Loops and Control Structure in R

Baburao Kamble

September,2014

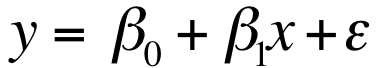
# Introduction to Regression

Regression analysis is fundamental and forms a major part of statistical analysis. Regression is a statistical technique to determine the linear relationship between two or more variables.

Regression analysis is used to predict the value of one variable (the dependent variable) on the basis of other variables (the independent variables).

Dependent variable: denoted Y

Independent variables: denoted X1, X2, …, Xn



which is referred to as simple linear regression. We would be interested in estimating β0 and β1 from the data we collect.

Mastering R programming languages require an analyst to master the essential challenge of learning loops.

In R Programming, control flow (or alternatively, flow of control) refers to the order in which the individual statements, instructions or function calls of an imperative or a declarative program are executed or evaluated. The R language has those traditional looping structures like the for loop and while loop, but that’s not all.

The kinds of control flow statements supported by R programming language can be categorized by its effect:

*Conditional execution:* Conditional execution refers to the ability to execute a statement or sequence of statements only if some condition holds.

*Repetitive execution (Loops)*: Looping constructs repetitively execute a statement or series of statements until a condition isn’t true. These include the *for* and *while* structures.

# Text Data Input/Output

Here we explore how to define a data set in an R session. We will explore common text data format (text, csv, excel)

## \*.TXT Data

A tab-separated values file is a simple text format for storing data in a tabular structure (e.g. database or spreadsheet data). Delimited text files (.txt), in which the TAB character (ASCII character code 009) typically separates each field of text. Tab-delimited files are text files organized around data that has rows and columns.

R uses read.table to reads a tab separated text file in table format and creates a data frame from it, with cases corresponding to lines and variables to fields in the file.

>read.table(file,header=FALSE, sep="\t", ,)

*file*: the name of the file which the data are to be read from.

*header*: a logical value indicating whether the file contains the names of the variables as its first line

the field separator character.

*sep*: the field separator character.