

컴퓨터 응용통계

1-3 R 입문

최경미

vector, matrix, data.frame

- 벡터 Vector

$$\vec{x} = (x_1, x_2, \dots, x_n)$$

- 행렬 Matrix

모든 성분이 숫자 또는 모든 성분이 문자

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} \underline{R_1} \\ \underline{R_2} \\ \vdots \\ \underline{R_m} \end{pmatrix} = (\underline{C_1} \quad \underline{C_2} \quad \dots \quad \underline{C_n})$$

- 데이터 프레임 Data.frame

열 별로 숫자 또는 문자

R Console 창에서 계산기처럼 사용하기

대소문자 구별

한 번 만든 변수는 지우기 전에 없어지지 않음

algebra

2+3

2^3 # $2^3 = 8$

1/2 # $\frac{1}{2} = 0.5$

3-2*1+4

sqrt(2) # $\sqrt{2}$

abs(-1) # $|-1| = 1$

factorial(5) # $5! = 5 \cdot 4 \cdot 3 \cdot 2 \cdot 1 = 120$

choose(5,2) # $\binom{5}{2} = \frac{5!}{2!3!} = 10$

변수 만들기

`x <- 3`

$x = 3$

`x+4`

$3 + 4 = 7$

폴더 관리

`ls()`

R 내부 객체를 보여줌. x 포함.

`rm(x)`

그 중에서 x를 삭제함

`ls()`

R 내부 객체를 보여줌. x 없음

`rm(list=ls())`

R 내부 모든 객체를 삭제. clear folder

벡터 vectors

z <- 3:13 # 3부터 13까지 1간격으로 벡터 만들기

z # z 확인

z <- seq(3,13,0.1) # 3부터 13까지 0.1 간격으로 벡터 만들기

z <- 2.2 *z # 벡터 전체 성분에 2.2 곱하기

length(z) # 벡터의 크기

z[2] # 벡터의 두번째 성분

c(1,3,4) # 1,3,4 숫자를 붙여서 벡터 한 개 만들기

z[c(1,3,4)] # z 벡터 중, 1,3,4 번째 성분 가져오기

z[-1] # 1번 삭제

z[-c(1,3)] # 1번 3번 삭제

fruits <- c("banana", "apple", "orange", "mango")

fruits[2]

fruits <- c(fruits, "lemon")

all <- c(z, fruits)

matrices (모두 숫자 또는 모두 문자)

$x \leftarrow c(1,2,3)$ #화면에서는 행벡터, 실제로는 열벡터 $x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$

$y \leftarrow c(4,5,6)$ #화면에서는 행벡터, 실제로는 열벡터 $y = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}$

$xy \leftarrow c(x,y)$ # 벡터붙이기 $xy = (1,2,3,4,5,6)$

$A \leftarrow cbind(x,y)$ # 열붙이기 $A = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}$

$B \leftarrow rbind(x,y)$ # 행붙이기 $B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$

$c(1,3,5,0)$ # 벡터 만들기

$A \leftarrow matrix(c(1,3,5,0),2,2, byrow=T)$ # 행렬 만들기
R에서 행렬은 모두 숫자(numeric)
또는 모두 문자 (character)

$A[1,2]$ # $A(1,2) = 4$

$A[,2]$ # 2열

$B[1,]$ # 1행

Data frame

```
#2-3 data frame
name <-c("jaehee","younghee","chulsu","sunhee")
age <- c(10,12,13,15)
family <- data.frame(name,age)
family
names(family)
family$name
family$age
family[,1]
family[,2]
family[1,2]

gender <- c("f","f","m","f")
family <- data.frame(family, gender)
family$gender <- NULL
```

```
# use of built-in data
```

```
library()
```

```
library(datasets)
```

```
data()
```

```
# 사용가능한 패키지 보기
```

```
# "datasets library" 로딩하기
```

```
# 사용가능한 무료 데이터 보기
```

```
# data.frame ( 숫자 문자 섞임. 한 열이 모두 문자 또는 모두 숫자)
```

```
mtcars
```

```
help(mtcars)
```

```
head(mtcars)
```

```
names(mtcars)
```

```
mtcars$am
```

```
mtcars$hp
```

```
table(mtcars$am)
```

```
boxplot(mtcars$hp)
```

```
summary(mtcars$hp)
```

```
boxplot(hp~am, data=mtcars)
```

```
# help 보기
```

```
# 첫 몇 줄 보기
```

```
# 열 이름 보기
```

```
# am 열(변수) 가져오기
```

```
# hp 열(변수) 가져오기
```

```
# 빈도표 만들기
```

```
# 상자도표 그리기
```

```
# 기술통계 구하기
```

```
# am=0/1 집단 별로 hp(마력)에 대한 상자도표 그리기
```


변수(열) 선택하기 또는 data.frame 쪼개기

head(mtcars)

mtcars[1:5,]

1,2,3,4,5 행 가져오기

mtcars[,2]

2열 가져오기

mtcars\$am

am 열 (변수) 가져오기

mtcars\$hp

hp 열 (변수) 가져오기

subset(mtcars, vs==1)

행 (표본) 선택

subset(mtcars, vs==1&am==0)

mtcars[,c("mpg","hp")]

열 (변수) 선택

subset(mtcars, vs==1&am==0, c("mpg","hp"))

새 변수 만들기

mtcars\$hp2 <- mtcars\$hp + 3

열 (변수)에 연산을 실시 후 저장하기

head(mtcars)

변수 없애기

mtcars\$hp2 <- NULL

열 (변수) 없애기

head(mtcars)

폴더 지정, 데이터 파일 읽기, 데이터 파일로 내보내기

```
write.csv(mtcars, "c:/temp/mtcars.csv")    # mtcars를 지정된 위치에 Excel로 내보내기  
mtcars2<-read.csv("c:/temp/mtcars.csv")  # 다시 읽어들이기
```

① 파일 => 작업디렉토리 변경 => 원하는 폴더로 변경

```
write.csv(mtcars, "mtcars.csv")           # family를 지정된 위치에 Excel로 내보내기  
mtcars2<-read.csv("mtcars.csv")          # 다시 읽어들이기
```

폴더 지정, 데이터 파일 읽기, 데이터 파일로 내보내기

```
write.table(mtcars, "c:/temp/mtcars.data") # mtcars를 지정된 위치에 Excel로 내보내기  
mtcars3<-read.table("c:/temp/mtcars.data") # 다시 읽어들이기
```

① 파일 => 작업디렉토리 변경 => 원하는 폴더로 변경

```
write.table(mtcars, "mtcars.data")         # family를 지정된 위치에 Excel로 내보내기  
mtcars3<-read.table("mtcars.data")        # 다시 읽어들이기
```