r-statistics.co (/) by Selva Prabhakaran

Search..

Tutorial

R Tutorial (R-Tutorial.html)

ggplot2

ggplot2 Short Tutorial (ggplot2-Tutorial-With-R.html)

ggplot2 Tutorial 1 - Intro (Complete-Ggplot2-Tutorial-Part1-With-R-Code.html)

ggplot2 Tutorial 2 - Theme (Complete-Ggplot2-Tutorial-Part2-Customizing-Theme-With-R-Code.html)

ggplot2 Tutorial 3 - Masterlist (Top50-Ggplot2-Visualizations-MasterList-R-Code.html)

ggplot2 Quickref (ggplot2-cheatsheet.html)

Foundations

Linear Regression (Linear-Regression.html)

Statistical Tests (Statistical-Tests-in-R.html)

Missing Value Treatment (Missing-Value-Treatment-With-R.html)

Outlier Analysis (Outlier-Treatment-With-R.html)

Feature Selection (Variable-Selection-and-Importance-With-R.html)

Model Selection (Model-Selection-in-R.html)

Logistic Regression (Logistic-Regression-With-R.html)

Advanced Linear Regression (Environments.html)

Advanced Regression Models

Advanced Regression Models (adv-regression-models.html)

Time Series

Time Series Analysis (Time-Series-Analysis-With-R.html)

Time Series Forecasting (Time-Series-Forecasting-With-R.html)

More Time Series Forecasting (Time-Series-Forecasting-With-R-part2.html)

High Performance Computing

Parallel computing (Parallel-Computing-With-R.html)

Strategies to Speedup R code (Strategies-To-Improve-And-Speedup-R-Code.html)

Useful Techniques

Association Mining (Association-Mining-With-R.html)

Multi Dimensional Scaling (Multi-Dimensional-Scaling-With-R.html)

Optimization (Profiling.html)

InformationValue package (Information-Value-With-R.html)

Stay up-to-date. Subscribe!
(https://docs.google.com/forms/d/1xkMYkLNFU9U39Dd8S_2JC0p8B5t6_Yq6zUQjanQQJpY/viewform)

Chat! (https://docs.google.com/forms/d/13GrkCFcNa-TOIllQghsz2SIEbc-YqY9eJX02B19l5Ow/viewform)

Contents

Time Series Forecasting

> This is a follow-up to the introduction to time series analysis (Time-Series-Analysis-With-R), but focused more on forecasting rather than analysis.

Simple Moving Average

Simple moving average can be calculated using `ma()` from forecast

```
sm <- ma(ts, order=12) # 12 month moving average
lines(sm, col="red") # plot
```

Exponential Smoothing

Simple, Double and Triple exponential smoothing can be performed using the HoltWinters() function. There are multiple implementations of the Holt Winters method – `hw()` {forecast} and `ets()`.

```
library(forecast)

# Simple exponential smoothing: Level Only
model <- hw(trainingData, initial = "optimal", h=(forecastPeriodLen), beta=NULL, gamma=N
ULL) # h is the no. periods to forecast

# Double Exponential smoothing: Level and Trend components
model <- hw(trainingData, initial = "optimal", h=(forecastPeriodLen), gamma=NULL)

# Holt Winters: Level, Trend and Seasonality
model <- hw(trainingData, initial = "optimal", h=(forecastPeriodLen))
plot(model)
accuracy(model) # calculate accuracy measures
```

ARIMA

The forecast package offers auto.arima() function to fit ARIMA models. It can also be manually fit using Arima(). A caveat with ARIMA models in R is that it does not have the functionality to fit long seasonality of more than 350 periods eg: 365 days for daily data or 24 hours for 15 sec data.

```
# Fit and forecast with auto.arima()
autoArimaFit <- auto.arima(tsData)
plot(forecast(autoArimaFit, h=20))

# Fit and forecast with Arima()
arimaFit <- Arima(tsData,order=c(3,1,0))
plot(forecast(arimafit,h=20))
```

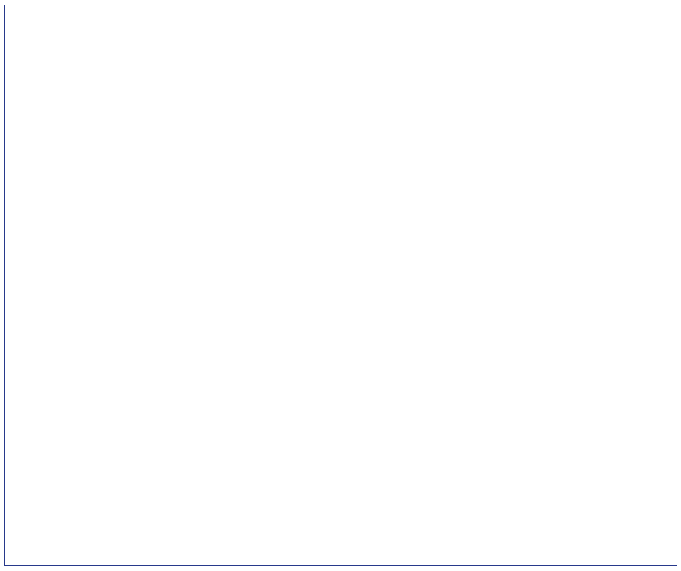
How To Forecast ARIMA Models With Long Seasonality (Greater Than 350 Periods)?

Upon plotting your Arima() forecast, you find a more or less flat forecast, it could be because of long seasonality. In such case, you can feed in the seasonality as an external regressor through the 'xreg' argument.

```
Fit <- Arima(tsData,order=c(3,1,0))  # fit Arima model
Fit <- auto.arima(tsData, seasonal=FALSE, xreg=fourier(tsData,4))  # fit auto.arima mode
l
plot(forecast(Fit,h=20))
pred <- predict (Fit, newxreg=newXregVar) # alternate way to forecast
plot(forecast(fit, h=h, xreg=fourierf(tsData,4,h))) # h is number of forecasts
```

If you are using a numeric vector as an external regressor (xreg), make sure you change it to a data.frame() before feeding it as an xreg parameter to auto.arima(). You can also use multiple external regressors by binding them together as a data.frame().

Some Useful External Regressors For Arima() and auto.arima()



Any dataframe with as many rows as length of ts data can be used as 'xreg' argument. A couple of common 'xreg's that are used to model seasonal effects are below.

```
Xreg1 <- seasonaldummy(tsData) # creates dummy binary variable for each period in a seas
on.
Xreg2 <- model.matrix(~ as.factor(weekday) + 0)) # weekday could be a monthday, hour-of-
day, holiday indicator etc ..
```

How To Model Time Series With Complex Seasonality Pattern?

Use the tbats() in forecast package. Time series with multiple-seasonality can be modelled with this method. Since this is a computationally intensive procedure, the in-built parallel processing facility may be leveraged.

```
tbatsFit <- tbats(tsData, use.parallel=TRUE, num.cores = 2) # fit tbats model
plot(forecast(fit)) # plot
components <- tbats.components(tbatsFit)
plot(components)
```

How To Find Confidence Intervals For My Forecasts?

The predict() function has the facility. By providing the argument 'prediction.interval=TRUE' and 'level = n', the prediction intervals for a given confidence is calculated. Below is a general format of the code.

```
model <- HoltWinters(TS) predict(model, 50, prediction.interval = TRUE, level= 0.99)  #
 prediction.interval = TRUE
```

More Useful Functions Related To Time Series

Functions	Description
accuracy()	accuracy measures of forecast

Functions	Description
BoxCox, invBoxCox()	Box-Cox transformation
decompose()	Decompose time series data into components
dm.test()	Diebold-Mariano test compares the forecast accuracy
monthdays()	number of days in seasonal series
na.interp()	interpolate missing values
seasadj()	Remove the seasonal components from a time series
seasonaldummy()	create matrix of seasonal indicator variables
seasonplot()	Plot seasonal effects