Lecture 17

Models with Quadratics

As your author states, "Quadratic functions are also used quite often in applied economics to capture decreasing or increasing marginal effects." (p. 189)

Consider the following quadratic regression function:

$$y = \beta_0 + \beta_1 X + \beta_2 X^2 + u \, .$$

In this case β_1 does not measure the effect of a change in X only. Instead

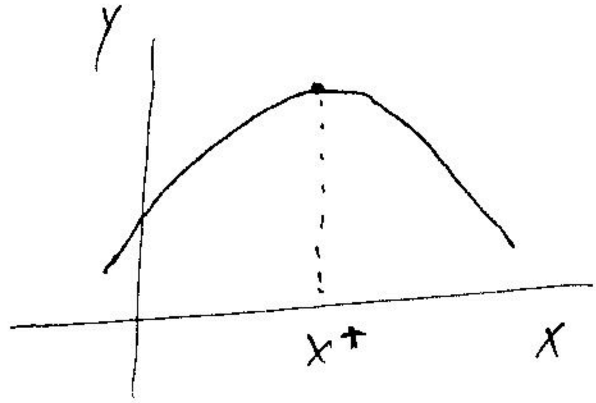
$$\Delta \hat{y} \approx (\hat{\beta}_1 + 2\hat{\beta}_2 X)\Delta X \, .$$

Therefore, the slope relationship between X and y depends on the value of X as well as $\hat{\beta}_1$ and $\hat{\beta}_2$.

Consider the single estimated quadratic regression model:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 X + \hat{\beta}_2 X^2 .$$

If $\hat{\beta}_2 < 0$ we have a quadratic function like

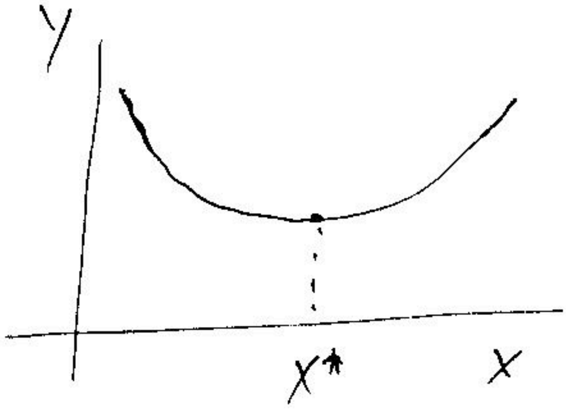


with the maximax value being realized at the X value of

$$X^* = \frac{-\hat{\beta}_1}{2\hat{\beta}_2} .$$

$$\left[\frac{\Delta y}{\Delta X} = \hat{\beta}_1 + 2\hat{\beta}_2 X^* = 0 \Rightarrow X^* = \frac{-\hat{\beta}_1}{2\hat{\beta}_2} \right]$$

If $\hat{\beta}_2 > 0$ the quadratic function looks like



Think of the former graph as y = profits and you are trying to determine the level of output x that maximizes profit.

In the latter graph think of y = average costs and you are trying to determine the level of output x that exhausts all of the economies of scale in production. For some good examples see Figures 6.1 and 6.2 and example 6.2 in Wooldridge.

Models with Interaction Terms

You should also note that in models with interaction terms the slope relationship between a variable involved in an interaction term is not straight forward. For example, consider

$$price = \beta_0 + \beta_1 sqrft + \beta_2 bdrms + \beta_3 sqrft \cdot bdrms + u$$

Here

$$\frac{\Delta price}{\Delta bdrms} = \beta_2 + \beta_3 sqrft.$$

Adjusted R-Squared

Motivation of Adjusted R^2

The usual R^2 can be written as

$$R^2 = 1 - \frac{SSR/N}{SST/N} \qquad (1)$$

The population R^2 can be written as

$$\rho^2 = 1 - \frac{\sigma_u^2}{\sigma_y^2}$$

Therefore, instead of ρ^2 using (1) and biased estimators of σ_u^2 and σ_y^2 we can use the unbiased estimators $\hat{\sigma}_u^2 = SSR/(N-k-1)$ and $\hat{\sigma}_y^2 = SST/(N-1)$ resulting in

$$\bar{R}^2 = 1 - \frac{\hat{\sigma}_u^2}{\hat{\sigma}_y^2} = 1 - \frac{SSR/(N-k-1)}{SST/(N-1)}$$

~~where $\hat{\sigma}_u^2$~~ $\bar{R}^2$ is the corrected or adjusted R^2. The shortcoming of R^2, at least as it relates to choosing variables to include in a regression is that R^2 is monotonically increasing in the number of explanatory variables that you have in your model. In fact if you have N observations, you can fit all points exactly with a multiple regression that is an $(N-1)$-order polynomial in X and $SSR = 0$ and $R^2 = 1$. But the resulting model will surely not be very interesting. Instead what we want is a model selection criterion that penalizes us for putting variables that don't increase R^2 (the fit) by very much. $\bar{R}^2$ penalizes us in such a way. Looking at the $\bar{R}^2$ criterion

$$\bar{R}^2 = 1 - \frac{SSR/(N-k-1)}{SST/(N-1)}$$

SST and $N-1$ do not change as k increases. But as k increases SSR falls but so does $(N-k-1)$. So $\bar{R}^2$ is not a monotonically increasing function of k. In contrast, as k increases we might expect, after some point, that adding another regressor will decrease $N-k-1$ more than SSR and $\bar{R}^2$ will actually decline.

So, if we use $\bar{R}^2$ as a model selection criterion we would prefer models with the highest adjusted R-squared value.

It should be noted, however, that the $\bar{R}^2$ model selection criterion is a pretty <u>liberal</u> selection criterion in that it can be shown that the adjusted R-squared criterion increases as long as we add an explanatory variable with a t-value whose absolute value is greater than one, that is

$$\bar{R}^2 \text{ increases when } |t_{\hat{\beta}_{k+1}}| > 1$$

but

$$\bar{R}^2 \text{ decreases when } |t_{\hat{\beta}_{k+1}}| < 1$$

Instead if one includes a new variable only if it is significant at the 5% level then with $N > 120$ we would require that $|t_{\hat{\beta}_{k+1}}| > 1.96$. Using $\bar{R}^2$ as a model selection criterion gives rise models having more explanatory variables than models based upon 5% statistical significance.

An other important limitation of using $\bar{R}^2$ to choose between models is that it cannot be used to choose between regression model with different forms of the dependent variable like y vs. log(y). We will see later how to deal with such comparisons.

Read the section in your textbook titled "Using Adjusted R-squared to Choose Between Non-Nested Models" (pp. 198-199).

Note two models are non-nested if neither equation is a special case of the other.