Time Series Analysis

Basics of Time Series Analysis

Nicoleta Serban, Ph.D.
Professor
Stewart School of Industrial and Systems Engineering

Decomposition: Seasonality Estimation Data Examples

Georgia Tech

About This Lesson

Data Example: Temperature in Atlanta, Georgia

Data: Average monthly temperature records starting in 1879 until 2016.

- Available from the iWearherNet.com
- The Weather Bureau (now the National Weather Service) began keeping weather records for Atlanta 138 years, 8 months and 19 days ago on October 1, 1878.
- Provided in Fahrenheit degrees

Are there seasonality and trend in the Atlanta temperature over the past 100 or more years?

Seasonality: Seasonal Models

```
library(dynlm)
## Estimate seasonality using Seasonal Means Model
## Drop January/with intercept
model1 = dynlm(temp~season(temp))
summary(model1)
## Seasonal mean effects/without intercept
model2 = dynlm(temp~ season(temp)-1)
summary(model2)
## Estimate seasonality using cos-sin model
model3=dynlm(temp~harmon(temp))
summary(model3)
model4=dynlm(temp~harmon(temp,2))
summary(model4)
```

Georgia Tech

Seasonality: Seasonal Means Model

	Estimate	Std. Error	t value	Pr(>|t|)
monthJanuary	43.2072	0.2725	158.5	<2e-16
monthFebruary	45.9587	0.2725	168.6	<2e-16
monthMarch	53.2304	0.2725	195.3	<2e-16
monthApril	61.6087	0.2725	226.1	<2e-16
monthMay	69.7696	0.2725	256.0	<2e-16
monthJune	76.6986	0.2725	281.4	<2e-16
monthJuly	79.0051	0.2725	289.9	<2e-16
monthAugust	78.2703	0.2725	287.2	<2e-16
monthSeptember	73.2986	0.2725	268.9	<2e-16
monthOctober	62.9616	0.2725	231.0	<2e-16
monthNovember	52.5493	0.2725	192.8	<2e-16
monthDecember	45.0725	0.2725	165.4	<2e-16

Multiple R-squared: 0.9975, Adjusted R-squared: 0.9974

$\hat{\mu}_{January} = 43.02$
$\hat{\mu}_{February} = 45.95$
$\hat{\mu}_{March} = 53.23$
$\hat{\mu}_{April} = 61.61$
$\hat{\mu}_{May} = 69.77$
$\hat{\mu}_{June} = 76.70$
$\hat{\mu}_{July} = 79.00$
$\hat{\mu}_{August} = 78.27$
$\hat{\mu}_{September} = 73.30$
$\hat{\mu}_{October} = 62.96$
$\hat{\mu}_{November} = 52.55$
$\hat{\mu}_{December} = 45.08$

$R^2 \sim 99.7\%$ *variability explained*

Georgia Tech

Seasonality: Cos-Sin Model

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	61.80254	0.08133	759.870	< 2e-16
har2cos(2*pi*t)	-18.30228	0.11502	-159.119	< 2e-16
har2sin(2*pi*t)	-0.69366	0.11502	-6.031	2.01e-09

Multiple R-squared: 0.9388, Adjusted R-squared: 0.9387

$R^2 \sim 93.9\%$ *variability explained*

	Estimate	Std. Error	t value	Pr(>|t|)
(Intercept)	61.80254	0.07888	783.490	< 2e-16
har2cos(2*pi*t)	-18.30228	0.11155	-164.065	< 2e-16
har2cos(4*pi*t)	-0.63031	0.11155	-5.650	1.88e-08
har2sin(2*pi*t)	-0.69366	0.11155	-6.218	6.36e-10
har2sin(4*pi*t)	0.96246	0.11155	8.628	< 2e-16

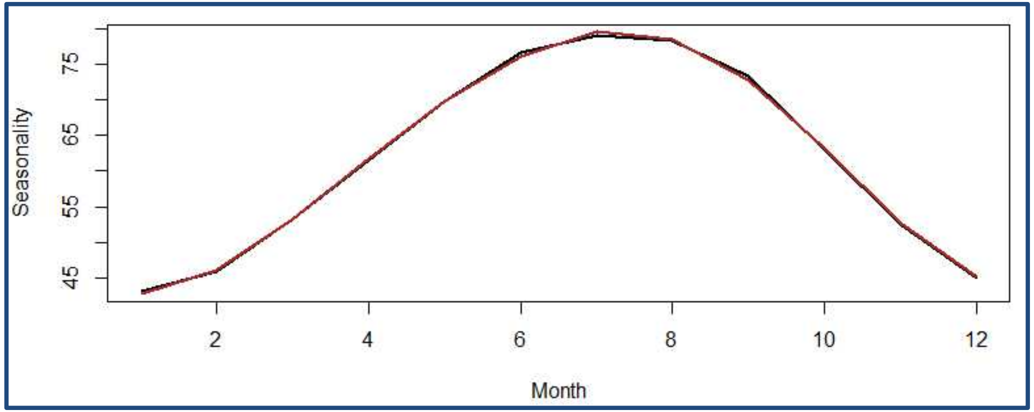
Multiple R-squared: 0.9425, Adjusted R-squared: 0.9424

$\hat{\beta}_{\cos(2pif)}$ = -18.30
$\hat{\beta}_{\sin(2pif)}$ = -0.693
$\hat{\beta}_{\cos(2pif/2)}$ = -0.630
$\hat{\beta}_{\sin(2pif/2)}$ = 0.962

$R^2 \sim 94.3\%$ *variability explained*

Seasonality: Compare Models

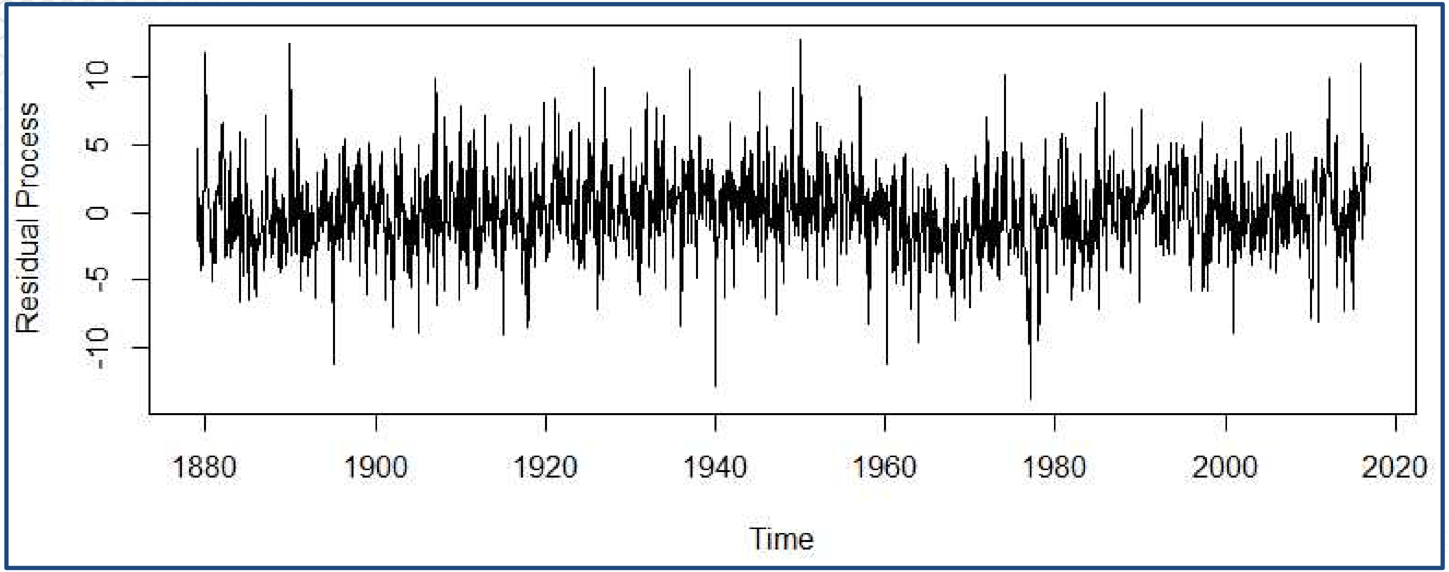
```
## Seasonal Means Model
st1 = coef(model2)
## Cos-Sin Model
st2 = fitted(model4)[1:12]
## Compare Seasonality Estimates
plot(1:12,st1,lwd=2,type="l",xlab="Month",ylab="Seasonality")
lines(1:12,st2,lwd=2, col="brown")
```



Seasonality & Trend: Parametric Model

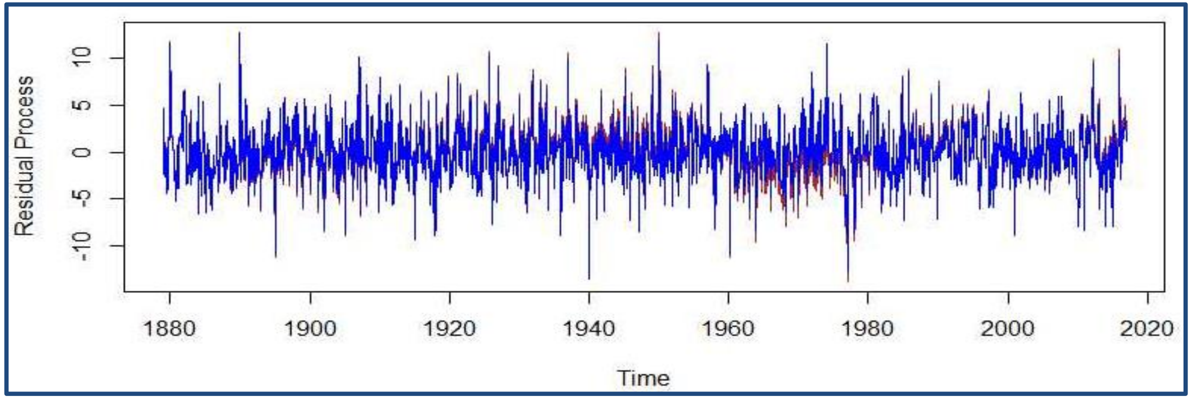
```
## Fit a parametric model for both trend and seasonality
x1 = time.pts
x2 = time.pts^2
lm.fit = dynlm(temp~x1+x2+harmon(temp,2))
summary(lm.fit)
dif.fit.lm = ts((temp-fitted(lm.fit)),start=1879,frequency=12)
ts.plot(dif.fit.lm,ylab="Residual Process")
```

Seasonality & Trend: Parametric Model



Seasonality & Trend: Compare Models

```
## Fit a non-parametric model for trend and linear model for seasonality
har2 = harmonic(temp,2)
gam.fit = gam(temp~s(time.pts)+har2)
dif.fit.gam = ts((temp-fitted(gam.fit)),start=1879,frequency=12)
## Compare approaches
ts.plot(dif.fit.lm,ylab="Residual Process",col="brown")
lines(dif.fit.gam,col="blue")
```



Georgia Tech

Summary