

Introducción a R

Modelos no paramétricos y de regresión

Enrique Reyes

05 de febrero de 2018

Matrices

Al igual que los vectores, este tipo de objetos sólo puede ser de un sólo tipo de variable, una matriz no puede almacenar en una columna datos de tipo numérico y en otra de tipo carácter, esta es una limitante que debemos tener presentes.

```
#Matriz vertical (Mn x 1)
```

```
M1=matrix(1:6); M1
```

```
##      [,1]
## [1,]    1
## [2,]    2
## [3,]    3
## [4,]    4
## [5,]    5
## [6,]    6
```

```
#Matriz General (Mn x m)
```

```
M2=matrix(1:6,nrow=2); M2
```

```
##      [,1] [,2] [,3]
## [1,]    1    3    5
## [2,]    2    4    6
```

```
M3=matrix(1:6,nrow=2,byrow=TRUE); M3
```

```
##      [,1] [,2] [,3]
## [1,]    1    2    3
## [2,]    4    5    6
```

```
M=matrix(1:6,nrow=3); M
```

```
##      [,1] [,2]
## [1,]    1    4
## [2,]    2    5
## [3,]    3    6
```

```
#Numero de datos en una matriz
```

```
length(M)
```

```
## [1] 6
```

```
#Tipo de matriz
```

```
mode(M); str(M)
```

```
## [1] "numeric"
```

```
## int [1:3, 1:2] 1 2 3 4 5 6
```

```
#Dimension
```

```
dim(M)
```

```
## [1] 3 2
#Buscar el nombre o etiqueta de nuestras columnas y renglones
M4=M; M4

##      [,1] [,2]
## [1,]    1    4
## [2,]    2    5
## [3,]    3    6

rownames(M)<-c("Hugo","Paco","Luis"); M

##      [,1] [,2]
## Hugo    1    4
## Paco    2    5
## Luis    3    6

colnames(M)<-c("Edad","Estatura"); M

##      Edad Estatura
## Hugo    1         4
## Paco    2         5
## Luis    3         6

dimnames(M4)<-list(c("Hugo","Paco","Luis"),c("Edad","Estatura")); M4

##      Edad Estatura
## Hugo    1         4
## Paco    2         5
## Luis    3         6

y1=1:24;y1

## [1]  1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16 17 18 19 20 21 22 23
## [24] 24

x1=matrix(1:24,nrow=1,byrow=T); x1

##      [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13]
## [1,]    1    2    3    4    5    6    7    8    9   10   11   12   13
##      [,14] [,15] [,16] [,17] [,18] [,19] [,20] [,21] [,22] [,23] [,24]
## [1,]   14   15   16   17   18   19   20   21   22   23   24

y1[1]

## [1] 1
x1[1,1]

## [1] 1
#Para saber si tenemos una matriz
is.matrix(y1)

## [1] FALSE
is.matrix(x1)

## [1] TRUE
as.matrix(y1,byrow=T)

##      [,1]
```

```
## [1,] 1
## [2,] 2
## [3,] 3
## [4,] 4
## [5,] 5
## [6,] 6
## [7,] 7
## [8,] 8
## [9,] 9
## [10,] 10
## [11,] 11
## [12,] 12
## [13,] 13
## [14,] 14
## [15,] 15
## [16,] 16
## [17,] 17
## [18,] 18
## [19,] 19
## [20,] 20
## [21,] 21
## [22,] 22
## [23,] 23
## [24,] 24
```

```
#Extracción de datos
```

```
M2[1,1]
```

```
## [1] 1
```

```
M3[1,3]
```

```
## [1] 3
```

```
#error es necesario saber la dimensión de nuestra matriz
```

```
#M1[1,2]
```

```
#Extracción por columna
```

```
M2[,2]
```

```
## [1] 3 4
```

```
#Extraccion de renglones
```

```
M2[1,]
```

```
## [1] 1 3 5
```

```
#
```

```
M3[]
```

```
##      [,1] [,2] [,3]
```

```
## [1,] 1    2    3
```

```
## [2,] 4    5    6
```

```
#Concatenación de columnas
```

```
cbind(M3,c(9,8))
```

```
##      [,1] [,2] [,3] [,4]
```

```
## [1,] 1    2    3    9
```

```
## [2,] 4    5    6    8
```

```
#Concatenación por renglon
```

```
M5=rbind(M,c(11,10)); M5
```

```
##      Edad Estatura
## Hugo    1         4
## Paco     2         5
## Luis     3         6
##          11        10
```

```
rownames(M5)[4]="Pedro"; M5
```

```
##      Edad Estatura
## Hugo    1         4
## Paco     2         5
## Luis     3         6
## Pedro   11        10
```

```
#Operaciones
```

```
#Transponer
```

```
t(M1)
```

```
##      [,1] [,2] [,3] [,4] [,5] [,6]
## [1,]    1    2    3    4    5    6
```

```
#Producto matricial
```

```
MC<-M1%*%t(M1); MC
```

```
##      [,1] [,2] [,3] [,4] [,5] [,6]
## [1,]    1    2    3    4    5    6
## [2,]    2    4    6    8   10   12
## [3,]    3    6    9   12   15   18
## [4,]    4    8   12   16   20   24
## [5,]    5   10   15   20   25   30
## [6,]    6   12   18   24   30   36
```

```
#inversa
```

```
solve(matrix(1:4,nrow=2))
```

```
##      [,1] [,2]
## [1,]   -2  1.5
## [2,]    1 -0.5
```

```
#solución de un sistema de ecuaciones
```

```
solve(matrix(1:4,nrow=2),c(1,-1))
```

```
## [1] -3.5  1.5
```

```
#Matriz diagonal
```

```
diag(1:4)
```

```
##      [,1] [,2] [,3] [,4]
## [1,]    1    0    0    0
## [2,]    0    2    0    0
## [3,]    0    0    3    0
## [4,]    0    0    0    4
```

```
#determinante
```

```
det(diag(1:4))
```

```
## [1] 24
```

```

datos<-matrix(c(20,65,174,22,70,180,19,68,170),nrow = 3,byrow=T)
datos

##      [,1] [,2] [,3]
## [1,]   20   65  174
## [2,]   22   70  180
## [3,]   19   68  170

dimnames(datos)<-list(c("hugo","paco","luis"),c("edad","peso","estatura"))
datos

##      edad peso estatura
## hugo   20   65     174
## paco   22   70     180
## luis   19   68     170

#Regresa todos los elementos de el renglon (obs) dado
#por nombre de sujeto
datos["hugo",]

##      edad      peso estatura
##        20        65     174

#por numero de renglon
datos[1,]

##      edad      peso estatura
##        20        65     174

#
datos["paco",]

##      edad      peso estatura
##        22        70     180

datos["luis",]

##      edad      peso estatura
##        19        68     170

#Regresa todos los elementos de la columna (var) dada
#por nombre de columna
datos[, "edad"]

## hugo paco luis
##   20   22   19

#por numero de columna
datos[,1]

## hugo paco luis
##   20   22   19

#
datos[, "peso"]

## hugo paco luis
##   65   70   68

datos[, "estatura"]

```

```

## hugo paco luis
## 174 180 170

#Extraccion por coordenada
datos[1,1]

## [1] 20
datos["hugo","edad"]

## [1] 20
datos[2,2]

## [1] 70
datos["paco","peso"]

## [1] 70
datos[3,3]

## [1] 170
datos["luis","estatura"]

## [1] 170
datos["luis",2]

## [1] 68
datos[2,"estatura"]

## [1] 180
#Extracción de más de una variable u observación
datos[,c("edad","estatura")]

##      edad estatura
## hugo    20      174
## paco    22      180
## luis    19      170
datos[c("hugo","luis"),]

##      edad peso estatura
## hugo    20   65      174
## luis    19   68      170
#Extraccion de los nombres de reglones y columnas
dimnames(datos)

## [[1]]
## [1] "hugo" "paco" "luis"
##
## [[2]]
## [1] "edad"    "peso"    "estatura"
colnames(datos)

## [1] "edad"    "peso"    "estatura"

```

```
rownames(datos)
```

```
## [1] "hugo" "paco" "luis"
```

```
#Calculo de media por variable
```

```
apply(datos,2,mean)
```

```
##      edad      peso  estatura  
## 20.33333 67.66667 174.66667
```

```
apply(datos,1,mean)
```

```
##      hugo      paco      luis  
## 86.33333 90.66667 85.66667
```

Arreglos

```
# Matrices y arreglos (muy similares)
```

```
array(1:6, c(2, 3))
```

```
##      [,1] [,2] [,3]  
## [1,]    1    3    5  
## [2,]    2    4    6
```

```
#Pero también se ven como una generalización de matrices
```

```
#array(datos,C(# de renglones,# de columnas,# de matrices))
```

```
array(1:12,c(2,3,2))
```

```
## , , 1
```

```
##
```

```
##      [,1] [,2] [,3]  
## [1,]    1    3    5  
## [2,]    2    4    6  
##
```

```
## , , 2
```

```
##
```

```
##      [,1] [,2] [,3]  
## [1,]    7    9   11  
## [2,]    8   10   12
```

```
array(1:12,c(4,3,1))
```

```
## , , 1
```

```
##
```

```
##      [,1] [,2] [,3]  
## [1,]    1    5    9  
## [2,]    2    6   10  
## [3,]    3    7   11  
## [4,]    4    8   12
```

```
array(1:12,c(2,2,3))
```

```
## , , 1
```

```
##
```

```
##      [,1] [,2]  
## [1,]    1    3  
## [2,]    2    4  
##
```

```

## , , 2
##
##      [,1] [,2]
## [1,]    5    7
## [2,]    6    8
##
## , , 3
##
##      [,1] [,2]
## [1,]    9   11
## [2,]   10   12

#Matrices simultaneas
dat<-array(c(45,46,65,55,170,167,48,49,68,56,169,165),c(2,3,2)); dat

## , , 1
##
##      [,1] [,2] [,3]
## [1,]   45   65  170
## [2,]   46   55  167
##
## , , 2
##
##      [,1] [,2] [,3]
## [1,]   48   68  169
## [2,]   49   56  165

dimnames(dat)<-list(SEXO=c("hombres","mujeres"),VARIABLES=c("edad","peso","estatura"),
                    ESTADO=c("CDMX","EdoMex"))
dat

## , , ESTADO = CDMX
##
##      VARIABLES
## SEXO      edad peso estatura
## hombres   45   65   170
## mujeres   46   55   167
##
## , , ESTADO = EdoMex
##
##      VARIABLES
## SEXO      edad peso estatura
## hombres   48   68   169
## mujeres   49   56   165

#Extraccion de etiquetas
dimnames(dat)

## $SEXO
## [1] "hombres" "mujeres"
##
## $VARIABLES
## [1] "edad"      "peso"      "estatura"
##
## $ESTADO
## [1] "CDMX"      "EdoMex"

```



```
#Extraccion de matrices
```

```
dat[,,"EdoMex"]
```

```
##          VARIABLES
## SEXO      edad peso estatura
## hombres   48   68   169
## mujeres   49   56   165
```

```
dat[,,"CDMX"]
```

```
##          VARIABLES
## SEXO      edad peso estatura
## hombres   45   65   170
## mujeres   46   55   167
```

```
#Extracción de columnas y renglones
```

```
dat["hombres",,]
```

```
##          ESTADO
## VARIABLES  CDMX EdoMex
## edad       45    48
## peso       65    68
## estatura   170   169
```

```
dat[,,"peso",]
```

```
##          ESTADO
## SEXO      CDMX EdoMex
## hombres   65    68
## mujeres   55    56
```

```
#Extraccion de más de una variable u observación
```

```
dat[,c("edad","estatura"),]
```

```
## , , ESTADO = CDMX
##
##          VARIABLES
## SEXO      edad estatura
## hombres   45    170
## mujeres   46    167
##
## , , ESTADO = EdoMex
##
##          VARIABLES
## SEXO      edad estatura
## hombres   48    169
## mujeres   49    165
```

```
#Calculo de minimo por sujeto
```

```
apply(dat,1,min)
```

```
## hombres mujeres
##      45      46
```

```
#Calculo de media por variable
```

```
apply(dat,2,mean)
```

```
##      edad      peso estatura
##    47.00    61.00   167.75
```

```
#Calculo de la maximo por matriz  
apply(dat,3,max)
```

```
## CDMX EdoMex  
## 170 169
```

Listas

```
familia<-list(padre="Juan",madre="María",edad_padres=c(30,29),num_hijos=3,  
             nom_hijos=c("Axel","Damian","Tania"),edad_hijos=c(7,5,3),ciudad="Madrid")  
familia
```

```
## $padre  
## [1] "Juan"  
##  
## $madre  
## [1] "María"  
##  
## $edad_padres  
## [1] 30 29  
##  
## $num_hijos  
## [1] 3  
##  
## $nom_hijos  
## [1] "Axel" "Damian" "Tania"  
##  
## $edad_hijos  
## [1] 7 5 3  
##  
## $ciudad  
## [1] "Madrid"
```

```
#Regresa el nombre de las "variables"  
names(familia)
```

```
## [1] "padre"      "madre"      "edad_padres" "num_hijos"  "nom_hijos"  
## [6] "edad_hijos" "ciudad"
```

```
#Extraccion de datos  
familia$padre
```

```
## [1] "Juan"
```

```
familia$edad_hijos
```

```
## [1] 7 5 3
```

```
familia[[1]]
```

```
## [1] "Juan"
```

```
familia[[3]]
```

```
## [1] 30 29
```

```
#Error, pues en listas no acepta extraccion simultanea  
#familia[[c(1,3)]]
```

```
#Esta es la forma correcta de haces extracciones de más de un elemento  
familia[c(1,3)]
```

```
## $padre  
## [1] "Juan"  
##  
## $edad_padres  
## [1] 30 29
```

```
familia[c(4,5,6)]
```

```
## $num_hijos  
## [1] 3  
##  
## $nom_hijos  
## [1] "Axel" "Damian" "Tania"  
##  
## $edad_hijos  
## [1] 7 5 3
```

```
#Otra forma de hacer extracciones en listas es  
familia$nom_hijos
```

```
## [1] "Axel" "Damian" "Tania"
```

```
familia$padre
```

```
## [1] "Juan"
```

```
familia$edad_padres[1]
```

```
## [1] 30
```