

# Math Final

Matthew Vanaman

05-19-2019

## 7.11

(a)

$$\frac{\partial}{\partial x_1} = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 x_{i2}^3 + \beta_5 x_{i3} + \beta_6 x_{i4} + \beta_7 x_{i3} x_{i4},$$

Derivative of a constant is zero, so drop  $\alpha$ :

$$= \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 x_{i2}^3 + \beta_5 x_{i3} + \beta_6 x_{i4} + \beta_7 x_{i3} x_{i4},$$

Break up across addition and subtraction, then bring out constants:

$$= \beta_1 \frac{\partial}{\partial x_1} x_{i1} + \beta_2 \frac{\partial}{\partial x_1} x_{i2} + \beta_3 \frac{\partial}{\partial x_1} x_{i2}^2 + \beta_4 \frac{\partial}{\partial x_1} x_{i2}^3 + \beta_5 \frac{\partial}{\partial x_1} x_{i3} + \beta_6 \frac{\partial}{\partial x_1} x_{i4} + \beta_7 \frac{\partial}{\partial x_1} x_{i3} \frac{\partial}{\partial x_1} x_{i4},$$

All other independent variables are treated as constants:

$$= \frac{\partial}{\partial x_1} \beta_1 x_{i1} + 1(0) + 1(0) + 1(0) + 1(0) + 1(0) + 1(0)(0),$$

Bring the constant out:

$$= \beta_1 \frac{\partial}{\partial x_1} x_{i1} = \beta_1(1) = \beta_1.$$

(b)

As long as the  $x$  variable is not exponentiated or multiplied, the coefficient is the partial derivative. This is not true for exponentiated terms. Take for example:

$$\begin{aligned} & \frac{d}{dx}(a + \beta x^2), \\ &= \frac{d}{dx}(a) + \frac{d}{dx}(\beta x^2), \\ &= \beta \frac{d}{dx}(x^2), \\ &= \beta(2x^{2-1}), \\ &= 2\beta x. \end{aligned}$$

This derivative does not yield the coefficient by itself, so you cannot interpret the coefficient as a straightforward linear relationship between  $x$  and  $y$  in the case of exponentiated terms. When the slope is exponential, the partial derivative gives you the instantaneous rate of change at each level of  $x$  which, unlike in the case of (a), will be different at each level of  $x$ .

Multiplicative relationships also cannot be interpreted as a straightforward linear relationship. Consider a model with two predictors,  $x_1$  and  $x_2$ , and their interaction:

$$\frac{\partial}{\partial x_1}(a + \beta_1 x_1 + \beta_2 x_2 + \beta_{12}(x_1 \times x_2)) = \beta_1 + \beta_{12} x_2.$$

Here, the first derivative with respect to  $x_1$  is a function of  $x_2$ . If there is an interaction, this partial derivative will change when evaluated across differing  $x_2$  values. The key thing to notice is that because this function incorporates information about  $x_2$ , it does not describe an exclusive relationship between  $x_1$  and  $y$  like function (a) does. Therefore, when an interaction term is present, the first partial derivative does not yield the regression coefficient  $\beta_1$ . However, if this derivative does not change across differing levels of  $x_2$ , there is no interaction. In cases where there is no interaction (i.e. in cases where  $\beta_{12}x_2$  comes out to zero), this function *would* describe the linear relationship between  $x_1$  and  $y$  in the way function (a) does. This would also be true whenever  $x_2 = 0$ . However, mathematically, this derivative does not directly represent the coefficient one sees in regression output for  $x_1$  in a main-effects only model, nor does it yield the coefficient for the interaction.

To get the coefficient for the interaction, you need to take the *cross-partial* derivative with respect to both  $x_1$  and  $x_2$ :

$$\frac{\partial^2}{\partial x_1 \partial x_2}(a + \beta_1 x_1 + \beta_2 x_2 + \beta_{12}(x_1 \times x_2)) = \beta_{12}.$$

Importantly, this coefficient does not describe an exclusive relationship between  $x_1$  and  $y$  since it is a function of both  $x_1$  and  $x_2$ . The cross-partial derivative introduces a third slope, which is the slope between  $x_1$  and  $y$  spanning across all levels of  $x_2$ . If this slope is zero, you have no interaction.

In sum, the coefficient seen in regression output describes a linear relationship between  $x$  and  $y$  in a linear main-effects model but the meaning of this coefficient changes when  $x$  is exponentiated or multiplied by another independent variable.

(c)

$$\frac{\partial}{\partial x_2} = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 x_{i2}^3 + \beta_5 x_{i3} + \beta_6 x_{i4} + \beta_7 x_{i3} x_{i4},$$

Derivative of a constant is zero, so drop  $\alpha$ :

$$= \frac{\partial}{\partial x_2}(\beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 x_{i2}^3 + \beta_5 x_{i3} + \beta_6 x_{i4} + \beta_7 x_{i3} x_{i4}),$$

Break up across addition and subtraction, then again across multiplication:

$$= \frac{\partial}{\partial x_2} \beta_1 x_{i1} + \frac{\partial}{\partial x_2} \beta_2 x_{i2} + \frac{\partial}{\partial x_2} \beta_3 x_{i2}^2 + \frac{\partial}{\partial x_2} \beta_4 x_{i2}^3 + \frac{\partial}{\partial x_2} \beta_5 x_{i3} + \frac{\partial}{\partial x_2} \beta_6 x_{i4} + \frac{\partial}{\partial x_2} \beta_7 x_{i3} x_{i4},$$

The independent variables all get treated as constants:

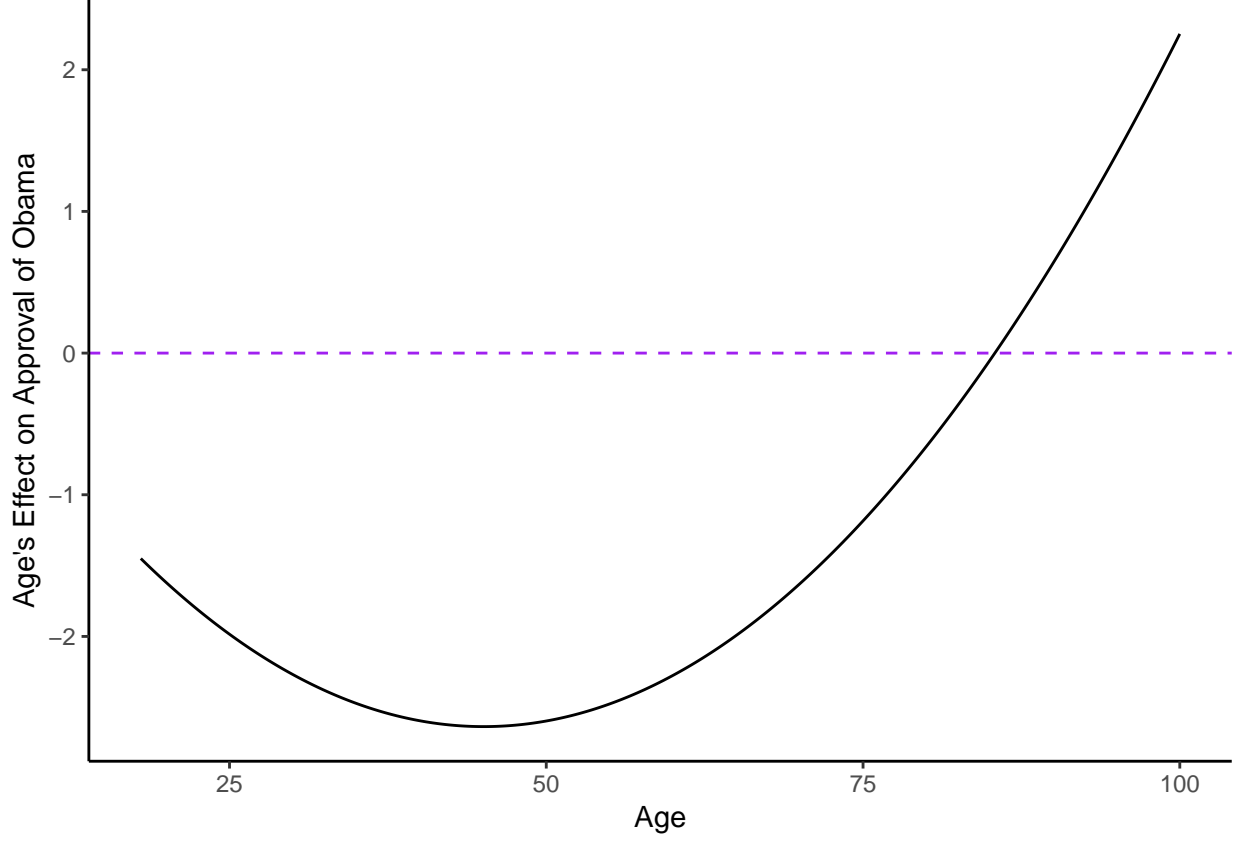
$$\begin{aligned} &= 0 + \frac{\partial}{\partial x_2} \beta_2 x_{i2} + \frac{\partial}{\partial x_2} \beta_3 x_{i2}^2 + \frac{\partial}{\partial x_2} \beta_4 x_{i2}^3 + 0 + 0 + 0 + 0, \\ &= \beta_2 \frac{\partial}{\partial x_2} x_{i2} + \beta_3 \frac{\partial}{\partial x_2} x_{i2}^2 + \beta_4 \frac{\partial}{\partial x_2} x_{i2}^3, \\ &= \beta_2 + 2\beta_3 x_{i2} + 3\beta_4 x_{i2}^2. \end{aligned}$$

(d)

On the x-axis, you will age. On the y-axis, you will have:

$$\begin{aligned} &0.653 + 2(0.073)x_{i2} + 3(0.00054)x_{i2}^2, \\ &= 0.6530.146x_{i2} + 0.00162x_{i2}^2. \end{aligned}$$

Voting age starts at 18, so make age run from 18 to 100, let's say.



There is a curvilinear relationship such that generally speaking, age is negatively associated with approval. That is, the slope between age and approval is negative everywhere except for those older than around 80 years old or so. At that age, the effect turns around, and age becomes positively predictive of approval. The effect of age on approval is most negative at middle age.

(e)

$$\frac{\partial}{\partial x_4} = \alpha + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 x_{i2}^3 + \beta_5 x_{i3} + \beta_6 x_{i4} + \beta_7 x_{i3} x_{i4},$$

Derivative of a constant is zero, so drop  $\alpha$ :

$$\begin{aligned} &= \frac{\partial}{\partial x_4} (\beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + \beta_4 x_{i2}^3 + \beta_5 x_{i3} + \beta_6 x_{i4} + \beta_7 x_{i3} x_{i4}), \\ &= \frac{\partial}{\partial x_4} (\beta_1 x_{i1} + \frac{\partial}{\partial x_4} \beta_2 x_{i2} + \frac{\partial}{\partial x_4} \beta_3 x_{i2}^2 + \frac{\partial}{\partial x_4} \beta_4 x_{i2}^3 + \frac{\partial}{\partial x_4} \beta_5 x_{i3} + \frac{\partial}{\partial x_4} \beta_6 x_{i4} + \frac{\partial}{\partial x_4} \beta_7 x_{i3} x_{i4}), \\ &= 0 + 0 + 0 + 0 + 0 + 0 + \frac{\partial}{\partial x_4} \beta_6 x_{i4} + \frac{\partial}{\partial x_4} \beta_7 x_{i3} x_{i4}, \\ &= \beta_6 \frac{\partial}{\partial x_4} x_{i4} + \beta_7 x_{i3} \frac{\partial}{\partial x_4} x_{i4}, \\ &= \beta_6 + \beta_7 x_{i3}. \end{aligned}$$

(f)

$$= \beta_6 + \beta_7 x_{i3} = 1 + 3x_{i3}.$$

Dummy code sex and make male = 0 and female = 1:

$$\text{male} = 1 + 3(0) = 1.$$

$$\text{female} = 1 + 3(1) = 3.$$

(g)

In both males and females, the derivative is positive in sign, indicating that there is a positive slope between pro-choice views and approval of Obama. However, there is also a sex difference: for females, the effect is 3 whereas for males it is 1; what this means is that the slope between abortion views and approval of Obama is steeper for women than for men. Because there is a difference in slopes of sex, we can say there is an interaction. In other words, there is a compounding effect of sex such that when you move from the male to female level, you see an increase in the strength of the relationship between abortion views and approval of Obama (descriptively speaking anyway).

## 9.10

(a)

$$\begin{aligned}
 X'Y &= \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 0 & 1 & 3 & 11 & 0 & 5 & 0 \\ 5 & 1 & 0 & 2 & 0 & 0 & 0 & 0 & 12 & 2 \end{bmatrix} \begin{bmatrix} 10 \\ 7 \\ 2 \\ 8 \\ 10 \\ 6 \\ 2 \\ 10 \\ 8 \\ 5 \end{bmatrix}, \\
 &= 1(10) + 1(7) + 1(2) + 1(8) + 1(10) + 1(6) + 1(2) + 1(10) + 1(8) + 1(5) = 68, \\
 &\quad \text{etc...} \\
 &= \begin{bmatrix} 68 \\ 121 \\ 179 \end{bmatrix}.
 \end{aligned}$$

(b)

$$X'X = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 1 & 2 & 0 & 1 & 3 & 11 & 0 & 5 & 0 \\ 5 & 1 & 0 & 2 & 0 & 0 & 0 & 0 & 12 & 2 \end{bmatrix} \times \begin{bmatrix} 1 & 2 & 5 \\ 1 & 1 & 1 \\ 1 & 2 & 0 \\ 1 & 0 & 2 \\ 1 & 1 & 0 \\ 1 & 3 & 0 \\ 1 & 11 & 0 \\ 1 & 0 & 0 \\ 1 & 5 & 12 \\ 1 & 0 & 2 \end{bmatrix},$$

$$1(1) + 1(1) + 1(1) + 1(1) + 1(1) + 1(1) + 1(1) + 1(1) + 1(1) + 1(1) = 10,$$

etc...

$$\begin{bmatrix} 10 & 25 & 22 \\ 25 & 165 & 71 \\ 22 & 71 & 178 \end{bmatrix}.$$

(c)

Follow the formula for the inverse of a large square matrix:

$$\frac{1}{|A|} \text{adj}(A) = \frac{1}{|X'X|} \text{adj}(X'X).$$

First we have to get the adjoint matrix. To get the adjoint matrix, we have to find the minor elements, followed by the cofactor matrix, then we can derive the adjoint matrix.

To get the minor elements, we find the determinant of each possible 2 x 2 matrix. The determinant matrix is:

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc,$$

which we will have to do for all of the possible matrices. Starting with element (1,1):

$$e_{11} = \begin{bmatrix} 165 & 71 \\ 71 & 178 \end{bmatrix} = 165(178) - 71(71) = 24,329.$$

We continue this all the way down the line:

$$e_{12} = \begin{vmatrix} 25 & 71 \\ 22 & 178 \end{vmatrix} = 2888.$$

$$e_{13} = \begin{vmatrix} 25 & 22 \\ 165 & 71 \end{vmatrix} = -1855.$$

$$e_{21} = \begin{vmatrix} 25 & 22 \\ 71 & 178 \end{vmatrix} = 2888.$$

$$e_{22} = \begin{vmatrix} 10 & 22 \\ 22 & 178 \end{vmatrix} = 1296.$$

$$e_{23} = \begin{vmatrix} 10 & 25 \\ 22 & 71 \end{vmatrix} = 160.$$

$$e_{31} = \begin{vmatrix} 25 & 22 \\ 165 & 71 \end{vmatrix} = -1855.$$

$$e_{32} = \begin{vmatrix} 10 & 25 \\ 22 & 71 \end{vmatrix} = 160.$$

$$e_{33} = \begin{vmatrix} 10 & 25 \\ 25 & 165 \end{vmatrix} = 1025.$$

$$\text{Minor Elements Matrix} = \begin{bmatrix} 24329 & 2888 & -1855 \\ 2888 & 1296 & 160 \\ -1855 & 160 & 1025 \end{bmatrix}.$$

To get the cofactor matrix, you just multiple the cells by -1 if their row and columns number add up to an odd number:

$$\text{Cofactor Matrix} = \begin{bmatrix} 24329 & -2888 & -1855 \\ -2888 & 1296 & -160 \\ -1855 & -160 & 1025 \end{bmatrix}.$$

Lastly, you derive the adjoint matrix by transposing the cofactor matrix. But since it's symmetric, the cofactor matrix equals the adjoint matrix:

$$\text{Adjoint Matrix} = \begin{bmatrix} 24329 & -2888 & -1855 \\ -2888 & 1296 & -160 \\ -1855 & -160 & 1025 \end{bmatrix}.$$

Now that we have the adjoint matrix, we use it to find the determinant of the  $X'X$  matrix. We do this by multiplying a row or column of  $X'X$  by its corresponding elements in the cofactor matrix and then sum the product.

$$\begin{bmatrix} 25 \\ 165 \\ 71 \end{bmatrix} \times \begin{bmatrix} -2888 \\ 1296 \\ -160 \end{bmatrix} = \sum \begin{bmatrix} -72200 \\ 213840 \\ -11360 \end{bmatrix} = 130280.$$

Last, now that we have  $|X'X|$ , we can use the formula for matrix inversion:

$$\begin{aligned} (X'X)^{-1} &= \frac{1}{|X'X|} \text{adj}(X'X), \\ &= \frac{1}{130280} \times \begin{bmatrix} 24329 & -2888 & -1855 \\ -2888 & 1296 & -160 \\ -1855 & -160 & 1025 \end{bmatrix} = \begin{bmatrix} 0.187 & -0.022 & -0.014 \\ -0.022 & 0.01 & -0.001 \\ -0.014 & -0.001 & 0.008 \end{bmatrix}. \end{aligned}$$

(d)

$$\begin{aligned} \hat{\beta} &= (X'X)^{-1}(X'Y), \\ &= \begin{bmatrix} 0.187 & -0.022 & -0.014 \\ -0.022 & 0.01 & -0.001 \\ -0.014 & -0.001 & 0.008 \end{bmatrix} \times \begin{bmatrix} 68 \\ 121 \\ 179 \end{bmatrix}, \\ &= \begin{bmatrix} 7.468 \\ -0.524 \\ 0.291 \end{bmatrix} = \begin{bmatrix} \alpha \\ \beta_1 \\ \beta_2 \end{bmatrix}. \end{aligned}$$

## Check with Regression

	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	7.4676	1.1477	6.51	0.0003
Sanctions	-0.5235	0.2649	-1.98	0.0886
Riots	0.2915	0.2356	1.24	0.2559

## 10.9