S-PLUS/R tutorial for G6503[2]

Erwann Rogard [1]

[1] Department of Statistics, Columbia University

October 19, 2006

[2] taught by Professor Heyde

Outline

- Getting started
- Box Jenkins/ARIMA example
- Example revisited with GARCH
- Resources

Outline

Getting started

- Install
  - Download R from `http://www.r-project.org/`
    - Start R and install **`tseries`**,**`moments`** and **`fSeries`** from Menu>Packages>Install Packages from CRAN
  - S-PLUS and **`finmetrics`**: I can (still) lend you the CD-ROMS during the TA sessions

- Start the software and
  - R: load **`tseries`** and **`moments`** from Menu>Packages>Load Package
  - S+: load **`finmetrics`** from File>load module

- Open the commands window (Menu>Window) and create a new script file
  - R: File>New Script
  - S+: File>New>Script File
- You can execute commands in either window but you can *save* the script. Type `print("hello")` in the script and type
  - R: Ctrl-R
  - S+: key F10

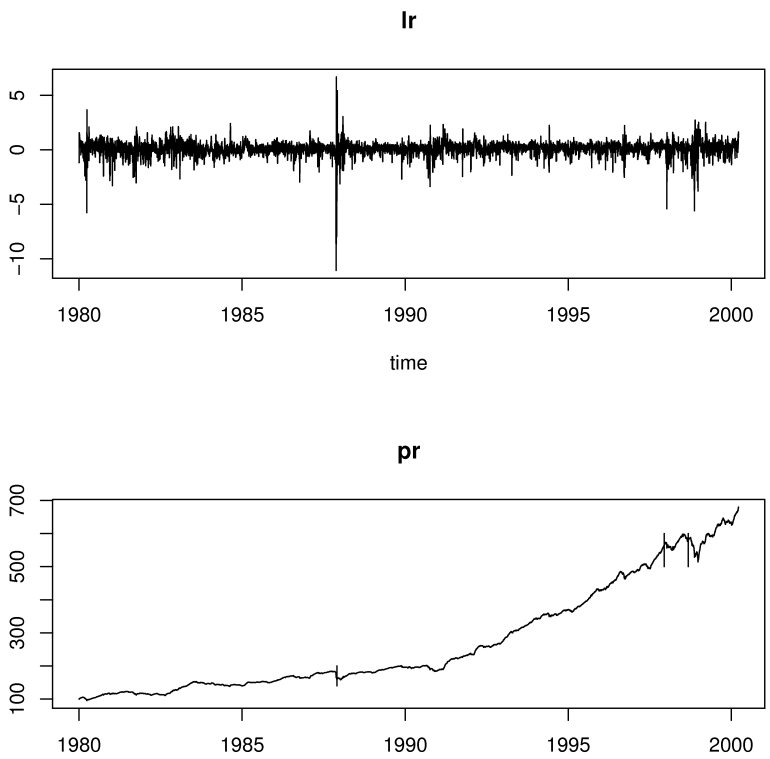
Outline

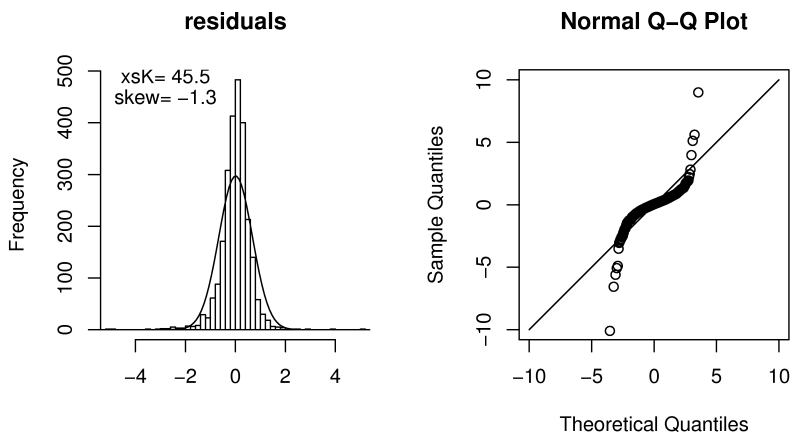
Example

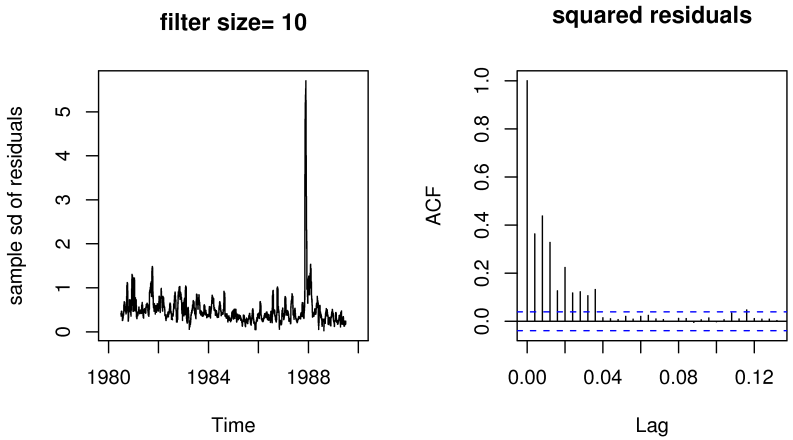
`http://faculty.chicagogsb.edu/ruey.tsay/teaching/fts/d-hwp3dx8099.dat`

- Column 3 contains daily log returns of the CRSP equal-weighted index from January 1980 to December 1999.
- Let's carry out the Box-Jenkins method in R/SPLUS to fit an ARIMA model to the period ending on the *last trading day of 1989*
  - identification
  - estimation
  - model checking
  - forecasting

Selected output




residuals
xsK= 45.5
skew= −1.3
Frequency
0
100
200
300
400
500
−4
−2
0
2
4
Normal Q−Q Plot
Sample Quantiles
−10
−5
0
5
10
Theoretical Quantiles


filter size= 10
sample sd of residuals
0
1
2
3
4
5
1980
1984
1988
Time
squared residuals
ACF
0.0
0.2
0.4
0.6
0.8
1.0
0.00
0.04
0.08
0.12
Lag

Review of Commands

- Data structures
  - Types
    `list`[[*name*]] `or list$`*name*
    `matrix`[*i*, *j*]
    `c(x,y,z)`[*i*]
    `array`
    `ts`
  - Ways to generate them
    `rep(x,n)`
    `seq(s,s+n) or s:(s+n)`
    `rnorm(100,mean=0,sd=1)`
    `paste("x=",x)`
    `ts(data)`
    `arima.sim(model=list(order=c(2,0,1),ar=c(a1,a2),ma=c(b1)))`
  - Vectorized operations faster than loops
    `for(in 1:n){x[i]<-exp(x[i])}`
    `x<-exp(x)`
    `filter`

- Times series

  `window` `arima`
  `filter` `predict`
  `diff` `tsdiag`
  `acf`

- Graphs

  `ts.plot` `lines`
  `acf` `par(mfrow=...)`
  `text` `hist`

- I/O

  `scan`
  `write`
  right click on a graph and Save as Postscript

Some differences (refer to the script)

R: `adf.test(ex1$pr)`
S+: `unitroot(ex1$pr,method="adf")`

R: `arima(train,order=c(i,0,j))`
S+: `arima.mle(train,model=list(order=c("ar"=3,"d"=0,"ma"=2)) )`

R: `ex1$fit$coef`
S+: `c(ex1$fit$model$ar,hw$fit$model$ma)`

R: `tsdiag(ex1$fit)`
S+: `arima.diag(ex1$fit)`

R: `ex1$fit$resid`
S+: `residuals(ex1$fit)`

R: `predict(ex1$fit,n.ahead=10)`
S+: `arima.forecast(train,model=ex1$fit$model,n=10)`

Outline

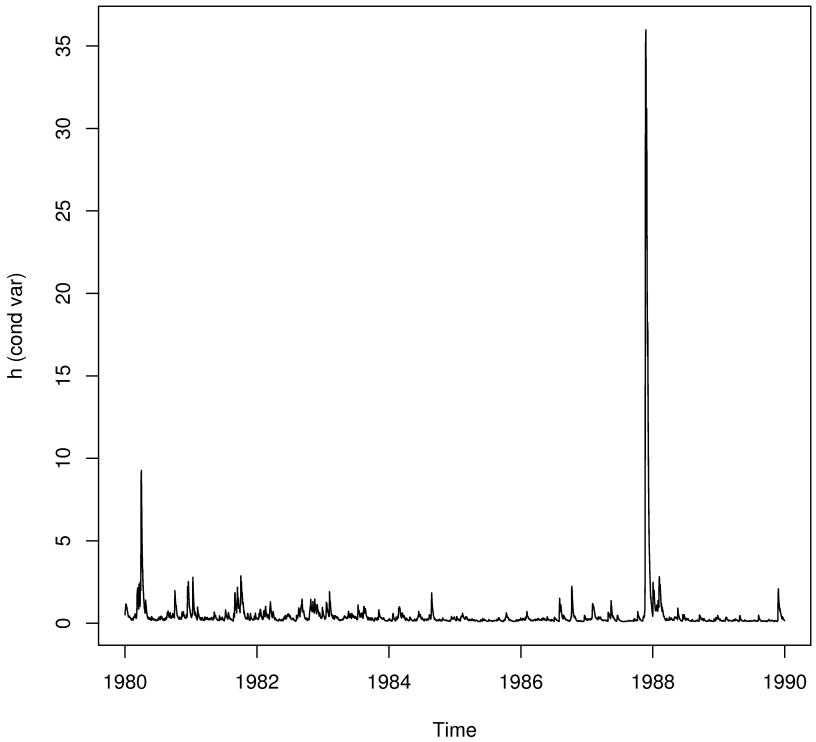
- Model checking revealed non–stationarity of the variance and non–independence of the squared residuals[18]. Negative skew and high kurtosis also suggest a GARCH model
- model

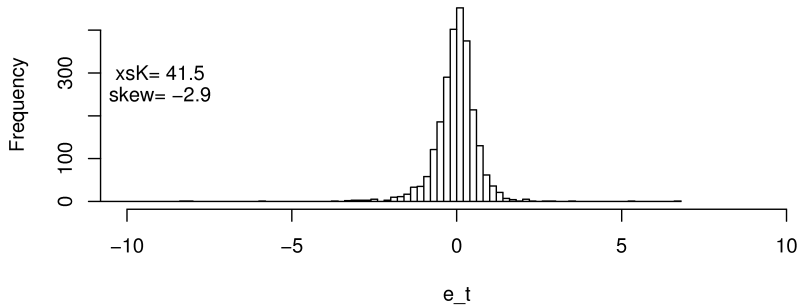
$$\epsilon_t|\mathcal{F}_t = \sigma_t z_t \tag{1}$$

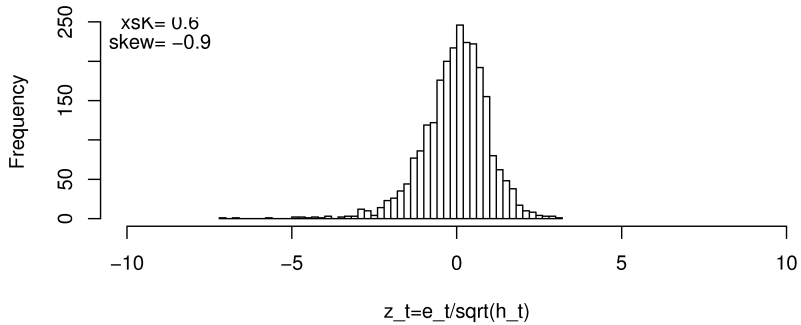
$$\sigma_t^2 = \omega + \alpha\epsilon_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{2}$$

- ```
  fit<-garchFit(train,formula.mean=~ arma(0,0),
  formula.var= garch(1,1))
  e<-fit@residuals
  h<-fit@fit$series$h
  ```

[18] for a precise interpretation of the correlation structure under GARCH, see Bollersev, 1986, Section 4
  ```


h (cond var)
0
5
10
15
20
25
30
35
1980
1982
1984
1986
1988
1990
Time


xsK= 41.5
skew= −2.9
Frequency
0
100
300
−10
−5
0
5
10
e_t


xsK= 0.6
skew= −0.9
Frequency
0
50
150
250
−10
−5
0
5
10
z_t=e_t/sqrt(h_t)

Outline

Selected resources

D. Wuertz, Y. Chalabi and Ladislav Luksan.

Parameter Estimation of ARMA Models with GARCH/APARCH Errors: an R an SPlus Software Implementation, Journal of Statistical Software

T. Bollersev.

GARCH, Journal of Econometrics, 1986