

Tips on Using *R*

February, 2011

Jin-Lung Lin

1 What is R?

R is an open-source (GPL) object oriented statistical package started by Robert Gentleman and Ross Ihaka of the Statistics Department of the University of Auckland in 1995 and currently maintained by the R core-development team, a hard-working, international team of volunteer developers. Syntactically and functionally R is very similar (if not identical) to S/Splus, the popular statistics packages.

2 Why using R?

According to Julian J. Faraway and many others, the advantages of R are:

1. Versatility

R is also a programming language and user is not limited by the procedures that are pre-programmed by a package. It is relatively easy to program new methods in R .

2. Interactivity

Data analysis is inherently interactive and R allows us to make changes interactively.

3. Portability

R is open-source software and may be freely redistributed. Linux, Macintosh, Windows and other UNIX versions are maintained and can be obtained from the R-project at www.r-project.org.

4. Popularity

R is most popular with researchers in Statistics and finance. There are already a lot of library for all areas in statistics and the number keeps increasing every day.

5. Interface

Foreign language interface is available.

6. Low (or no) cost

R is free (Who says there is no free lunch ?).

Disadvantages of R

1. It is not so easy to learn to use R as it is command-driven rather than menu-driven.
2. Inefficient in handling large dataset as compared with SAS.

3 Obtaining R

1. browse <http://www.r-project.org/>
2. click CRAN
3. select
<http://cran.csie.ntu.edu.tw/>
4. click Windows (95 and later)
5. click base
6. download R-2.12.2-win32.exe (As of Feb. 26, 2011)
7. also download appropriate Rblas.dll from /bin/windows/contrib/ATLAS. Be sure that you choose the right CPU version for your computer.

4 Installing R

1. Execute R-2.12.2-win32.exe. I accept all the defaults during installation process.
2. Replace the original C:/Program Files/R/R-2.12.2/bin/Rblas.dll(<200K) with the Rblas.dll (about 2Meg) optimized for your CPU. This will typically speed up the computation twice faster.
3. Edit C:/Program Files/R/R-2.12.2/etc/Rconsole by changing 'language = ' to 'language = en' provided you like the English environment.
4. Activate R, click on the button 'Package' and select 'Install packages'. Select 'http://cran.csdb.cn/' as reposit site and install all packages. It takes about 20 minutes.

5 Installing Tinn-R

Tinn-R is a small editor written specifically for running R. Obtaining Tinn-R from <http://sourceforge.net/projects/tinn-r>. Just run the setup file. The newest version is <http://sourceforge.net/projects/tinn-r> 2.0.0.7 released on Sep 5, 2008.

6 Manuals for R

Several excellent manuals are available at www.r-project.org

1. An Introduction to R:
under 'Manuals' at Documentation
2. Simple R
under 'contributed documentation'
3. Econometrics in R
under contributed documentation
4. Some-useful-R-pointers
available at <C:/Program Files/R/R-2.5.1/library/bayesm/doc>
5. A Brief Guide to R for Beginners in Econometrics by Mahmood Arai
available at http://people.su.se/~ma/R_intro/
6. R by examples
<http://www.mayin.org/ajayshah/KB/R/index.html>

7 Objects and files

R regards things as objects. A dataset, vector, matrix, results of a regression, a plot etc. are all objects. One or several objects can be saved in a file. A file containing R-data is not an object but a set of objects. Basically all commands you use are functions. A command: something(object), does something on an object. This means that you are going to write lots of parentheses. Check that they are there and check that they are in the right place.

Extensions for files

It is a good practice to use the extension `< R >` for your files including R-codes. A file `< lab1.R >` is then a text-file including R-codes. The extension `< rda >` is appropriate for work images (i.e files created by `< save() >`). The file `< lab1.rda >` is then a file including R-objects. The default name for the saved work image is `¡.RData¡`. Be careful not to name a file as `¡.RData¡` when you use `¡RData¡` as extension, since you will then overwrite the `¡.Rdata¡` file.

Saving and loading objects and images of working spaces

```
load("DataWageMacro.rda")
save(lnu, file="mydata.rda")
save.image()
save.image("myimage.rda")
```

8 Packages for R

1. R functions for doing forecasting written by Rob Hyndman
http://www-personal.buseco.monash.edu.au/~hyndman/index.php?option=com_content&task=view&id=44&Itemid=71
2. For finance
Rmetrics at <http://www.itp.phys.ethz.ch/econophysics/R/>
3. Task view on finance
<http://cran.r-project.org/src/contrib/Views/Finance.html>
4. Task view on Econometrics
<http://cran.r-project.org/src/contrib/Views/Econometrics.html>
5. task view on Bayesian
<http://cran.r-project.org/src/contrib/Views/Bayesian.html>
6. task view on graphics
<http://cran.r-project.org/src/contrib/Views/Graphics.html>
7. task view on Multivariate
<http://cran.r-project.org/src/contrib/Views/Multivariate.html>

9 Obtaining help within R

1. *help* or *?*

A useful trip is to type `'?command'`, say `?rnorm`, marked the example codes, and then paste them into R.

2. *help.search*, *apropos*, or *find*

Useful when not sure which command to look at. For example,
`apropos("cor")`

```
[1] ".__C__recordedplot" "cancor"          "cor"
[4] "cor.test"           "cov2cor"         "Harman23.cor"
[7] "Harman74.cor"       "recordGraphics"  "recordPlot"
```

3. Searching for help over the net

For example, `RSiteSearch("Kalman Filter")`

4. Read FAQs at www.r-project.org

5. Use the R site search at www.r-project.org for answers.