Statistics One

Lecture 11
Multiple Regression

1

Three segments

- Multiple Regression (MR)
- Matrix algebra
- Estimation of coefficients

2

Lecture 11 ~ Segment 1

Multiple Regression

3

Multiple Regression

- Important concepts/topics
  - Multiple regression equation
  - Interpretation of regression coefficients

4

Simple vs. multiple regression

- Simple regression
  - Just one predictor (X)
- Multiple regression
  - Multiple predictors (X1, X2, X3, …)

5

Multiple regression

- Multiple regression equation
  - Just add more predictors (multiple Xs)

$$\hat{Y} = B_0 + B_1X_1 + B_2X_2 + B_3X_3 + \ldots + B_kX_k$$

$$\hat{Y} = B_0 + \Sigma(B_kX_k)$$

6

Multiple regression

- Multiple regression equation

$\hat{Y}$ = predicted value on the outcome variable Y
B_0 = predicted value on Y when all X = 0
X_k = predictor variables
B_k = unstandardized regression coefficients
Y - $\hat{Y}$ = residual (prediction error)
k = the number of predictor variables

7

Model R and R^2

- R = multiple correlation coefficient
  - R = $r_{\hat{Y}Y}$
  - The correlation between the predicted scores and the observed scores
- R^2
  - The percentage of variance in Y explained by the model

8

Multiple regression: Example

- Outcome measure (Y)
  - Faculty salary (Y)
- Predictors (X1, X2, X3)
  - Time since PhD (X1)
  - Number of publications (X2)
  - Gender (X3)

9

Summary statistics

	M	SD
Salary	$64,115	$17,110
Time	8.09	5.24
Publications	15.49	7.51

10

Multiple regression: Example

- Gender
  - Male = 0
  - Female = 1

11

Multiple regression: Example

- lm(Salary ~ Time + Pubs + Gender)
- $\hat{Y}$ = 46,911 + 1,382(Time) + 502(Pubs) + -3,484(Gender)

12

Table of coefficients

	B	SE	t	β	p
B_0	46,911				
Time	1,382	236	5.86	.42	< .01
Pubs	502	164	3.05	.22	< .01
Gender	-3,484	2,439	-1.43	-.10	.16

$$\hat{Y} = 46{,}911 + 1{,}382(\text{Time}) + 502(\text{Pubs}) + -3{,}484(\text{Gender})$$

13

Multiple regression: Example

- What is $46,911?
- What is $502?
- Who makes more money, men or women?
- According to this model, is the gender difference statistically significant?
- What is the strongest predictor of salary?

14

Multiple regression: Example

- $46,911 is the predicted salary for a male professor who just graduated and has no publications (predicted score when all X=0)
- $502 is the predicted change in salary associated with an increase of one publication, for professors who have been out of school for an average amount of time, averaged across men and women

15

Multiple regression: Example

- Who makes more money, men or women?
  - Trick question: Based on the output we can't answer this question
- According to this model, is the gender difference statistically significant?
  - No

16

Multiple regression: Example

- What is the strongest predictor of salary?
  - Time

17

Segment summary

- Important concepts/topics
  - Multiple regression equation
  - Interpretation of regression coefficients

18

END SEGMENT

19

Lecture 11 ~ Segment 2

Matrix algebra

20

Matrix algebra

- Important concepts/topics
  - – Matrix addition/subtraction/multiplication
  - – Special types of matrices
    - Correlation matrix
    - Sum of squares / Sum of cross products matrix
    - Variance / Covariance matrix

21

Matrix algebra

- A matrix is a rectangular table of known or unknown numbers, for example,

$$M = \begin{pmatrix} 1 & 2 \\ 5 & 1 \\ 3 & 4 \\ 4 & 2 \end{pmatrix}$$

22

Matrix algebra

- The size, or order, of a matrix is given by identifying the number of rows and columns. The order of matrix M is 4x2

$$M = \begin{pmatrix} 1 & 2 \\ 5 & 1 \\ 3 & 4 \\ 4 & 2 \end{pmatrix}$$

23

Matrix algebra

- The transpose of a matrix is formed by rewriting its rows as columns

$$M = \begin{pmatrix} 1 & 2 \\ 5 & 1 \\ 3 & 4 \\ 4 & 2 \end{pmatrix} \qquad M^T = \begin{pmatrix} 1 & 5 & 3 & 4 \\ 2 & 1 & 4 & 2 \end{pmatrix}$$

24

Matrix algebra

- Two matrices may be added or subtracted only if they are of the same order

$$N = \begin{pmatrix} 2 & 3 \\ 4 & 5 \\ 1 & 2 \\ 3 & 1 \end{pmatrix}$$

25

Matrix algebra

- Two matrices may be added or subtracted only if they are of the same order

$$M + N = \begin{pmatrix} 1 & 2 \\ 5 & 1 \\ 3 & 4 \\ 4 & 2 \end{pmatrix} + \begin{pmatrix} 2 & 3 \\ 4 & 5 \\ 1 & 2 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 5 \\ 9 & 6 \\ 4 & 6 \\ 7 & 3 \end{pmatrix}$$

26

Matrix algebra

- Two matrices may be multiplied when the number of columns in the first matrix is equal to the number of rows in the second matrix.
- If so, then we say they are *conformable* for matrix multiplication.

27

Matrix algebra

- Matrix multiplication:

$R = M^T * N \quad R_{ij} = \sum (M^T_{ik} * N_{kj})$

28

Matrix algebra

$$R = M^T * N = \begin{pmatrix} 1\ 5\ 3\ 4 \\ 2\ 1\ 4\ 2 \end{pmatrix} * \begin{pmatrix} 2 & 3 \\ 4 & 5 \\ 1 & 2 \\ 3 & 1 \end{pmatrix} = \begin{pmatrix} 37 & 38 \\ 18 & 21 \end{pmatrix}$$

29

Matrix algebra

- In the next ten slides we will go from a raw dataframe to a correlation matrix!

30

Raw dataframe

$$X_{np} = \begin{pmatrix} 3 & 2 & 3 \\ 3 & 2 & 3 \\ 2 & 4 & 4 \\ 4 & 3 & 4 \\ 4 & 4 & 3 \\ 5 & 4 & 3 \\ 2 & 5 & 4 \\ 3 & 3 & 2 \\ 5 & 3 & 4 \\ 3 & 5 & 4 \end{pmatrix}$$

31

Row vector of sums

$$T_{1p} = 1_{1n} * X_{np} = \begin{pmatrix} 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1\ 1 \end{pmatrix} * \begin{pmatrix} 3 & 2 & 3 \\ 3 & 2 & 3 \\ 2 & 4 & 4 \\ 4 & 3 & 4 \\ 4 & 4 & 3 \\ 5 & 4 & 3 \\ 2 & 5 & 4 \\ 3 & 3 & 2 \\ 5 & 3 & 4 \\ 3 & 5 & 4 \end{pmatrix} = \begin{pmatrix} 34\ 35\ 34 \end{pmatrix}$$

32

Row vector of means

$$M_{1p} = T_{1p} * N^{-1} = \begin{bmatrix} 34 & 35 & 34 \end{bmatrix} * 10^{-1} = \begin{bmatrix} 3.4 & 3.5 & 3.4 \end{bmatrix}$$

33

Matrix of means

$$M_{np} = 1_{n1} * M_{1p} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} * \begin{bmatrix} 3.4 & 3.5 & 3.4 \end{bmatrix} = \begin{pmatrix} 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \end{pmatrix}$$

34

Matrix of deviation scores

$$D_{np} = X_{np} - M_{np} = \begin{pmatrix} 3 & 2 & 3 \\ 3 & 2 & 3 \\ 2 & 4 & 4 \\ 4 & 3 & 4 \\ 4 & 4 & 3 \\ 5 & 4 & 3 \\ 2 & 5 & 4 \\ 3 & 3 & 2 \\ 5 & 3 & 4 \\ 3 & 5 & 4 \end{pmatrix} - \begin{pmatrix} 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \end{pmatrix} = \begin{pmatrix} -.4 & -1.5 & -.4 \\ -.4 & -1.5 & -.4 \\ -1.4 & .5 & .6 \\ .6 & -.5 & .6 \\ .6 & .5 & -.4 \\ 1.6 & .5 & -.4 \\ -1.4 & 1.5 & .6 \\ -.4 & -.5 & -1.4 \\ 1.6 & -.5 & .6 \\ -.4 & 1.5 & .6 \end{pmatrix}$$

35

SS / SP matrix

$$S_{xx} = D_{pn}^{T} * D_{np} =$$

$$\begin{pmatrix} -.4 & -.4 & -1.4 & .6 & .6 & 1.6 & -1.4 & -.4 & 1.6 & -.4 \\ -1.5 & -1.5 & .5 & -.5 & .5 & .5 & 1.5 & -.5 & -.5 & 1.5 \\ -.4 & -.4 & .6 & .6 & -.4 & -.4 & .6 & -1.4 & .6 & .6 \end{pmatrix} * \begin{pmatrix} -.4 & -1.5 & -.4 \\ -.4 & -1.5 & -.4 \\ -1.4 & .5 & .6 \\ .6 & -.5 & .6 \\ .6 & .5 & -.4 \\ 1.6 & .5 & -.4 \\ -1.4 & 1.5 & .6 \\ -.4 & -.5 & -1.4 \\ 1.6 & -.5 & .6 \\ -.4 & 1.5 & .6 \end{pmatrix}$$

$$= \begin{pmatrix} 10.4 & -2.0 & -.6 \\ -2.0 & 10.5 & 3.0 \\ -.6 & 3.0 & 4.4 \end{pmatrix}$$

Variance / Covariance matrix

$$C_{xx} = S_{xx} * N^{-1} = \begin{pmatrix} 10.4 & -2.0 & -.6 \\ -2.0 & 10.5 & 3.0 \\ -.6 & 3.0 & 4.4 \end{pmatrix} * 10^{-1} = \begin{pmatrix} 1.04 & -.20 & -.06 \\ -.20 & 1.05 & .30 \\ -.06 & .30 & .44 \end{pmatrix}$$

37

SD matrix

$$S_{xx} = (\mathrm{Diag}(C_{xx}))^{1/2} = \begin{pmatrix} 1.02 & 0 & 0 \\ 0 & 1.02 & 0 \\ 0 & 0 & .66 \end{pmatrix}$$

38

Correlation matrix

$$R_{xx} = S_{xx}^{-1} * C_{xx} * S_{xx}^{-1} =$$

$$\begin{pmatrix} 1.02^{-1} & 0 & 0 \\ 0 & 1.02^{-1} & 0 \\ 0 & 0 & .66^{-1} \end{pmatrix} * \begin{pmatrix} 1.04 & -.20 & -.06 \\ -.20 & 1.05 & .30 \\ -.06 & .30 & .44 \end{pmatrix} * \begin{pmatrix} 1.02^{-1} & 0 & 0 \\ 0 & 1.02^{-1} & 0 \\ 0 & 0 & .66^{-1} \end{pmatrix}$$

$$= \begin{pmatrix} 1.00 & -.19 & -.09 \\ -.19 & 1.00 & .44 \\ -.09 & .44 & 1.00 \end{pmatrix}$$

39

Matrix algebra

- Important concepts/topics
  - Matrix addition/subtraction/multiplication
  - Special types of matrices
    - Correlation matrix
    - Sum of squares / Sum of cross products matrix
    - Variance / Covariance matrix

40

END SEGMENT

Lecture 11 ~ Segment 3

Estimation of coefficients

Estimation of coefficients

- The values of the coefficients (B) are estimated such that the model yields optimal predictions
  - Minimize the residuals!

Estimation of coefficients

- The sum of the squared (SS) residuals is minimized
  - SS.RESIDUAL = $\Sigma(\hat{Y} - Y)^2$

Estimation of coefficients

- Standardized and in matrix form, the regression equation is $\hat{Y} = B(X)$, where
  - $\hat{Y}$ is a [N x 1] vector
    - N = number of cases
  - B is a [k x 1] vector
    - k = number of predictors
  - X is a [N x k] matrix

45

Estimation of coefficients

- $\hat{Y} = B(X)$
  - To solve for B
  - Replace $\hat{Y}$ with Y
  - Conform for matrix multiplication:
    - $Y = X(B)$

46

Estimation of coefficients

- $Y = X(B)$
- Now let's make X square and symmetric
- To do this, pre-multiply both sides of the equation by the transpose of X, X^T

47

Estimation of coefficients

- $Y = X(B)$ becomes
- $X^T(Y) = X^T(XB)$
- Now, to solve for B, eliminate X^TX
- To do this, pre-multiply by the inverse, $(X^TX)^{-1}$

48

Estimation of coefficients

- $X^TY = X^T(XB)$ becomes
- $(X^TX)^{-1}(X^TY) = (X^TX)^{-1}(X^TXB)$
  - Note that $(X^TX)^{-1}(X^TX) = I$
  - And $IB = B$
- Therefore, $(X^TX)^{-1}(X^TY) = B$

49

Estimation of coefficients

- $B = (X^TX)^{-1}(X^TY)$

50

Estimation of coefficients

- $B = (X^TX)^{-1}(X^TY)$
- Let's use this formula to calculate B's from the raw data matrix used in the previous segment

51

Raw data matrix

$$X_{np} = \begin{pmatrix} 3 & 2 & 3 \\ 3 & 2 & 3 \\ 2 & 4 & 4 \\ 4 & 3 & 4 \\ 4 & 4 & 3 \\ 5 & 4 & 3 \\ 2 & 5 & 4 \\ 3 & 3 & 2 \\ 5 & 3 & 4 \\ 3 & 5 & 4 \end{pmatrix}$$

52

Row vector of sums

$$T_{1p} = 1_{1n} * X_{np} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} * \begin{bmatrix} 3 & 2 & 3 \\ 3 & 2 & 3 \\ 2 & 4 & 4 \\ 4 & 3 & 4 \\ 4 & 4 & 3 \\ 5 & 4 & 3 \\ 2 & 5 & 4 \\ 3 & 3 & 2 \\ 5 & 3 & 4 \\ 3 & 5 & 4 \end{bmatrix} = \begin{bmatrix} 34 & 35 & 34 \end{bmatrix}$$

53

Row vector of means

$$M_{1p} = T_{1p} * N^{-1} = \begin{bmatrix} 34 & 35 & 34 \end{bmatrix} * 10^{-1} = \begin{bmatrix} 3.4 & 3.5 & 3.4 \end{bmatrix}$$

54

Matrix of means

$$M_{np} = 1_{n1} * M_{1p} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} * \begin{bmatrix} 3.4 & 3.5 & 3.4 \end{bmatrix} = \begin{bmatrix} 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \end{bmatrix}$$

55

Matrix of deviation scores

$$D_{np} = X_{np} - M_{np} = \begin{bmatrix} 3 & 2 & 3 \\ 3 & 2 & 3 \\ 2 & 4 & 4 \\ 4 & 3 & 4 \\ 4 & 4 & 3 \\ 5 & 4 & 3 \\ 2 & 5 & 4 \\ 3 & 3 & 2 \\ 5 & 3 & 4 \\ 3 & 5 & 4 \end{bmatrix} - \begin{bmatrix} 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \\ 3.4 & 3.5 & 3.4 \end{bmatrix} = \begin{bmatrix} -.4 & -1.5 & -.4 \\ -.4 & -1.5 & -.4 \\ -1.4 & .5 & .6 \\ .6 & -.5 & .6 \\ .6 & .5 & -.4 \\ 1.6 & .5 & -.4 \\ -1.4 & 1.5 & .6 \\ -.4 & -.5 & -1.4 \\ 1.6 & -.5 & .6 \\ -.4 & 1.5 & .6 \end{bmatrix}$$

56

SS & SP matrix

$S_{xx} = D_{pn}^{T} * D_{np} =$

$$\begin{pmatrix} -.4 & -.4 & -1.4 & .6 & .6 & 1.6 & -1.4 & -.4 & 1.6 & -.4 \\ -1.5 & -1.5 & .5 & -.5 & .5 & .5 & 1.5 & -.5 & -.5 & 1.5 \\ -.4 & -.4 & .6 & .6 & -.4 & -.4 & .6 & -1.4 & .6 & .6 \end{pmatrix} * \begin{pmatrix} -.4 & -1.5 & -.4 \\ -.4 & -1.5 & -.4 \\ -1.4 & .5 & .6 \\ .6 & -.5 & .6 \\ .6 & .5 & -.4 \\ 1.6 & .5 & -.4 \\ -1.4 & 1.5 & .6 \\ -.4 & -.5 & -1.4 \\ 1.6 & -.5 & .6 \\ -.4 & 1.5 & .6 \end{pmatrix}$$

$$= \begin{pmatrix} 10.4 & -2.0 & -.6 \\ -2.0 & 10.5 & 3.0 \\ -.6 & 3.0 & 4.4 \end{pmatrix}$$

SS & SP matrix

Since we used deviation scores:
Substitute S_{xx} for $X^{T}X$
Substitute S_{xy} for $X^{T}Y$

Therefore,

$B = (S_{xx})^{-1}S_{xy}$

Estimation of coefficients

$B = (S_{xx})^{-1}S_{xy}$

$$B = \begin{pmatrix} 10.5 & 3.0 \\ 3.0 & 4.4 \end{pmatrix}^{-1} \begin{pmatrix} -2.0 \\ -.6 \end{pmatrix} = \begin{pmatrix} -0.19 \\ -0.01 \end{pmatrix}$$

Estimation of coefficients

```
Call:
lm(formula = demo$Y ~ demo$X1 + demo$X2)

Residuals:
    Min      1Q  Median      3Q     Max
-1.3011 -0.6855 -0.3091  0.6458  1.6909

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  4.086022   2.034669   2.008   0.0846 .
demo$X1     -0.188172   0.411447  -0.457   0.6613
demo$X2     -0.008065   0.635598  -0.013   0.9902
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 1.196 on 7 degrees of freedom
Multiple R-squared: 0.03665,  Adjusted R-squared: -0.2386
F-statistic: 0.1332 on 2 and 7 DF,  p-value: 0.8775
```

END SEGMENT

END LECTURE 11