Stats 401 Lab 4

401 GSI team

Outline

- Homework feedback
- Review on summation sign
- Sample quiz

Homework feedback - mad_libs in swirl

```
mad_libs <- function(...){
  # unpacking argument
  args <- list(...)

  # assign them to variables
  place <- args[["place"]]
  adjective <- args[["adjective"]]
  noun <- args[["noun"]]

  paste("News from", place, "today where", adjective,
        "students took to the streets in protest of the new
        noun, "being installed on campus.")
}
```

Homework feedback - NA and 0

NA stands for not available, which means the value is missing. NA is not equal to 0. (Suppose x=0, then the value of 0 is known, hence not missing).

```
X = cbind(c(NA,0),c(1,2));X
```

```
##      [,1] [,2]
## [1,]   NA    1
## [2,]    0    2
```

```
Y = matrix(1:4, nrow = 2);Y
```

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

Homework feedback - NA and 0

```
# NA plus/time other value will give us NA
X + Y
```

```
##      [,1] [,2]
## [1,]   NA    4
## [2,]    2    6
```

```
X %*% Y
```

```
##      [,1] [,2]
## [1,]   NA   NA
## [2,]    4    8
```

```
# can use is.na() to check for NA
is.na(X)
```

```
##       [,1]  [,2]
## [1,]  TRUE FALSE
## [2,] FALSE FALSE
```

Homework feedback

Any other questions about the homework?

Review - summation sign

$\sum_{i=1}^{n} x_i = x_1 + x_2 + ... + x_n$. You can always do this expansion if you are uncertain what to do.

Useful results to remember:

- $\sum_{i=1}^{n} cx_i = c \sum_{i=1}^{n} x_i$
- $\sum_{i=1}^{n} c = nc$
- $\frac{d}{dc} \sum_{i=1}^{n} f(x_i, c) = \sum_{i=1}^{n} \frac{d}{dc} f(x_i, c)$

Review - summation sign

Example : Calculate $\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

Review - summation sign

Example : Calculate $\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$ where $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.
Solution:

$$\begin{aligned}
\frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2 &= \frac{1}{n}\sum_{i=1}^{n}(x_i^2 - 2x_i\bar{x} + \bar{x}^2) \\
&= \frac{1}{n}\sum_{i=1}^{n}x_i^2 - \frac{2}{n}\bar{x}\sum_{i=1}^{n}x_i + \frac{1}{n}\sum_{i=1}^{n}\bar{x}^2) \\
&= \frac{1}{n}\sum_{i=1}^{n}x_i^2 - 2(\bar{x})^2 + (\bar{x})^2 \\
&= \frac{1}{n}\sum_{i=1}^{n}x_i^2 - (\bar{x})^2
\end{aligned} \tag{1}$$

Review - summation sign

Example : Suppose $\mathbf{c} = (c_1, ..., c_p)$ and $\mathbf{v} = (v_1, ..., v_p)$, Use Σ notation to evaluate the matrix product $\mathbf{c}^\top \mathbf{v}$

Review - summation sign

Example : Suppose $\mathbf{c} = (c_1, ..., c_p)$ and $\mathbf{v} = (v_1, ..., v_p)$, Use Σ notation to evaluate the matrix product $\mathbf{c}^\top \mathbf{v}$
Solution: $\mathbf{c}^\top \mathbf{v} = \sum_{i=1}^{p} c_i v_i$

Quiz outline

- Test the skills covered in HW 1 to 4
- 50 minutes; start at the beginning of next lab
- Closed book
- In today's lab we will do a sample quiz, which will be similar to the real quiz next week

Sample quiz - Matrix exercises

Suppose we define $\mathbb{A}$ and $\mathbb{B}$ as follows,

```
A
```

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

```
B
```

```
##      [,1] [,2]
## [1,]    1    0
## [2,]   -2    1
```

Calculate the matrices returned by following r command:

1. A %*% B
2. t(A)
3. solve(B)

Sample quiz - Summation exercises

1. Calculate $\sum_{i=k}^{k+5}(i+3)$
2. Calculate $\frac{d}{dm}\sum_{i=1}^{n}(y_i - mx_i)^2$

Sample quiz - R exercises

Which of the following code successfully construct the matrix

$$\mathbb{A} = \begin{bmatrix} 1 & 1 \\ 2 & 2 \\ 3 & 3 \end{bmatrix}$$

A. A <- matrix(c(1,1,2,2,3,3) ,nrow=3)

B. A <- cbind(c(1,1),c(2,2),c(3,3))

C. A <- t(matrix(c(1,1,2,2,3,3) ,nrow=2))

D. A <- c(c(1:3),c(1:3))

Sample quiz - Fitting a linear model by least squares

We look at the uswage data. Recall that

```
## Warning: package 'faraway' was built under R version 3.3
```

```
head(uswages, n=4)
```

```
##          wage educ exper race smsa ne mw so we pt
## 6085   771.60   18    18    0    1  1  0  0  0  0
## 23701  617.28   15    20    0    1  0  0  0  1  0
## 16208  957.83   16     9    0    1  0  0  1  0  0
## 2720   617.28   12    24    0    1  1  0  0  0  0
```

We want to fit a linear model using wage as response, educ and exper as predictors.

Sample quiz - Fitting a linear model by least squares

Which of the following code succesfully construct the matrix $\mathbb{X}$.

A. X <- matrix(uswages$educ, uswages$exper)

B. X <- matrix(rep(1,nrow(uswages)), uswages$educ, uswages$exper)

C. X <- cbind(rep(1,nrow(uswages)), uswages$educ, uswages$exper)

D. X <- cbind(uswages$educ, uswages$exper)

Sample quiz - Fitting a linear model by least squares

If we want to fit the model using R function lm(), which of the following call is correct?

A. lm(wage ~ ., data = uswages)

B. lm(y ~ x, data = uswages)

C. lm(wage = educ + exper, data = uswages)

D. lm(wage educ + exper, data = uswages)

Explain briefly how you would check whether your proposed solution in is correct in R.