R programming for beginners

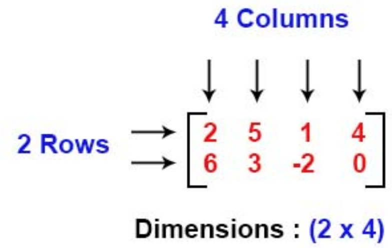
Ni Shuai

Computational Genome Biology
German Cancer Research Center (DKFZ)

November, 2016

Matrix

- A matrix is a rectangular array of numbers or symbols arranged in rows and columns.
- The number of rows and columns determine the size of the matrix.



- A matrix with m rows and n columns is called an m-by-n matrix, while m and n are called its dimensions.
- Each row and column of a matrix is a vector
- Vectors can be regarded as a 1-row or 1-column matrix

Build a marix

Matrices can be constructed with the matrix() function:

```
A = matrix( data=c(2, 4, 3, 1, 5, 7), nrow=3, ncol=2, byrow=FALSE)
```

where the arguments mean:

Data	An optional vector of data to fill in the matrix
nrow	The desired number of rows of the matrix
ncol	The desired number of columns of the matrix
byrow	In which way will the matrix be filled by 'Data' (by row or by column)

Build a marix

- Elements in a matrix must be of the same type*
- The number of elements in the matrix should be equal to the product of its dimensions.
- Either dimention has to be specified, R will calculate the other automatically.

Try me:

```
x=matrix(1:12, nrow=2); # Fill the matrix by column by default
x=matrix(1:12, nrow=4);
x=matrix(1:12, ncol=2); # Same as x=matrix(1:12, nrow=6)
x=matrix(1:12, ncol=4); # Same as x=matrix(1:12, nrow=3)
x=matrix(1:12, nrow=2, byrow=TRUE);
x=matrix(1:12, nrow=4, byrow=TRUE);
```

Matrix indexing

- The [ ] operater is also used to index a matrix, but since matrices are two dimentional, one needs to specigy both row number and column number of an element
- To access the element at row i and column j of a matrix X, simply use A[i,j]
- To access all elements in row i, simply omit the the column index: A[i,]

Examples:

```
X=matrix(1:12, nrow=4);X      # Create a matrix X

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

Matrix indexing

Extract a vector:

```
X[1,2]              # The element in the 1st row and 2nd column

## [1] 5

X[3, ]              # The 3rd row of matrix X

## [1]  3  7 11

X[ ,1]              # The 1st column of matrix X

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

Exercise:
change the value of X in row 2, column 3 to be 100.
change the values of X in the 3rd column to be its 3rd column plus 100.

Matrix indexing

Subset to a smaller matrix:

```
X[2,1:2]                  # The the first 2 columns of the 1st row

## [1] 2 6

X[c(2,4),c(1,3)]          # The 2nd and last row of 1st and last column

##      [,1] [,2]
## [1,]    2   10
## [2,]    4   12

X[-1,]                    # Exclude the first row

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

Functions for matrix

dim()	List the dimensions of a matrix.
rownames()	Retrieve or set the row names of a matrix
colnames()	Retrieve or set the column names of a matrix
t()	Calculate the transpose of a matrix
nrow(), ncol()	Number of rows or columns of a matrix
rbind(), cbind	combine two matrices by row or by column
as.vector()	coerce a matrix into a plane vector by column
rowMeans(), colMeans()	report mean of each row or column of a matrix
rowSums(), colSums()	report sum of each row or column of a matrix
%*%	Matrix multiplication

Row and column names

Assign row and column names to matrices

```
rownames(X)=paste0('foo',1:4)
colnames(X)=paste0('bar',1:3); X

##      bar1 bar2 bar3
## foo1    1    5    9
## foo2    2    6   10
## foo3    3    7   11
## foo4    4    8   12
```

Use names to subset matrices, make sure names are enclosed by single or double quotes

```
X[c('foo1','foo3'),]

##      bar1 bar2 bar3
## foo1    1    5    9
## foo3    3    7   11
```

PEMDAS in matrix

Basic mathematical operations also hold in matrix like in vectors, for example:

```
S=matrix(1:8, nrow=2)
S+S

##      [,1] [,2] [,3] [,4]
## [1,]    2    6   10   14
## [2,]    4    8   12   16

S*3

##      [,1] [,2] [,3] [,4]
## [1,]    3    9   15   21
## [2,]    6   12   18   24

S^3

##      [,1] [,2] [,3] [,4]
## [1,]    1   27  125  343
## [2,]    8   64  216  512
```

Combine matrices

Use "cbind( )" to combine two matrices with same number of rows

```
Y=matrix(rnorm(16), nrow=4)
Z=cbind(X,Y);Z

##      bar1 bar2 bar3
## foo1    1    5    9  2.5140297 -0.41007391  0.02117209  0.5658298
## foo2    2    6   10  1.3475018  0.39538087  1.54117168 -1.5791798
## foo3    3    7   11 -0.1203425  0.03155866 -2.23743212  2.3958135
## foo4    4    8   12 -0.9108418  0.47753508 -0.72007441 -1.9817583

colnames(Z)

## [1] "bar1" "bar2" "bar3" ""     ""     ""     ""
```

Exercise: The last 4 column names of Z are currently empty, try to fill it with a vector 'norm1, norm2, norm3 and norm4'

Matrix

Exercise:

- Creat the following matrix A

$$\begin{bmatrix} 1 & 8 & 4 \\ 3 & 9 & 3 \\ 0 & -5 & -1 \end{bmatrix}$$

1) Calculate the sum of the second column
2) Replace the third column of A by the sum of its second and third column
3) Double the first two rows of A

Matrix

Exercise:

- Creat the following matrix B with 11 rows

$$\begin{bmatrix} 3 & 4 & 5 & \dots & 8 \\ 3 & 4 & 5 & \dots & 8 \\ \multicolumn{5}{c}{\dotfill} \\ 3 & 4 & 5 & \dots & 8 \end{bmatrix}$$

1) Check how many columns does matrix B have using dim()
2) Double the last column of B and bind it to B as its first column
3) Give a name to each column of B

Matrix

Exercise:

- Use rnorm() to generate 100 random number from a normal distribution, put them into a 10 by 10 matrix C
- Check how many column have a mean value larger than 0
- Check which row in matrix C has the larget mean
- What is the larget number in C