Introduction to R

P. Adames

11/14/2020

Contents

1 Introduction to R

The introduction to R document in the official repository of packages for the R language has 105 pages and 14 sections (https://cran.r-project.org/doc/manuals/r-release/R-intro.pdf). It covers data structures, reading and writing (I/O) to files, probability distributions, statistical models, conditional statements and other forms of controlling the execution of a program, user functions, graphics, extension via packages, and commands to interact with the host operating system.

Undoubably too much for a 45 minute session. Therefore the focus will be on two aspects:

- language specific syntax
- basic I/O and interoperability with the operating system

These two aspects can give the student a quick sense of usability for the R langauage. This approrach leaves an essential strengtgh of R for later, its nature as a statistical computing and graphics language.

2 Language Syntax

Programing languages model information using data structures. At its core, R represents everything either as an object or a function. There are two immediate consequences from this: there no absolute need to use for loops —although they exist— because functions are designed to work very effectively on vectorized data structures, and scalars only exist as unidimensional data structures of length 1 (Ihaka and Gentleman 1996).

We will emphasize these two points to set the stage for elegant and fast R code. R was a functional language long before this became a buzz word in Computer Science.

2.1 Data structures

R uses the data structures illustrated in the table below from (Wickham 2015, 13). Vectors, matrices, adn arrays can store values of the same data type, whether they are are numeric, integer, logical, or character. In contrast, lists and data frames can store mixed data types. For 3d and higher dimensionality the native data structure is an object of class *array*.

	Homogeneous	Heterogeneous
1d	Atomic vector	List
2d	Matrix	Data frame
nd	Array	

Because every data in R is an object, all of these structures have attributes like name, dimension, and a `class`. They can also receive and carry user-defined attributes.

2.2 Data types

- *numeric:* double precision floating point real numbers.
- *integer:* model integer numbers.
- *logical:* can be TRUE, T, 1 or FALSE, F, 0.
- *character:* models aa text character in an encoding.

2.3 Built in functions

The magic of the generic dispatch mechanism of S3 R objects.

2.3.1 Iteration over data structures

2.3.2 I/O

2.3.3 Interoperatibility with the host operating system

2.4 User-functions

2.5 Execution control

References

Ihaka, Ross, and Robert Gentleman. 1996. "R: A Language for Data Analysis and Graphics." *Journal of Computational and Graphical Statistics* 5 (3). [American Statistical Association, Taylor & Francis, Ltd., Institute of Mathematical Statistics, Interface Foundation of America]: 299–314. http://www.jstor.org/stable/1390807.

Wickham, Hadley. 2015. *Advanced R*. Chapman & Hall/Crc the R Series (Crc Press).