Multiple Linear Regression Model Selection

William Brasic

The University of Arizona

Models with Quadratics

Definition 1: Models with Quadratics

A linear regression model with quadratics is when the outcome variable is a squared function of at least one of the independent variables.

Models with Quadratics

Example 1: Models with Quadratics

An example of a linear regression model with quadratics is given by

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 + u_i.$$

Models with Quadratics

Question 1

If $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 u_i$, then what is the marginal impact of x_2 on y?

Models with Quadratics

Question 1

If $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i2}^2 u_i$, then what is the marginal impact of x_2 on y?

Answer to Question 1

We take the derivative as usual, but it looks different now:

$$\frac{\partial y}{\partial x_2} = \beta_2 + 2\beta_3 x_2.$$

Models with Quadratics

Question 2

When would a quadratic linear regression model be useful?

Models with Quadratics

Question 2

When would a quadratic linear regression model be useful?

Answer to Question 2

When we think our outcome variable is initially an increasing function of a covariate, but after some point begins decreasing.

- Think about when a person's income is the outcome variable and their age is a covariate.

Models with Quadratics

Property 1: Models with Quadratics

Suppose we estimate the impact of years of experience on wage and obtain the regression equation

$$\widehat{wage}_i = \widehat{\beta}_0 + \widehat{\beta}_1 exper_i + \widehat{\beta}_2 exper_i^2.$$

We would expect $\widehat{\beta}_2$ to be negative.

- The negative component accounts for experience having a diminishing effect on wage.

Models with Quadratics

Question 3

Suppose we estimate the impact of years of experience on yearly income and obtain the regression equation

$$\begin{aligned}\widehat{income}_i &= \widehat{\beta}_0 + \widehat{\beta}_1 exper_i + \widehat{\beta}_2 exper_i^2 \\ &= 30000 + 10000 exper_i - 250 exper_i^2.\end{aligned}$$

1. What does this model predict someone with zero years of experience will make per year?
2. What is the effect of the first year of experience on yearly income?
3. What is the effect of the fifth year of experience on yearly income?

Models with Quadratics

Answer to Question 3

$$\frac{\partial \widehat{income}}{\partial exper} = 10000 - 500exper.$$

1. When an individual has zero years of experience, the model predicts their yearly income is $\$30,000$.
2. The effect of the first year of experience on yearly income is an increase of $10000 - 500(1) = \$9,500$
3. The effect of the fifth year of experience on yearly income is an increase of $10000 - 500(5) = \$7,500$.

Models with Quadratics

Question 4

Suppose we estimate the impact of years of experience on yearly income and obtain the regression equation

$$\begin{aligned}\widehat{income}_i &= \widehat{\beta}_0 + \widehat{\beta}_1 exper_i + \widehat{\beta}_2 exper_i^2 \\ &= 30000 + 10000 exper_i - 250 exper_i^2\end{aligned}$$

For what value of experience is yearly income predicted to be maximized?

Models with Quadratics

Answer to Question 4

Use the first order condition!

$$\frac{\partial \widehat{wage}}{\partial exper} = 10000 - 500exper = 0 \iff 10000 = 500exper$$
$$\iff exper^* = 20.$$

Thus, the model predicts a person will maximize their yearly income with twenty years of experience.

Models with Interactions

Definition 2: Models with Interactions

A linear regression model with interactions is when the outcome variable is a function of the product of at least two covariates.

Models with Interactions

Example 2: Models with Interactions

An example of a linear regression model with interactions is given by

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i1} x_{i2} + u_i.$$

Models with Interactions

Question 5

If $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i1} x_{i2} u_i$, then what is the marginal impact of x_2 on y?

Models with Interactions

Question 5

If $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i1} x_{i2} u_i$, then what is the marginal impact of x_2 on y?

Answer to Question 5

We take the derivative as usual, but it looks different now:

$$\frac{\partial y}{\partial x_2} = \beta_2 + \beta_3 x_1.$$

Models with Interactions

Question 6

When would a linear regression model with interactions be useful?

Models with Interactions

Question 6

When would a linear regression model with interactions be useful?

Answer to Question 6

When we think the effect of one covariate on our outcome depends on another covariate.

Models with Interactions

Property 2: Models with Interactions

Suppose we estimate the impact of the number of bedrooms a home has on its price and obtain the regression equation

$$\widehat{price}_i = \widehat{\beta}_0 + \widehat{\beta}_1 sqrft + \widehat{\beta}_2 bdrms + \widehat{\beta}_3 sqrft * bdrms.$$

1. $\widehat{\beta}_3 > 0$ implies the estimated effect of additional bedrooms is larger for bigger homes.
2. $\widehat{\beta}_3 < 0$ implies the estimated effect of additional bedrooms is smaller for bigger homes.

Models with Interactions

Question 7

Suppose we estimate the impact of the square footage of a home on its price and obtain the regression equation

$$\widehat{price}_i = 20000 + 250sqrft + 20000bdrms + 100sqrft * bdrms.$$

What is the estimated effect of an additional square foot on the price of a home with four bedrooms?

Models with Interactions

Answer to Question 7

$$\frac{\partial \widehat{income}}{\partial exper} = 250 + 100bdrms$$

The effect of an additional square foot on the price of a four bedroom home is an increase in the homes price of $250 + 100(4) = \$650$.

Average Partial Effects (APE)

Question 8

If we have the model $y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \beta_3 x_{i1}^2 + u_i$, we know upon estimation we will be able to obtain

$$\frac{\partial \widehat{y}}{\partial x_1} = \widehat{\beta}_1 + 2\widehat{\beta}_3 x_1.$$

To get the estimated effect of x_1 on y, what should we use for x_1?

- In a model with quadratics, interactions, etc., this isn't an issue and we would simply get the estimated effect as $\widehat{\beta}_1$.

Average Partial Effects (APE)

Definition 3: Average Partial Effects (APE)

The average partial effect (APE) of x_1 on y is given by:

1. Taking the derivative of y with respect to x_1
2. If an x_1 term still exists, enter its mean to obtain the APE

- Could we also use the median?

Average Partial Effects (APE)

Example 3: Average Partial Effects (APE)

Suppose we estimate the impact of years of experience on yearly income and obtain the regression equation

$$\begin{aligned}\widehat{income}_i &= \widehat{\beta}_0 + \widehat{\beta}_1 exper_i + \widehat{\beta}_2 exper_i^2 \\ &= 30000 + 10000 exper_i - 250 exper_i^2.\end{aligned}$$

If the average experience is ten years, then the APE of years of experience on yearly income is

$$\begin{aligned}\left.\frac{\partial \widehat{income}}{\partial exper}\right|_{\overline{exper}} &= 10000 - 500\overline{exper} \\ &= 5000.\end{aligned}$$

Nested Models

Example 4: Nested Models

The model

$$income_i = \beta_0 + \beta_1 exper_i + u_i$$

is nested within

$$income_i = \beta_0 + \beta_1 exper_i + \beta_2 age_i + \beta_3 educ_i + v_i$$

as the former contains a subset of the latter's covariates.

- We know we can use an F-test to choose between these models.

Non-Nested Models

Definition 4: Non-Nested models

Two models are non-nested when neither model is a special case of the other.

Example 5: Non-Nested models

The models

$$income_i = \beta_0 + \beta_1 age_i + \beta_2 exper_i + u_i$$
$$income_i = \beta_0 + \beta_1 age_i + \beta_2 educ_i + v_i$$

are non-nested.

- Can't use an F-test here so what do we do?

Non-Nested Models

Property 3: Non-Nested Models

To choose between two non-nested models, we can:

1. Estimate both and obtain their adjusted R-squared
2. Choose the model with the higher adjusted R-squared

- Recall that adjusted R-squared is a more robust version of the regular R-squared. Why?
- This is an example of a model selection technique
  - Model selection is critical in machine learning.

Thank You!