Fitting Time Series

Below is the output of help(package="ts"), which shows us the routines we have available in R.

ts  Time series functions

Description:

Package: ts
Version: 1.4.0
Priority: base
Title: Time series functions
Author: Collected by Brian Ripley. See SOURCES.
Maintainer: R Core Team <R-core@r-project.org>
Description: Time series functions.
License: GPL Version 2 or later.

Index:

Preliminary time-series package for R
--------------------------------------

This is a preliminary version of a time-series package.

Functions in base R:
---------------------

ts Create a (univariate or multivariate) ts object
[.ts Subsetting method for ts objects.
as.ts, is.ts Coercion and membership functions
plot, lines, print methods

cbind.ts cbind method for time series (aligns time bases)
na.omit.tsna.omit method for time series: omits at ends only
Ops.ts arithmetic (such as + - * /) for time series

aggregate Computes summaries (e.g. sum) over disjoint time intervals
diff Lagged differences of a time series
derd Time of last observation
frequency Number of observations per unit of time
deltat Return time interval between observations
time Create time series giving the times of observations
cycle Create time series giving the positions in a cycle of a time series
start Time of first observation
tsp, tsp<- Get and set time-series attributes
window Subset to a time window

Functions in package ts:
-----------------------
acf Autocovariance and autocorrelation function
ar Wrapper for autoregression estimation functions
ar.burg Estimate autoregression model by Burg's method
ar.ols Estimate autoregression model by ordinary least squares
ar.mle Estimate autoregression model by maximum likelihood
ar.yw Estimate autoregression model by solving Yule-Walker equations
arima0 ARIMA modelling -- provisional version
Box.test Box-Pierce and Ljung-Box tests of independence
ccf Cross-covariance and cross-correlations for two series
cpgram Plot cumulative periodogram of univariate time series
diffinv Discrete integration, the inverse of diff()
embed Embedding a time series
filter Linear filtering on a time series
kernapply Apply kernel smoothers
kernel Smoothing kernel objects (and (modified) Daniell, Fejer and Dirichlet kernels)
lag Compute lagged version of time series
na.contiguous Find longest contiguous stretch of non-NAs
pacf Partial autocorrelation function
plot.acf Plot autocorrelation function
plot.spec Plot spectral density estimate, coherency and phase.
PP.test Phillips-Perron test for unit roots
predict methods for ar and arima0
spec.ar        Estimate spectral density by autoregression
spec.pgram     Estimate spectral density from periodogram
spec.taper     Taper by cosine bell
spectrum       Wrapper for spectral density estimation functions
stl            Seasonal decomposition using loess
toeplitz       Generate Toeplitz matrix
ts.intersect   Bind time series as multivariate ts over the common
time base      
ts.plot        Plot several time series with different time bases
ts.union       Bind time series as multivariate ts over their total
time base      

In some cases the visual output will closer to that of S(-PLUS) if
options(ts.S.compat=TRUE) has been set.

Datasets in base R:
---------------------

airmiles    Passenger-Miles on US Airlines 1937-1960
co2         Moana Loa Atmospheric CO2 Concentrations
nhtemp      Yearly Average Temperatures in New Haven CT
presidents  Quarterly Approval Ratings for US Presidents
sunspots    Monthly Mean Relative Sunspot Numbers 1749-1983
uspop       Populations Recorded by the US Census

Datasets in package ts:
-----------------------

beavers      time series of body temperatures of two beavers
BJsales      sales data with leading indicator from Box &
Jenkins      
EuStockMarkets daily closing prices of major European stock
indices, 1991-8
LakeHuron    level of Lake Huron 1875-1972
lh           dataset on luteinizing hormone from Diggle (1990)
lynx         Annual Canadian Lynx trappings 1821-1934
nottem       monthly time-series of temperatures in Nottingham,
1920-1939
sunspot yearly sunspot data, 1700-1988
   monthly sunspot data, 1749-1997
treering yearly tree ring data, -6000-1979
UKDriverDeaths time series on UK road deaths of drivers from
   Harvey (1989)
UKLungDeaths   time-series on UK lung deaths 1974-9 from Diggle
   (1990)
USAccDeaths    US accidental deaths 1973-8

R gives three different methods for fitting AR models: ordinary least squares, “Burg’s Method”,
and Maximum likelihood. Usually, you will get similar results for all three.

We’ll continue to investigate the LakeHuron data. First, note that the acf (autocorrelation
function) suggests that the series is stationary, and possibly an AR model (since the correlations
gradually die down):

Further, we can fit the partial correlation function, which would tell us which order the AR model
should be if, in fact, an AR model is correct:
Note that the x axis begins at $k = 1$ (not $k = 0$ as in the acf plot), and that the last non-zero value is at $k=2$, which suggests fitting an AR(2) model.

In fact, this doesn’t matter much. R uses something called the AIC criterion to determine which order is best. Basically, it fits the AR(p) model for $p = 1, \ldots, 10 \log(N)$ (base 10) and calculates AIC for each. The “best” model is the one that minimizes the AIC. Strictly speaking the AIC applies only to the MLE fit, but empirically works in more general situations. These are details that won’t affect your use of the packages.

Here are the outputs from calls to `ar(LakeHuron)`, `ar.burg(LakeHuron)`, and `ar.mle(LakeHuron)`. Note that they are remarkably similar:

OLS:

```
> out$order
[1] 2
> out$ar
, , 1
```

```
[,1]
```
To confirm that this is a good model, we can plot the residuals. They should be white noise, which means their acf should be 0 for k <>0.

```r
> plot(out$resid)
```
acf(out$resid, na.action=na.omit)
This is, indeed, white noise. Note that I had to add the term `na.action=na.omit` to the call to `acf`. This is because the residuals had missing values, and in fact the first p terms will also be missing since predictions can’t be fit for them. The “plot” function deals with missing values automatically, but the default for the `acf` function is to fail if it sees a missing value.