**Case Study – Time Series Analysis**

* 1. **Estimation and Removal of Deterministic Components**
     1. **Testing the Presence of Trend, it’s Estimation and Removal**
        1. **Relative Ordering Test for the Presence of Trend**

* + - 1. **Estimation using Least Squares**

**Case 1:** Test for the presence of trend and estimate it if it’s present for consumption expenditure (in million dollars) for the United States for 1944 to 2000 using appropriate test and method. Obtain the de-trended consumption series. Also provide a simple trend based forecast for the consumption expenditure for the next 5 years.

R Code and Output:

#case1

l<-read.csv("F:/R\_workbook/Programming and Predictive Modeling using R/3. Case Studies/7. Time Series Analysis/Data/case1.csv")

View(l)

plot(l,type="b")

ts<-ts(l,start=1947,end=2000)

class(ts)

plot.ts(ts,type="b")

library(devtools)

install\_github("abhishekumrawal/trendseason")

library(trendseason)

ro.test(ts)

t1<-time(ts)

lse<-lm(ts~t1)

abline(a=lse$coefficients[1],b=lse$coefficients[2])

lse\_trend<-lse$fitted.values

plot(lse\_trend)

detrended=as.vector(ts)-lse$fitted.values

plot(detrended)

library(forecast)

require(forecast)

forecast(lse$fitted.values,h=5)

Output:

> ro.test(ts)

Relative Ordering Test for Presence of Trend

Null Hypothesis: Absence of Trend, and

Alternative Hypothesis: Presence of Trend.

Test Statistic: 9.876

p\_value: 0

No. of Discordants: 1029

Expected No. of Discordants: 2889

> forecast(lse$fitted.values,h=5)

Year

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

55 2001 2001 2001 2001 2001

56 2002 2002 2002 2002 2002

57 2003 2003 2003 2003 2003

58 2004 2004 2004 2004 2004

59 2005 2005 2005 2005 2005

Consumption..C.

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

55 5462.495 5462.495 5462.495 5462.495 5462.495

56 5556.100 5556.100 5556.100 5556.100 5556.100

57 5649.706 5649.706 5649.706 5649.706 5649.706

58 5743.311 5743.311 5743.311 5743.311 5743.311

59 5836.916 5836.916 5836.916 5836.916 5836.916

Conclusion:

The forecasting value is given above.

**Case 2:** Test for the presence of trend and estimate it if it’s present for the following world development indicators for India: (time period)

1. Gross National Income (GNI) per capita based on Purchasing Power Parity (PPP) Exchange Rates (ER) measured in current USD,
2. Population Total,
3. Gross Domestic Product (GDP) (current USD),
4. Gross Domestic Product (GDP) Growth (annual %) and
5. Life Expectancy at birth (years)

Obtain the de-trended indicators.

R Code and Output:

#case2

l1<-read.csv("F:/R\_workbook/Programming and Predictive Modeling using R/3. Case Studies/7. Time Series Analysis/Data/case2.csv")

View(l1)

s<-l1[,c(1,2)]

c<-ts(s,start=2005,end=2013)

plot.ts(c,type="b")

ro.test(c)

t<-time(c)

ls1<-lm(c~t)

summary(ls1)

plot(ls1$fitted.values,type="b",pch=20)

#-------------------------------------------------------------------------------

s1<-l1[,c(1,3)]

c1<-ts(s1,start=2005,end=2013)

plot.ts(c1,type="b")

ro.test(c1)

t1<-time(c1)

ls2<-lm(c1~t1)

summary(ls2)

plot(ls2$fitted.values,type="b",pch=20)

#----------------------------------------------------------------------------------

s2<-l1[,c(1,4)]

c2<-ts(s2,start=2005,end=2013)

plot.ts(c2,type="b")

ro.test(c2)

t2<-time(c2)

ls3<-lm(c2~t2)

summary(ls3)

plot(ls3$fitted.values,type="b",pch=20)

#----------------------------------------------------------------------------------

s3<-l1[,c(1,5)]

c3<-ts(s3,start=2005,end=2013)

plot.ts(c3,type="b")

ro.test(c3)

t3<-time(c3)

ls4<-lm(c3~t3)

summary(ls4)

plot(ls4$fitted.values,type="b",pch=20)

#-----------------------------------------------------------------------------------

s4<-l1[,c(1,6)]

c4<-ts(s4,start=2005,end=2013)

plot.ts(c4,type="b")

ro.test(c4)

t4<-time(c4)

ls5<-lm(c4~t4)

summary(ls5)

plot(ls5$fitted.values,type="b",pch=20)

ouputs:

various plots were made to analyse the trend and the result of it are discussed in conclusion.

Conclusion:

Every variable has an upward trend except Gross Domestic Product (GDP) Growth (annual %).

**Estimation using Moving Averages**

**Case 3:** Test for the presence of trend and estimate it if it’s present for the annual sales measured in million USD for a trading company for 1994-2013. Obtain the de-trended sales. Also provide a simple trend based forecast for the annual sales for the next 3 years.

R code:

#case3

o<-read.csv("F:/R\_workbook/Programming and Predictive Modeling using R/3. Case Studies/7. Time Series Analysis/Data/case3.csv")

View(o1)

o1<-ts(o,start=1994,end=2013)

plot.ts(o1,type="b",pch=20)

ro.test(o1)

library(TTR)

MAtrend<-SMA(o1[,2],n=5)

plot((o1),type="l")

plot(MAtrend,type="l")

forecast(MAtrend,h=3)

Output:

> ro.test(o1)

Relative Ordering Test for Presence of Trend

Null Hypothesis: Absence of Trend, and

Alternative Hypothesis: Presence of Trend.

Test Statistic: -0.8622

p\_value: 0.1943

No. of Discordants: 427

Expected No. of Discordants: 390

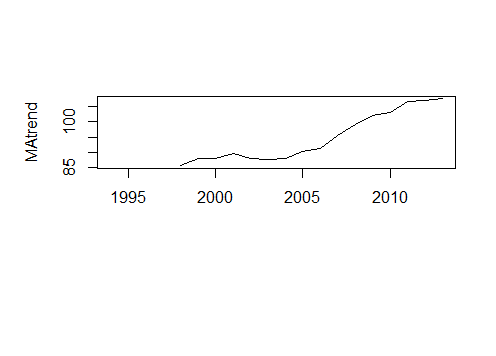
> forecast(MAtrend,h=3)

Point Forecast Lo 80 Hi 80 Lo 95 Hi 95

2014 108.3451 105.9813 110.7089 104.73000 111.9602

2015 108.7913 103.1509 114.4317 100.16506 117.4176

2016 109.2375 100.3705 118.1045 95.67665 122.7984



Conclusion:

The forecast is given above.

**Testing the Presence of Seasonality, it’s Estimation and Removal**

**Freidman’s Test for the Presence of Seasonality**

**Additive Decomposition using Least Squares**

**Case 4:** Test for the presence of trend and seasonality, and estimate them if they are present for the monthly World Airline Passengers from 1949-1960 using appropriate tests and methods. Obtain the additive decomposition of the original series viz. estimated trend, estimated seasonality, and estimated random component. Give a deterministic components based forecast for the monthly World Airline Passengers for the next 5 months.

R code:

#case4

k<-read.csv("F:/R\_workbook/Programming and Predictive Modeling using R/3. Case Studies/7. Time Series Analysis/Data/case4,5,7.csv")

View(k1)

l<-ts(k1,start=1949,end=1960,frequency=12)

plot.ts(l,type="b",pch=20)

ro.test(l)

ti<-time(l)

s<-lm(l~ti)

summary(s)

plot(s$fitted.values)

detren<-as.vector(l)-s$fitted.values

plot(detren)

library(trendseason)

friedman.test(detren,12)

require(forecast)

month<-seasonaldummy(l)

m<-lm(l~ti+month)

output:

> ro.test(l)

Relative Ordering Test for Presence of Trend

Null Hypothesis: Absence of Trend, and

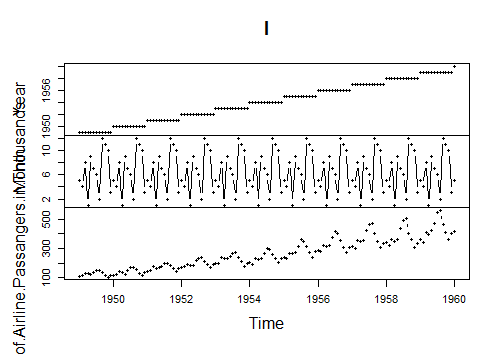
Alternative Hypothesis: Presence of Trend.

Test Statistic: -1

p\_value: 0.1587

No. of Discordants: 1

Expected No. of Discordants: 0.5



Conclusion:

Here we find a trend, estimate and eliminate that and then check for seasonality on detrended data.

**Additive Decomposition using STL in R**

**Case 5:** Test for the presence of trend and seasonality, and estimate them if they are present for the monthly World Airline Passengers from 1949-1960 using appropriate tests and methods. Obtain the additive decomposition of the original series viz. estimated trend, estimated seasonality, and estimated random component. Give a deterministic components based forecast for the monthly World Airline Passengers for the next 5 months.

R code:

#case5

View(k1)

l<-ts(k1,start=1949,end=1960,frequency=12)

plot.ts(l,type="b",pch=20)

ro.test(l)

ti<-time(l)

s<-lm(l~ti)

summary(s)

plot(s$fitted.values)

detren<-as.vector(l)-s$fitted.values

plot(detren)

library(trendseason)

friedman.test(detren,12)

decomp<-stl(l,t.window =12,s.window = 12 )

forecast(decomp)

Output:

Response Year :

Call:

lm(formula = Year ~ ti)

Residuals:

Min 1Q Median 3Q Max

-0.49007 -0.24161 0.00685 0.25530 0.48883

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) 11.608574 15.405714 0.754 0.452

ti 0.993828 0.007882 126.086 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 0.2908 on 131 degrees of freedom

Multiple R-squared: 0.9918, Adjusted R-squared: 0.9918

F-statistic: 1.59e+04 on 1 and 131 DF, p-value: < 2.2e-16

Response No..of.Airline.Passangers.in.Thousands :

Call:

lm(formula = No..of.Airline.Passangers.in.Thousands ~ ti)

Residuals:

Min 1Q Median 3Q Max

-86.688 -25.704 -4.379 20.729 139.287

Coefficients:

Estimate Std. Error t value Pr(>|t|)

(Intercept) -59739.522 2219.247 -26.92 <2e-16 \*\*\*

ti 30.700 1.135 27.04 <2e-16 \*\*\*

---

Signif. codes: 0 ‘\*\*\*’ 0.001 ‘\*\*’ 0.01 ‘\*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 41.9 on 131 degrees of freedom

Multiple R-squared: 0.848, Adjusted R-squared: 0.8469

F-statistic: 731 on 1 and 131 DF, p-value: < 2.2e-16

Conclusion:

As we can see that response variable is highly dependent on time, we find the trend and eliminate it and after that

We decompose the time series using stl().

**Case 6:** Test for the presence of trend and seasonality, and estimate them if they are present for the quarterly demand for an industrial good measured in thousand units for a manufacturing company for 2001-2005 using appropriate tests and methods. Obtain the additive decomposition of the original series viz. estimated trend, estimated seasonality, and estimated random component. Give a deterministic components based forecast for the quarterly demand for the industrial good for the next 2 quarters.

R code:

#case6

k<-read.csv("F:/R\_workbook/Programming and Predictive Modeling using R/3. Case Studies/7. Time Series Analysis/Data/case6.csv")

View(k)

tk<-ts(k,start=2005,end=2014,frequency=12)

plot.ts(tk)

ro.test(tk)

library(trendseason)

friedman.test(tk,4)

deco<-decompose(tk,type = "additive")

plot(deco)

Output:

> friedman.test(tk,4)

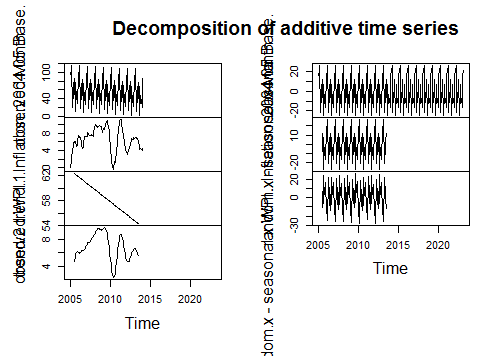
Freidman (JASA) Test for Presence of Seasonality

Null Hypothesis: Absence of Seasonality, and

Alternative Hypothesis: Presence of Seasonality.

Test Statistic: 18.4789 (Chi Sqaure with 3 df)

p\_value: 4e-04



Conclusion:

The decomposition of various deterministic parts are shown above.

* + - 1. **Additive Decomposition using Decompose in R**

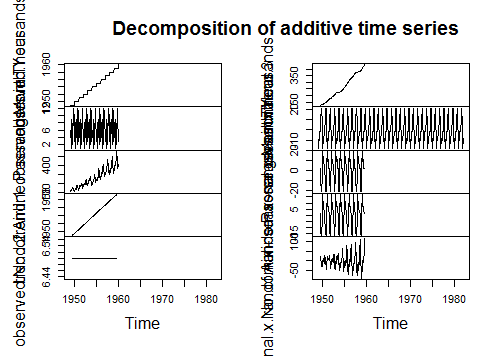
**Case 7:** Test for the presence of trend and seasonality, and estimate them if they are present for the monthly World Airline Passengers from 1949-1960 using appropriate tests and methods. Obtain the additive decomposition of the original series viz. estimated trend, estimated seasonality, and estimated random component. Give a deterministic components based forecast for the monthly World Airline Passengers for the next 5 months.

R code:

data.dec<-decompose(l,type = "additive")

plot(data.dec)

output:



Conclusion:

The deterministic parts are shown above.

**Case 8:** Test for the presence of trend and seasonality, and estimate them if they are present for the monthly Wholesale Price Index (WPI) – Inflation, Base year 2004-05 for India using appropriate tests and methods. Obtain the additive decomposition of the original series viz. estimated trend, estimated seasonality, and estimated random component. Give a deterministic components based forecast for the monthly Whole Sale Price Index for the next 5 months.

R code:

#case8

c<-read.csv("F:/R\_workbook/Programming and Predictive Modeling using R/3. Case Studies/7. Time Series Analysis/Data/case8.csv")

View(c)

tm<-ts(c,start=2001,end=2005,frequency=4)

plot.ts(tm)

ro.test(tm)

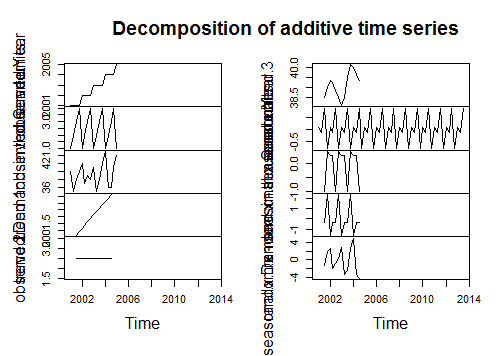
library(trendseason)

friedman.test(tm,4)

dec<-decompose(tm,type="additive")

plot(dec)

Output:



Conclusion:

The deterministic components are shown above.

* + - 1. **Comparison of different methods of Additive Decomposition in R**

**Case 9:** Compare the solutions for Case 4, Case 5 and Case 7.

Ans: here, we can see that additive decomposition using stl is better than additive decomposition using least squares and additive decomposition using decompose() itself. Hence, we use stl().