

A Short Note on R Programming

1 Download and install R

Download from <http://cran.cnr.berkeley.edu> and then install according to instruction.

2 The R environment

1. To start an R session, double click R on PC or type command *R* in UNIX (this later approach is not recommended as you will have to learn UNIX)
2. To stop an R session, in R prompt type
`> quit()`
3. You can input R statement by statement in R prompt or store all commands/statements in a file (say *myRsource.R*) first and then in R prompt type
`> source("myRsource.R")`
4. Commands are separated either by a semi-colon or by a newline.
5. Use `#` for comments, everything from `#` to the end of the line is a comment.
6. For help on a particular command, in R prompt type
`> help(commandname)`
or google with relevant keywords

3 Misc R programming tasks

1. Sequence of data.
`c()` – to create a list of data
`rep()` – to repeat a value for many times
`seq()` – to generate a sequence of values
2. Arrays and matrices.
`array()`, `matrix()` – create array or matrix
`det()` – compute the determinant of a matrix
`svd()` – singular value decomposition of a matrix
`qr()` – QR decomposition of a matrix
`t()` – matrix transpose
`dim()` – return dimension of a matrix
`crossprod()` – cross product of matrices
`outer()` – outer product of matrices
`A * B` – multiplication of two matrices
`A%*%B` – multiply matrices elementwise

3. Conditional statement
if expression1 else expression2
4. Loop (several ways to do this)
while(condition) {expression}
for (name in expression1) expression2
tapply()
5. Functions for input/output with file.
read.table() – to read input from a file usually a text file
scan() – to read input from a file
load() – To read input from a file, applies only to files created with *save()*
save() – to save a variable to a file
6. R graphics.
plot(), points(), text(), abline(), legend(), title(), axis(), par();
qqnorm() – to generate quantile-quantile plot w.r.t normal distribution
qqline() – to add a straight line to a qqplot
qqplot() – to generate quantile-quantile plot for two distributions
hist() – to generate histogram
contour() – to generate contours.
To save plots (in postscript format) do the following
> postscript("filename.ps")
> plot()
> dev.off()
To save in .pdf, .png, or .jpg format, you will replace *postscript()* with *pdf(), png(), or jpeg()*.
7. To use packages (for those already installed), type command
> library(packagename)
8. Handling missing values and etc.
is.na() - to determine if a data item is NA
is.nan() - to determine if a data item is NaN
is.numeric() - to determine if a data item is numeric
is.character() - to determine if a data item is a character
9. Various statistics functions.
max() – max of a list of data
min() – min of a list of data
sum() – addition of a list of data
var() – variance to a list of data
sd() – standard deviation of a list of data
mean() – mean of a list of data
sample() – to sample with or without replacement from a list of data
rnorm() – to generate a list of data from normal distribution
rgamma() – to generate data from the Gamma distribution
10. Write your own functions.
myFunctionName <- function(arg1, arg2, ...) expression

4 For more information

1. *An Introduction to R* from R-Project web site <http://www.r-project.org>
2. *R Reference Manual* from R-Project web site
3. W. N. Venables and B. D. Ripley. *Modern Applied Statistics with S-Plus*. Springer-Verlag.

4. www.google.com