

# A Quick Introduction to R

Zhanwu Liu  
Department of Statistics  
Carnegie Mellon University

January 29, 2010

## 1 Introduction to R

### 1.1 R and package installation on Ubuntu

A detailed instruction can be found here:

```
http://cran.r-project.org/bin/linux/ubuntu/README,  
and I used  
deb http://lib.stat.cmu.edu/R/CRAN/bin/linux/ubuntu karmic/  
in the file  
/etc/apt/sources.list  
Install the packages r-base and r-recommended.
```

### 1.2 Use of R and package

#### 1.2.1 Start

To use R interactively, type R in any terminal window. Then you can either type your R commands, or use `source` command to read in and execute an R file (like `source("filename.R")`).

Alternatively, you can run R in batch mode. In Linux console, type  
R CMD BATCH filename.R &

#### 1.2.2 Library/packages

The major package we are using is `mgcv`

There are two ways to use R packages, *library* is used more when you type the command, while *require* is used more in functions.

- *require(packagename)*
- *library(packagename)*

### 1.2.3 Getting help

Following is the resource you can use for getting R help

- Help for a function, for example *rnorm*, type `"?rnorm"` in the console
- Help for a package, for example *mgcv* package, type `"library(help=mgcv)"` or `"help('mgcv-package')"`

## 1.3 Matrix, vector, operators and build-in functions

A brief list of operators/operations

- Using index to get elements of matrix. Examples of indices  $A[1,5]$ ,  $A[2, c(1 : 8)]$ ,  $A[2, ]$ ;  $A[-1, ]$  will return  $A$  with the first row removed.
- `B=matrix(0, 5, 7)` will generate a matrix  $B$  that has 5 rows, 7 cols, all elements 0.
- $t(B)$  is the transpose
- $diag(B)$  will take the diagonal terms of  $B$  (if is square); note that  $diag(5)$  will generate  $I_5$  identity matrix,  $diag(1 : 10)$  will generate a  $10 \times 10$  diagonal matrix, with diagonal matrices 1 to 10.
- Matrix product  $A \times B$  is `A%*%B` in R.
- `unique(x)` returns the unique elements of vector  $x$ .
- Elements in a list: either by numbers `list1[[1]]` or by names `list1$name1`.
- `for ( i in 1:10) { commands }` is the for loop
- Note that dot is allowed in object name, for example, `sd.factor` is the name of one parameter.
- `apply` is used to perform row or column computation for matrix. For example `apply(A, 1, sum)` returns the sum of each row of  $A$ .

## 1.4 Input/Output

Ways to deal with different files

- `save`, `load` deal with R specific binary files (usually with `.RData` extension).
- `read.table`, `write.table` to read/write formatted ascii files
- `source` to read and run R scripts

## 2 Some other commands

### 2.1 Random numbers

`rnorm(n, mean, sd)` will generate  $n$  Gaussian random variables, `mean` is vector `mean`, `sd` is vector `sd`

### 2.2 Visualization

See the file `run_test.R` for examples.

- Quick plot using *plot* and *points*, *abline*
- Heatmap using *image*
- Output using *postscript* command

## 3 The *mgcv* package

The package *mgcv* is the package I used for *gam*. The best reference for this package is a book written by Simon N Wood(<http://www.maths.bath.ac.uk/~sw283/>), the name of the book is *Generalized Additive Models: An Introduction with R*

See the code for details.

## 4 List of files

Here is the list of files

- `common_functions.R`, provides functions `vecnorm()` and `Mat.normalization()`
- `simulate_GaussianSpeed_8direction.R`, provides function `simu.8d()`
- `simulate_realkinematics.R`, provides function `simu.realpos()`
- `fit_noCV.R`, provides function `fit_noCV()`
- `fit_CV.R`, provides function `fit_CV()`
- `run_test.R`, run tests
- `pos.mat.RData`, Rdata file contains the hand positions from 8 center-outer trials, binary format to be opened by R directly. It contains a matrix `pos.mat` with four columns, the first three are the  $x, y$  and  $z$  coordinates, while the last one is the identifier for trial.