Normalized: $\sum_{r=-s}^{s} a_r = 1$

Examples

1) Moving Average Filter

$$\widehat{T}_t = \frac{1}{2s+1} \sum_{r=-s}^{s} X_{t+r}$$

- What is $\{\widehat{T}_t\}$ if $\{X_t\} = \{1.1, 2.2, 2.7, 4.1, 5.2, 5.8\}$ and $s = 1$?

- If $X_t = \alpha + \beta t + N_t$ (Trend+Noise),

$$\begin{aligned} \widehat{T}_t &= S_m(X_t) = \frac{1}{2s+1} \sum_{r=-s}^{s} [\alpha + \beta(t+r) + N_{t+r}] \\ &\approx \alpha + \beta t \quad \left(\text{smoothing: } \frac{1}{2s+1} \sum_{r=-s}^{s} N_{t+r} \approx 0 \right) \end{aligned}$$

2. Filtering

Examples

2) Spencer 15-point filter:

$$(a_0, a_1, ..., a_7) = \frac{1}{320}(74, 67, 46, 21, 3, -5, -6, -3) \ , \quad a_r = a_{-r}$$

- Property: Does not distort a cubic trend:
  - If $X_t = T_t + N_t$, where $T_t = at^3 + bt^2 + ct + d$, then

$$\begin{aligned} \widehat{T}_t = S_m(X_t) &= \sum_{r=-7}^{7} a_r T_{t+r} + \sum_{r=-7}^{7} a_r N_{t+r} \\ &= at^3 + bt^2 + ct + d + \underline{\sum_{r=-7}^{7} a_r N_{t+r}} \\ &\approx T_t \qquad \left(\sum_{r=-7}^{7} a_r N_{t+r} \approx 0\right) \end{aligned}$$

  - We say: the cubic trend T_t "passes through" the filter S_m

2. Filtering

Exercise: Spencer 15-point filter:

$$(a_0, a_1, ..., a_7) = \frac{1}{320}(74, 67, 46, 21, 3, -5, -6, -3) \ , \quad a_r = a_{-r}$$

What is the filtered/smoothed series if the input is

- $X_t = ct$?
- $X_t = bt^2$?
- $X_t = at^3$?

2. Filtering

Theorem 1

A k^{th} order polynomial passes through a filter (i.e., $S_m(X_t) = \sum_{r=-s}^{s} a_r X_{t+r} = X_t$ for $X_t = c_0 + c_1 t + \cdots + c_k t^k$) if and only if

1. $\sum_{r=-s}^{s} a_r = 1$
2. $\sum_{r=-s}^{s} r^j a_r = 0$ *for* $j = 1, 2, ..., k$

- Application: Design a filter that passes through a quadratic trend.
    - Find $\{a_r\}$ such that $\sum_{r=-s}^{s} a_r = 1$, $\sum_{r=-s}^{s} r a_r = \sum_{r=-s}^{s} r^2 a_r = 0$.
    - Three equations $\Rightarrow$ Three unknowns $\Rightarrow$ Try $s = 1$, $\{a_{-1}, a_0, a_1\}$.
    - Not satisfied? Try larger s.

3. Differencing

- Differencing

$$\begin{aligned} \text{First Order:} \qquad \Delta X_t &= X_t - X_{t-1} \\ \text{Second Order:} \qquad \Delta^2 X_t &= \Delta(\Delta X_t) \end{aligned}$$

  - Definition: Backshift operator (B):
    - $BX_t = X_{t-1}$
    - $B^k X_t = X_{t-k}$, $k = 1, 2, \dots$
  - Definition: Differencing operator (Δ):
    - $\Delta X_t = (1 - B)X_t$
    - $\Delta^k X_t = (1 - B)^k X_t$, $k = 1, 2, \dots$
- Exercise: What is $\{\Delta X_t\}$ if $\{X_t\} = \{1.1, 2.2, 2.7, 4.1, 5.2, 5.8\}$

3. Differencing removes trend

- If $X_t = \alpha + \beta t$,
  - $\Delta X_t = X_t - X_{t-1} = \alpha + \beta t - [\alpha + \beta(t-1)] = \beta$ (no Trend!)

- If $X_t = t^p$,
  - $\Delta X_t = X_t - X_{t-1} = t^p - (t-1)^p = pt^{p-1} - C_2^p t^{p-2} + \cdots$

- In general, if $X_t = T_t + N_t$ and $T_t = \sum_{j=0}^{p} a_j t^j$,
  - $\Delta^p X_t = p! a_p + \Delta^p N_t$
  - If the trend is a p^{th} degree polynomial, then the trend can be eliminated in differencing p times.

- Example:
  - `n=100; t=1:n; x=5-2*t+3*t∧2-4*t∧3+10*rnorm(n)`
  - `d1=diff(x);d2=diff(d1);d3=diff(d2);d4=diff(d3);d5=diff(d4)`
  - `par(mfrow=c(2,3))`
  - `ts.plot(x);ts.plot(d1);ts.plot(d2);ts.plot(d3);ts.plot(d4);ts.plot(d5)`

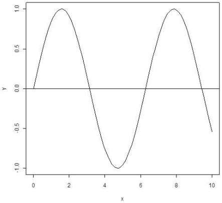
Outline

Seasonal Cycles

- General Decomposition

$$X_t = T_t + S_t + N_t$$

- Seasonal component S_t: period$=d$



  - Requirements
    1. $S_{t+d} = S_t$ (period d: repeating itself after time d)
    2. $\sum_{i=1}^{d} S_i = 0$ (common effect is explained by T_t)
  - e.g. season: d=4, month: d=12, week: d=7

Estimating/Removing Seasonal effect

Difficulties when both T_t and S_t exist:

- Need to separate the effect of trend T_t and seasonal effect S_t

Available Methods:

1. Least Squares Method
2. Filtering using Moving Average
3. Seasonal Differencing

1. Least Squares Method

- Idea :
  - Modeling seasonal effect: $S_t = \alpha_1 1_{\{s=1\}} + \alpha_2 1_{\{s=2\}} + \cdots + \alpha_d 1_{\{s=d\}}$
  - Modeling trend: $T_t = \beta_1 t + \cdots + \beta_k t^k$ (no constant term)
  - Find α_is , β_js such that

$$RSS = \sum_{t=1}^{n} (X_t - S_t - T_t)^2 \text{ is minimized}$$

- Method: $X_t = \alpha_1 1_{\{s=1\}} + \cdots + \alpha_d 1_{\{s=d\}} + \beta_1 t + \cdots + \beta_k t^k + N_t$
  - Regression: e.g. $d = 3$, $n = kd + 2$ for some integer k
    - $Y = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ \vdots \\ X_n \end{pmatrix}$; $\mathbf{X} = \begin{pmatrix} 1 & 0 & 0 & 1 & \cdots & 1 \\ 0 & 1 & 0 & 2 & \cdots & 2^k \\ 0 & 0 & 1 & 3 & \cdots & 3^k \\ 1 & 0 & 0 & 4 & \cdots & 4^k \\ \vdots & & & & \vdots & \\ 0 & 1 & 0 & n & \cdots & n^k \end{pmatrix}$

    $\Rightarrow \widehat{S} + \widehat{T} = \mathbf{X}\hat{\beta} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Y$
    - Let $\bar{\alpha} = \sum_{i=1}^{d} \hat{\alpha}_i / d$
    - Estimated Trend $\widehat{T_t}$: $\bar{\alpha} + \hat{\beta}_1 t + \cdots + \hat{\beta}_k t^k$
    - Estimated Seasonal Effect $\widehat{S_i}$: $\hat{\alpha}_i - \bar{\alpha}$ (so that $\sum_{i=1}^{d} \hat{S}_i = 0$)

$$X_t = \alpha_1 1_{\{s=1\}} + \cdots + \alpha_d 1_{\{s=d\}} + \beta_1 t + \cdots + \beta_k t^k + N_t$$

- - $Y = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ \vdots \\ X_n \end{pmatrix}$; $\mathbf{X} = \begin{pmatrix} 1 & 0 & 0 & 1 & \cdots & 1 \\ 0 & 1 & 0 & 2 & \cdots & 2^k \\ 0 & 0 & 1 & 3 & \cdots & 3^k \\ 1 & 0 & 0 & 4 & \cdots & 4^k \\ \vdots & & & & \vdots & \\ 0 & 1 & 0 & n & \cdots & n^k \end{pmatrix}$

  $\Rightarrow$ $\widehat{S} + \widehat{T} = \mathbf{X}\hat{\beta} = \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'Y$
  - Let $\bar{\alpha} = \sum_{i=1}^{d} \hat{\alpha}_i / d$
  - Estimated Trend $\widehat{T_t}$: $\bar{\alpha} + \hat{\beta}_1 t + \cdots + \hat{\beta}_k t^k$
  - Estimated Seasonal Effect $\widehat{S_i}$: $\hat{\alpha}_i - \bar{\alpha}$
- R-Implementation:
  - `s1=rep(c(1,0,0),33);s2=rep(c(0,1,0),33);s3=rep(c(0,0,1),33)`
  - `n=99;k=2;d=3; Y=3+2*(1:n)+9*s1-9*s3+10*rnorm(n)`
  - `X=cbind(s1,s2,s3);for (j in 1:k){ X=cbind(X,(1:n)∧(j))}`
  - `Reg.coef=solve(t(X)%*%X,t(X)%*%Y); a0=mean(Reg.coef[1:3])`
  - `Trend=a0+X[,4:5]%*%Reg.coef[4:5]`
  - `Season=X[,1:3]%*%Reg.coef[1:3]-a0`
  - `ts.plot(cbind(Y,Trend,Season,Trend+Season),col=1:4)`

2. Filtering by Moving Average

- STEP 1: Estimate the trend T_t by a special moving average filter
  - the filter must cover a complete cycle (length d) with equal weights

  $\Rightarrow$ the estimated trend is free from seasonal effect because $\sum_{j=1}^{d} S_j = 0$

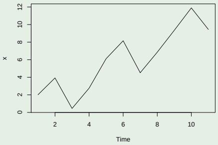
$$\widehat{T}_t = \begin{cases} \frac{1}{d}\left(\frac{1}{2}X_{t-q} + X_{t-q+1} + \ldots + \frac{1}{2}X_{t+q}\right) & , \quad \text{if } d = 2q \\ \frac{1}{d}\sum_{r=-q}^{q} X_{t+r} & , \quad \text{if } d = 2q+1 \end{cases}$$

- STEP 2: Estimate the seasonal component $(j = 1, \ldots, d)$

$$\widehat{S}_i = \frac{\sum_{t=i,d+i,2d+i,\ldots}(D_t - \overline{D})}{n_i}\,, \; D_t = X_t - \widehat{T}_t\,, \; \overline{D} = \frac{1}{n_d}\sum D_t\,,$$

  - n_i = number of season i observed
  - n_d = total number of D_ts
- RESULT: $X_t = \widehat{T}_t + \widehat{S}_t + \widehat{N}_t, \quad (\widehat{N}_t = X_t - \widehat{T}_t - \widehat{S}_t)$
- REMARK: May apply a better filter to $X_t - \widehat{S}_t$ to get an improved $\widetilde{T}_t$, then iterate Step 2 and improved filter until they converge.

Example 2



Consider the data set

$$(X_1, \ldots, X_{11}) = (2.1, 3.9, 0.5, 2.8, 6.1, 8.2, 4.5, 6.9, 9.3, 11.9, 9.4)$$

1. What is d?
2. Find $\widehat{T}$
3. Find $\widehat{S}_i$ for $i = 1, \ldots, d$

3. Seasonal Differencing

- Seasonal Differencing

$$\begin{aligned}\Delta^d X_t &= \left(1 - B^d\right) X_t \\ &= X_t - X_{t-d}\end{aligned}$$

- Seasonal differencing removes seasonal effects:
  If $X_t = S_t + N_t$ and period$= d$,
  then $\Delta^d X_t = S_t - S_{t-d} + N_t - N_{t-d} = N_t - N_{t-d}$. (Recall $S_t = S_{t-d}$)

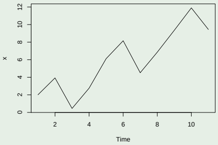
- Seasonal differencing also reduce polynomial trend by one degree:

$$\begin{aligned}\Delta^d t^p &= t^p - (t-d)^p = dt^{p-1} + \cdots \\ \Delta^d t &= t - (t-d) = d(\text{no } t)\end{aligned}$$

- Drawbacks
  1. Lose d data points
     - data $= (X_1, X_2, ..., X_n) \Rightarrow$ differenced data $= (\Delta X_{d+1}, ..., \Delta X_n)$
  2. No estimated seasonal effect $\widehat{S}_t$ is obtained.

3. Seasonal Differencing

Example 3



Consider the data set

$$(X_1, \ldots, X_{11}) = (2.1, 3.9, 0.5, 2.8, 6.1, 8.2, 4.5, 6.9, 9.3, 11.9, 9.4)$$

1. Draw the seasonal differenced series.

Agenda

Example in R

- Data :
  Quarterly operating revenues of Washington power company 80-86
  http://www.sta.cuhk.edu.hk/NHCHAN/TSBook2nd/dataset.html

- Step 0: Time Series Plot

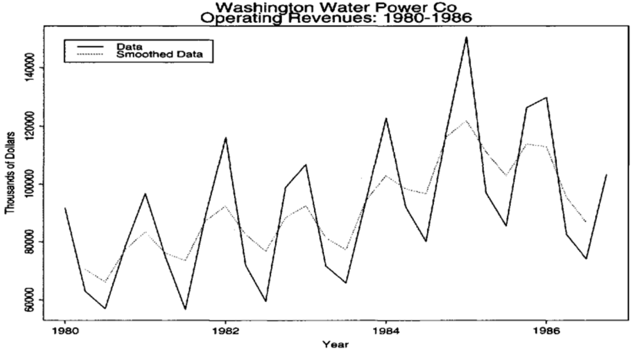


Fig. 1.2 Time series plots.

Step 0: Time Series Plot

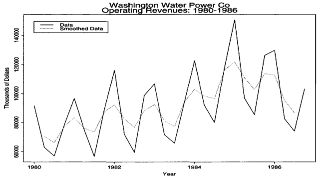


Fig. 1.2 Time series plots.

R-Codes for reading data and time series plot

```
x=read.delim("C://washpower.dat",head=FALSE)
x=ts(x, frequency = 4, start = c(1980, 1))
ts.plot(x,main="Wasington Water Power Co")
```

- Observations
  1) Slightly increasing trend
  2) Annual Cycle $(d = 4)$
     - Lower in summer
     - Higher in winter (heating?)
- Goal: Decompose $X_t = T_t + S_t + N_t$ by three methods

Method 1: Least Squares Method

Model: $$X_t = \alpha_1 1_{\{s=1\}} + \cdots + \alpha_4 1_{\{s=4\}} + \beta_1 t + \beta_2 t^2 + N_t$$

- R-Implementation:
  - Define seasonal indicator variables:

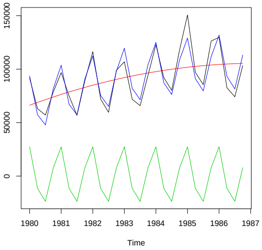
    ```
    s1=rep(c(1,0,0,0),7);s2=rep(c(0,1,0,0),7);s3=rep(c(0,0,1,0),7);s4=rep(c(0,0,0,1),7)
    ```

  - Define $\mathbf{X}$ matrix:

    ```
    X=cbind(s1,s2,s3,s4);for (j in 1:2){ X=cbind(X,(1:28)∧(j))}
    ```

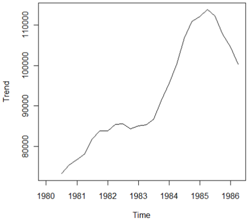
  - Least Squares fitting: `Reg.coef=solve(t(X)%*%X,t(X)%*%x);`
  - Results:

    ```
    a0=mean(Reg.coef[1:4]); T1=a0+X[,5:6]%*%Reg.coef[5:6];
    S1=X[,1:4]%*%Reg.coef[1:4]-a0; ts.plot(x,T1,S1,T1+S1,col=1:4)
    ```



- Step 1:
  Estimate trend $\widehat{T}_t$ by a filter covering a complete cycle $(d = 4)$

$$
\begin{aligned}
\widehat{T}_3 &= \frac{\frac{1}{2}X_1 + X_2 + X_3 + X_4 + \frac{1}{2}X_5}{4} \\
\widehat{T}_4 &= \frac{\frac{1}{2}X_2 + X_3 + X_4 + X_5 + \frac{1}{2}X_6}{4} \\
\cdots & \quad \cdots \\
\widehat{T}_{26} &= \frac{\frac{1}{2}X_{24} + X_{25} + X_{26} + X_{27} + \frac{1}{2}X_{28}}{4}
\end{aligned}
$$




```
n=length(x)
# make empty vector to store trend
T2=rep(NA,n)
filter=c(0.125,0.25,0.25,0.25,0.125)
# compute filtered values recursively
radius=2;start=3;end=n-2
for ( k in start:end) {
T2[k]=filter%*%x[(k-radius):(k+radius)]
}
ts.plot(T2,ylab="Trend")
```

Method 2: Filtering by Moving Average – Step 2

- Step 2: Estimate seasonal effect from the trend-removed series D_t

$$D_t = X_t - \widehat{T}_t\,, \quad \overline{D} = \frac{1}{24}\sum_{i=3}^{26} D_i$$

Estimating the seasonal part S_i

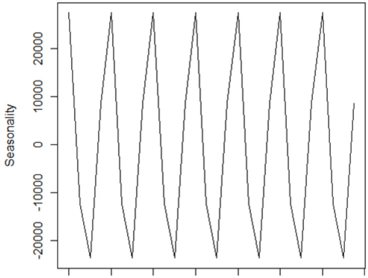
$$\widehat{S}_1 = [(D_5 - \overline{D}) + (D_9 - \overline{D}) + \cdots + (D_{25} - \overline{D})]/6$$

$$\widehat{S}_2 = [(D_6 - \overline{D}) + (D_{10} - \overline{D}) + \cdots + (D_{26} - \overline{D})]/6$$

$$\widehat{S}_3 = [(D_3 - \overline{D}) + (D_7 - \overline{D}) + \cdots + (D_{23} - \overline{D})]/6$$

$$\widehat{S}_4 = [(D_4 - \overline{D}) + (D_8 - \overline{D}) + \cdots + (D_{24} - \overline{D})]/6$$

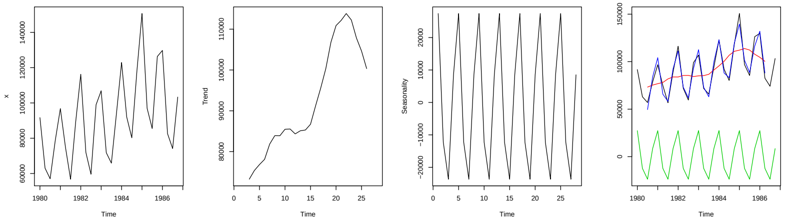
$$\Rightarrow S_{i+4j} = S_i, \quad \text{all } i, j$$




```
d=4;          # set period=4
D=x-T2;
# na.rm=T means remove NA entries
D.bar=mean(D,na.rm=T)
# data of same season put in same col.
S.mat=matrix(D-D.bar,ncol=d,byrow=T)
# compute column(2) mean
S2=apply(S.mat,2,mean,na.rm=T)
# make S the same length as x
S2=rep(S2,n/d)
ts.plot(S2,ylab="Seasonality")
```

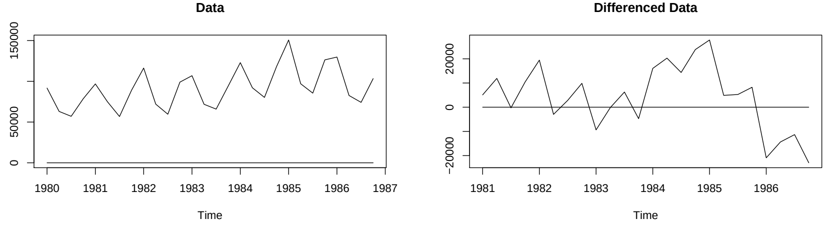
Method 2: Filtering by Moving Average – Results

```
par(mfrow=c(1,4))  # make 1×4 plots
ts.plot(x)
ts.plot(T2,ylab="Trend")
ts.plot(S2,ylab= "Seasonality")
ts.plot(x,T2,S2,T2+S2,col=1:4)
```



$$Y_t = (1 - B^4)X_t = X_t - X_{t-4}$$

```
dx=diff(x,4)
 par(mfrow=c(1,2))
 ts.plot(x,cex=2, main="Data")       #(cex controls the size of title)
 ts.plot(dx,cex=2,main="Differenced Data")
```



Comparing the three Methods

1. Method 1 (Least Squares) vs Method 2 (Filtering)
   - Mean Squared error (MSE):

```
E1=mean((x-T1-S1)∧ 2)
E2=mean((x-T2-S2)[3:26]∧ 2)
```

   $E1 = 65649889 > 28719343 = E2$

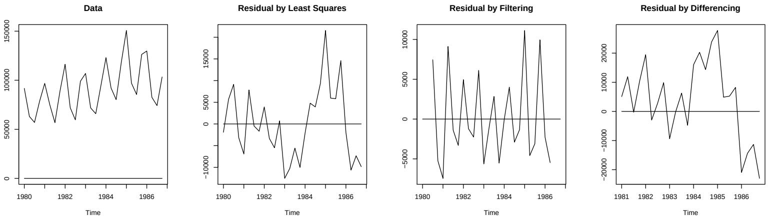
   - Method 1
     - uses a simpler model
     - allows extrapolation (predicting the future)
     - has no lost of data point
     - worse fit in terms of MSE
   - Method 2
     - uses a complicated model
     - does not allow extrapolation
     - loses 4 data points
     - better explains the data based on MSE.

2. Method 3 (Differencing) has no estimated Trend and Seasonal Effect

   $\Rightarrow$ Not directly comparable to Methods 1 and 2

   - Compare by Residual plots

```
par(mfrow=c(1,4))
ts.plot(x,cex=2,main="Data")
ts.plot(x-T1-S1,cex=2, main="Residual by Least Squares")
ts.plot(x-T2-S2,cex=2, main="Residual by Filtering")
ts.plot(dx,cex=2,main="Residual by Differencing")
```



   - Random Enough!

Agenda

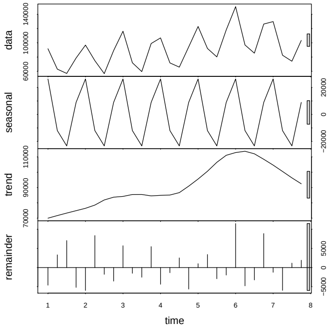
Command stl in R

```
stl(data,s.window="periodic")
```

- data is a "ts" object with a specified frequency (period)
- Method is similar to filtering with iterations between T_t and S_t

```
x=read.delim("C://washpower.dat",head=FALSE)
x=ts(x[1:28,1], frequency = 4)
decomp=stl(x,s.window="periodic")
plot(decomp)
```

```
# Store the results
trend=decomp$time.series[,2]
seasonal=decomp$time.series[,1]
y=decomp$time.series[,3]    # Noise
```



Agenda

Summary of Chapter 1

1) Time Series
   - Observations with time dependent structure
   - $\mathrm{Cov}(X_t, X_{t+k}) \neq 0$
2) Decomposition

$$X_t = \underset{\text{(Trend)}}{T_t} + \underset{\text{(Seasonality)}}{S_t} + \underset{\text{(Noise)}}{N_t}$$

   i) Removing Trend only
      - Least Squares Method: Minimizes $\sum_{t=1}^{n}(X_t - T_t)^2$
      - Filters: $\widehat{T}_t = S_m(X_t) = \sum_{r=-q}^{s} a_r X_{t+r}$
      - Differencing $\Delta X_t = (1 - B)X_t = X_t - X_{t-1}$
   ii) Removing Trend and Seasonality
      - Least Squares Method with seasonal indicators
      - Seasonal Filtering:
        Step 1: Moving average filter with length d to obtain $\widehat{T}_t$
        Step 2: Estimate $\widehat{S}_t$ from $X_t - \widehat{T}_t$
      - Seasonal Differencing $\Delta^d X_t = (1 - B^d)X_t = X_t - X_{t-d}$

Decomposition $X_t = \widehat{T}_t + \widehat{S}_t + \widehat{N}_t$

What to do after decomposition or differencing?

- Check the residuals
  - $\widehat{N}_t = X_t - \widehat{T}_t - \widehat{S}_t$ or
  - $\widehat{N}_t = X_t - X_{t-d}$

  to detect further structure
- Sophisticated models for the microscopic structure in $\widehat{N}_t$ will be discussed in later chapters