# Stats 401 Lab 3

Ed Wu

9/21/2018

# Announcements

- Homework 2 is due today
- Homework without a "Sources" section will receive a zero
- Make sure to staple your homework
- Quiz 1 is on October 5

# Basic matrix computation

- Addition
- Scalar multiplication
- Matrix multiplication
- Inverse
- Transpose

# Addition

- We can add two matrices by adding them together element-wise
- Let $\mathbb{A} = [a_{ij}]_{n \times p}$ and $\mathbb{B} = [b_{ij}]_{n \times p}$, then $\mathbb{A} + \mathbb{B} = [a_{ij} + b_{ij}]_{n \times p}$

For example,

$$\mathbb{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

and

$$\mathbb{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

Then

$$\mathbb{A} + \mathbb{B} = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} \\ a_{21} + b_{21} & a_{22} + b_{22} \end{bmatrix}$$

# Addition

```r
# generate matrices A and B
A = matrix(c(3,-2,-1,4,1,2),nrow=2);A
```

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

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

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

```r
A + B
```

```
##      [,1] [,2] [,3]
## [1,]    4    2    6
## [2,]    0    8    8
```

# Scalar multiplication

- We can multiply a scalar and a matrix together by multiplying each element of the matrix by the scalar
- Let $\mathbb{A} = [a_{ij}]_{n \times p}$ and $s$ be a scalar. Then $s\mathbb{A} = [sa_{ij}]_{n \times p}$

For example,

$$\mathbb{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Then

$$s\mathbb{A} = \begin{bmatrix} sa_{11} & sa_{12} \\ sa_{21} & sa_{22} \end{bmatrix}$$

# Scalar multiplication

```
# Use same matrix A
A
```

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

```
# 5 times A
5 * A
```

```
##      [,1] [,2] [,3]
## [1,]   15   -5    5
## [2,]  -10   20   10
```

# Transpose

- We can transpose a matrix by writing its rows as columns (or columns as rows)
- If $\mathbb{A} = [a_{ij}]_{n\times p}$, then $\mathbb{A}^\top = [a_{ji}]_{p\times n}$

For example,

$$\mathbb{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Then

$$\mathbb{A}^\top = \begin{bmatrix} a_{11} & a_{21} \\ a_{12} & a_{22} \end{bmatrix}$$

# Transpose

We can transpose in R using the function t()

```r
# Recall we have matrix A
A
```

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

```r
# A transpose
C = t(A);C
```

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

# Matrix multiplication

- While matrix addition and scalar multiplication behave as we might expect (element-wise), matrix multiplication is a bit different
- Matrix multiplication does not commute:

$$\mathbb{A}\mathbb{B} \neq \mathbb{B}\mathbb{A}$$

- We can multiply matrices together if the number of columns of the left matrix equals the number of rows of the right matrix

# Matrix multiplication

If $\mathbb{A} = [a_{ij}]_{n \times p}$ and $\mathbb{B} = [b_{ij}]_{p \times q}$, then $\mathbb{A}\mathbb{B} = [c_{ij}]_{n \times q}$ where $c_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}$

For example,

$$\mathbb{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

and

$$\mathbb{B} = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

Then

$$\mathbb{A}\mathbb{B} = \begin{bmatrix} a_{11} b_{11} + a_{12} b_{21} & a_{11} b_{12} + a_{12} b_{22} \\ a_{21} b_{11} + a_{22} b_{21} & a_{21} b_{12} + a_{22} b_{22} \end{bmatrix}$$

# Matrix multiplication

```
# Recall we have matrix B and C
B
```

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

```
C
```

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

Let's calculate BC by hand

# Matrix multiplication

Matrix multiplication is performed in R with the command %*%

```r
# Check with R
B %*% C
```

```
##      [,1] [,2]
## [1,]    5   20
## [2,]    8   24
```

```r
# notice that matrix multiplication is not commutative
C %*% B
```

```
##      [,1] [,2] [,3]
## [1,]   -1    1    3
## [2,]    7   13   19
## [3,]    5   11   17
```

# Identity matrix

- The $n \times n$ identity matrix is the $n \times n$ matrix with 1's on the diagonal and zeros elsewhere. For example, the $2 \times 2$ identity matrix is given by

$$\mathbb{I}_{2\times 2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

- The identity matrix plays the same role as the value 1 does in scalar multiplication. Multiplying a matrix by the identity matrix (of the appropriate dimension) returns the same matrix. For example, if $\mathbb{A}$ is an $n \times 2$ matrix

$$\mathbb{A}_{n\times 2}\mathbb{I}_{2\times 2} = \mathbb{A}_{n\times 2}$$

# Matrix inverse

- The scalar $a$ has an inverse $a^{-1} = \frac{1}{a}$ because $a \times a^{-1} = 1$
- For a matrix $\mathbb{A}$, we call $\mathbb{A}^{-1}$ the inverse of A if $\mathbb{A}\mathbb{A}^{-1} = \mathbb{I}$

# Matrix inverse

For a $2 \times 2$ matrix, we have the following formula for the inverse. Suppose

$$\mathbb{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

Then

$$\mathbb{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

- $det(\mathbb{A}) = a_{11}a_{22} - a_{12}a_{21}$ is called the determinant of $\mathbb{A}$
- If $det(\mathbb{A}) = 0$, then $\mathbb{A}$ is not invertible

# Matrix inverse

We can invert matrices in R using the solve() function

```r
# Generate a matrix
D = matrix(c(1,1,1,3,2,1,3,2,2), nrow=3);D
```

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

```r
# Obtain the inverse of D
solve(D)
```

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

# The Linear Model

Suppose we have collected our response variable $y_1, y_2, \ldots, y_n$ and for each unit $i$, we have $p$ explanatory variables $x_{i1}, x_{i2}, \ldots, x_{ip}$. We can write out the linear model using subscript notation:

$$
\begin{aligned}
y_1 &= b_1 x_{11} + b_2 x_{12} + \cdots + b_p x_{1p} + e_1 \\
y_2 &= b_1 x_{21} + b_2 x_{22} + \cdots + b_p x_{2p} + e_2 \\
&\vdots \\
y_n &= b_1 x_{n1} + b_2 x_{n2} + \cdots + b_p x_{np} + e_n
\end{aligned} \tag{LM1}
$$

# The linear model using matrix notation

- The linear model can also be written in matrix notation
- Define the (column) vectors $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, $\mathbf{e} = (e_1, e_2, \ldots, e_n)$, and $\mathbf{b} = (b_1, b_2, \ldots, b_p)$
- Let the matrix of explanator variables be

$$\mathbb{X} = [x_{ij}]_{n \times p} = \begin{bmatrix} | & | & \ldots & | \\ \mathbf{x}_1 & \mathbf{x}_2 & \ldots & \mathbf{x}_p \\ | & | & \ldots & | \end{bmatrix}$$

where each $\mathbf{x}_j$ is the column vector $(x_{1j}, x_{2j} \ldots, x_{nj})$ corresponding to the $j$-th variable

# The linear model using matrix notation

From class, we know that LM1 is equivalent to

$$\mathbf{y} = \mathbb{X}\mathbf{b} + \mathbf{e}$$

We consider the right hand side first:

$$\mathbb{X}\mathbf{b} + \mathbf{e} = \begin{bmatrix} x_{11} & x_{12} & \dots & x_{1p} \\ x_{21} & x_{22} & \dots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \dots & x_{np} \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_p \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

# The linear model using matrix notation

From class, we know that LM1 is equivalent to

$$\mathbf{y} = \mathbb{X}\mathbf{b} + \mathbf{e}$$

We consider the right hand side first:

$$\mathbb{X}\mathbf{b} + \mathbf{e} = \begin{bmatrix} b_1 x_{11} + b_2 x_{12} + \cdots + b_p x_{1p} \\ b_1 x_{21} + b_2 x_{22} + \cdots + b_p x_{2p} \\ \vdots \\ b_1 x_{n1} + b_2 x_{n2} + \cdots + b_p x_{np} \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}$$

# The linear model using matrix notation

From class, we know that LM1 is equivalent to

$$\mathbf{y} = \mathbb{X}\mathbf{b} + \mathbf{e}$$

We consider the right hand side first:

$$\mathbb{X}\mathbf{b} + \mathbf{e} = \begin{bmatrix} b_1 x_{11} + b_2 x_{12} + \cdots + b_p x_{1p} + e_1 \\ b_1 x_{21} + b_2 x_{22} + \cdots + b_p x_{2p} + e_2 \\ \vdots \\ b_1 x_{n1} + b_2 x_{n2} + \cdots + b_p x_{np} + e_n \end{bmatrix}$$

# The linear model using matrix notation

We therefore see that $\mathbf{y} = \mathbb{X}\mathbf{b} + \mathbf{e}$ is equivalent to LM1:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} b_1 x_{11} + b_2 x_{12} + \cdots + b_p x_{1p} + e_1 \\ b_1 x_{21} + b_2 x_{22} + \cdots + b_p x_{2p} + e_2 \\ \vdots \\ b_1 x_{n1} + b_2 x_{n2} + \cdots + b_p x_{np} + e_n \end{bmatrix}$$

# The linear model using matrix notation

Often, we include an intercept term in the model

Suppose we have $p-1$ predictors. Then we can write $\mathbf{x}_p = (1, \ldots, 1)$ and the resulting linear model will be:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} b_1 x_{11} + b_2 x_{12} + \cdots + b_{p-1} x_{1,p-1} + b_p + e_1 \\ b_1 x_{21} + b_2 x_{22} + \cdots + b_{p-1} x_{2,p-1} + b_p + e_2 \\ \vdots \\ b_1 x_{n1} + b_2 x_{n2} + \cdots + b_{p-1} x_{n,p-1} + b_p + e_n \end{bmatrix}$$

# In-lab activity (part 1)

Suppose we collect data on 5 students. We have the response variable final project score $\mathbf{y} = (90, 65, 69, 79, 85)$, exam 1 score $(87, 86, 73, 65, 90)$, exam 2 score $(100, 70, 76, 76, 90)$

1. Write out the matrix of the explanatory variables assuming the linear model (a) **does not** contain an intercept and (b) **does** contain an intercept
2. Write the same matrices in R. Call the version without an intercept "exams" and the version with an intercept "X"

# In-lab activity (part 1)

1(a)

$$\begin{bmatrix} 87 & 100 \\ 86 & 70 \\ 73 & 76 \\ 65 & 76 \\ 90 & 90 \end{bmatrix}$$

1(b)

$$\begin{bmatrix} 87 & 100 & 1 \\ 86 & 70 & 1 \\ 73 & 76 & 1 \\ 65 & 76 & 1 \\ 90 & 90 & 1 \end{bmatrix}$$

# In-lab activity (part 1)

Question 2

```
exams = matrix(c(87,86,73,65,90,100,70,76,76,90),nrow = 5)
exams
```

```
##      [,1] [,2]
## [1,]   87  100
## [2,]   86   70
## [3,]   73   76
## [4,]   65   76
## [5,]   90   90
```

# In-lab activity (part 1)

Question 2

```
X = cbind(exams,rep(1,5))
X
```

```
##      [,1] [,2] [,3]
## [1,]   87  100    1
## [2,]   86   70    1
## [3,]   73   76    1
## [4,]   65   76    1
## [5,]   90   90    1
```

# Linear model in R

We use the `lm()` command to create a linear model in R.

First, create the explanatory variable:

```
project = c(90,65,69,79,85)
```

# Linear model in R

Next, fit the linear model

```
lmod1 = lm(project ~ exams)
lmod1
```

```
##
## Call:
## lm(formula = project ~ exams)
##
## Coefficients:
## (Intercept)       exams1       exams2
##     23.4567      -0.2502       0.9006
```

"project ∼ exams" is the formula we give to R. It tells us the response variable is "project" and the explanatory variables are contained in "exams". By default, R assumes we want an intercept term so we use the no intercept data in the formula.

# Linear model in R

We can also give the function a data frame. First we create the data frame:

```
df = data.frame(cbind(project,exams))
df
```

```
##   project V2  V3
## 1      90 87 100
## 2      65 86  70
## 3      69 73  76
## 4      79 65  76
## 5      85 90  90
```

# Linear model in R

Next, we fit the linear model

```
lmod2 = lm(project ~ ., data = df)
lmod2
```

```
##
## Call:
## lm(formula = project ~ ., data = df)
##
## Coefficients:
## (Intercept)           V2           V3
##     23.4567      -0.2502       0.9006
```

In this case "project $\sim$ ., data = df" tells R that the data are contained in df, where "project" is the name of the response variable, and "." tells us to use all the remaining variables as predictors. Once again, R includes the intercept for us

# In-lab activity (part 2)

We have calculated the coefficients for the linear model (including an intercept term) using the lm() function and the data from part 1

```r
coef(lmod1)
```

```
## (Intercept)      exams1      exams2
##  23.4566800  -0.2502367   0.9006347
```

From lecture, we know the formula for the coefficients is $\mathbf{b} = (\mathbb{X}^\top \mathbb{X})^{-1} \mathbb{X}^\top \mathbf{y}$. Use R to calculate this quantity for the data from part 1 and compare to the coefficients above

Reminders:

- Make sure to include the intercept term
- solve() inverts a matrix, t() transposes a matrix, and %*% multiplies matrices together

# In-lab activity (part 2)

```r
coefficients = solve(t(X) %*% X) %*% t(X) %*% project
coefficients
```

```
##              [,1]
## [1,] -0.2502367
## [2,]  0.9006347
## [3,] 23.4566800
```

# Lab ticket

1. Suppose $\mathbb{A}$ is a $4 \times 6$ matrix and $\mathbb{B}$ is a $3 \times 6$ matrix.

- Does $\mathbb{A}\mathbb{B}$ exist? If so, what is the dimension of $\mathbb{A}\mathbb{B}$?
- Does $\mathbb{A}\mathbb{B}^\top$ exist? If so, what is the dimension of $\mathbb{A}\mathbb{B}^\top$?

2. Suppose our data are as follows: response variable $\mathbf{y} = (50, 40, 48)$ and one predictor $\mathbf{x} = (12, 6, 10)$

- What is the linear model (with an intercept) in matrix notation? Make sure to write out the full matrices