Reviewing Calculus

ECON 340: Economic Research Methods   Instructor: Div Bhagia

1 The Concept of a Derivative

Given some function,

$$y = f(x)$$

We are often interested in how the value for this function changes given a small change in x.

Note that for large changes in x, say x increases by Δx, y increases by Δy as follows:

$$\Delta y = f(x + \Delta x) - f(x)$$

Then average change in y per-unit change in x is given by:

$$\frac{\Delta y}{\Delta x} = \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

But as I said, we are interested in small changes in x. For this purpose we have the concept of derivatives. The derivative of y with respect to x, denoted by dy/dx, tells us how y changes with respect to a small change in x.

$$\frac{dy}{dx} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x + \Delta x) - f(x)}{\Delta x}$$

So essentially the derivative is the average change in y per unit change in x for very very small changes in x (i.e. when $\Delta x \to 0$ or Δx *goes to* 0.).

For most purposes, we don't need to deal with this clumsy definition of the derivative. We have some set of rules that we can follow to find derivatives. Memorizing these (by practicing) will make your life as a student of economics much easier.

2 How to Find a Derivative

Constant function rule

$$y = k \quad \rightarrow \quad \frac{dy}{dx} = 0$$

Example. For $y = 2$, $\frac{dy}{dx} = 0$.

Power function rule

$$y = x^n \quad \rightarrow \quad \frac{dy}{dx} = nx^{n-1}$$

Example. For $y = x^4$, $\frac{dy}{dx} = 4x^3$.

Generalized power function rule

$$y = cx^n \quad \rightarrow \quad \frac{dy}{dx} = cnx^{n-1}$$

Example. For $y = 3x^2$, $\frac{dy}{dx} = 6x$.

Sum-Difference Rule

$$\frac{d}{dx}[f(x) \pm g(x)] = \frac{df}{dx} \pm \frac{dg}{dx}$$

Example. For $y = 3x^2 + 5x + 2$, $\frac{dy}{dx} = 6x + 5 + 0 = 6x + 5$.

Chain Rule

If we have two functions:

$$z = f(y), y = g(x) \quad \rightarrow \quad \frac{dz}{dx} = \frac{dz}{dy} \cdot \frac{dy}{dx}$$

Example. For $z = y^3$ and $y = 1 + 2x$. Then,

$$\frac{dz}{dx} = 3y^2 \cdot 2x = 6x(1 + 2x)^2$$

Derivative of a Log Function

$$\frac{d}{dx}\log(x) = \frac{1}{x}$$

Note that, then by the chain rule:

$$\frac{d}{dx}\log(f(x)) = \frac{f'(x)}{f(x)}$$

3 Interpreting Coefficients in Linear Regression Model

Consider the following estimated model:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X$$

Now say X changes by ΔX. Then the change in the predicted value of Y:

$$\Delta\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X - [\hat{\beta}_0 + \hat{\beta}_1(X + \Delta X)] = \hat{\beta}_1 \Delta X$$

Since, $\Delta\hat{Y} = \hat{\beta}_1 \Delta X$, if X changes by 1 unit i.e., $\Delta X = 1$, the predicted value of Y changes by $\hat{\beta}_1$. Alternatively, we can write:

$$\frac{\Delta\hat{Y}}{\Delta X} = \hat{\beta}_1$$

Note that here,

$$\frac{\Delta\hat{Y}}{\Delta X} = \frac{d\hat{Y}}{dX}$$

This is because it's a linear function. Generally the above approach doesn't give us the correct correct answer for small changes in X and we have to rely on the derivative.

For the regression model with two variables:

$$\hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2$$

Say we change X_1 by ΔX_1 but keep X_2 constant, then change in $\hat{Y}$:

$$\Delta \hat{Y} = \hat{\beta}_0 + \hat{\beta}_1 X_1 + \hat{\beta}_2 X_2 - \hat{\beta}_0 - \hat{\beta}_1 (X_1 + \Delta X_1) - \hat{\beta}_2 X_2 = \hat{\beta}_1 \Delta X_1$$

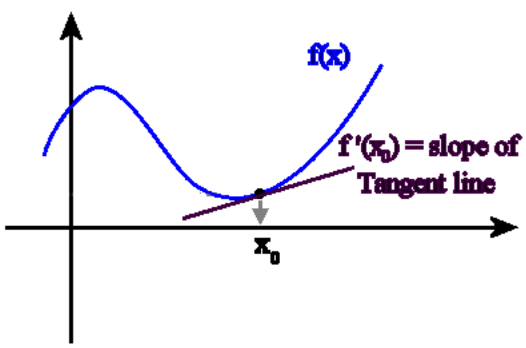
So now keeping X_2 constant, the predicted value of Y changes by $\hat{\beta}_1$ if X changes by 1 unit. Here again, we have

$$\frac{\Delta \hat{Y}}{\Delta X} |_{X_2 \text{ is constant}} = \frac{d\hat{Y}}{dX} |_{X_2 \text{ is constant}}$$

Again, this only works because the function is linear in parameters, but with quadratic or log-functions we will need to take the derivative.

4 Finding Maximum and Minimum of Functions

Note that the derivative is the slope of the tangent line at a point on the function.



- $dy/dx > 0$, the function is increasing
- $dy/dx < 0$, the function is decreasing
- $dy/dx = 0$, the function is constant

Note that at a peak or trough of function, the tangent has to be horizontal, that is $dy/dx = 0$. So when we try to find the maximum or minimum value of a function, we look at points where $dy/dx = 0$. In the derivation for the OLS estimator, we minimized the sum of squared residuals $\sum_i u_i^2$ as a function of OLS coefficients β_0 and β_1.