Intermediate Econometrics

Professor: Xiaojun Song

Timekeeper: Rui Zhou

Spring 2023

目录

5 MLR Analysis: Further Issues

5.1 More on Functional Form

5.1.1 Variables Included

Adding regressors may reduce the error variance, but may also exacerbate multicollinearity problems. Variables that are uncorrelated with other regressors should be added because they reduce error variance without increasing multicollinearity. However, such uncorrelated variables may be hard to find. If excessive numbers of variables are held fixed, this may betray your research purpose!

5.1.2 Log-Even Model

It's learned that logarithmic functional forms are more convenient for percentage and elasticity interpretation, and slope coefficients of logged variables are invariant to rescalings. Noteworthy that, taking logs often eliminates problems with outliers and helps to secure normality and homoskedasticity. However, variables measured in percentage points like education rate, or in units such as years should not be logged, and variables should be non-negative!

When $\log y$ is the dependent variable is helpful, this helps to predict y, say $\log y = \beta_0 + \beta_1 x_1 + ... + \beta_k x_k + u$. Equivalently, rewrite this as $y = \exp(\beta_0 + \beta_1 x_1 + ... + \beta_k x_k) \cdot \exp(u)$. We will show that it is hard to reverse the log-operation when constructing predictions.

Under the additional assumption that u is independent of $x_1, ..., x_k$:

$$E(y|x) = \exp(\beta_0 + \beta_1 x_1 + ... + \beta_k x_k) \cdot E(\exp(u))$$

Then from the PRF, the SRF would be

$$\hat{y} = \exp(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + ... + \hat{\beta}_k x_k) \cdot (\frac{1}{n} \sum_{i=1}^{n} \exp(\hat{u}_i))$$

From Jensen's inequality ,

$$E(e^u) \geq e^{E(u)} = 1$$

For $\frac{1}{n} \sum_{i=1}^{n} exp(\hat{u}_i)$, such a nonlinear function, it is hard to guarantee it to be an unbiased estimator of $\exp(u)$. However, for cases with large sample, consistency still holds, then plim $\hat{y} = y$.

5.1.3 Quadratic Form

Quadratic form fits better when dependent variable and independent variable(s) are nonlinearly-correlated. However, pay attention to the pinnacle of the quadratic function and your interpretation on both sides!

Sometimes, the "unwanted" relationship between DV and IV on the "unwanted" side does not make sense. In this case, be careful about how many observations in the sample lie in the "unwanted" side (before or after the turnaround point). If indeed that "unwanted" side's relationship does not make sense, and the proportion of sample lying there is non-negligible, then the regression model has to be refined!

5.1.4 Interaction Terms

Interaction term is a representation of mediation effect, and interaction effects make our interpretation of parameters more complicated.

Start from a concrete example, if we are to construct the following regression with interaction term

$$\begin{aligned} y &= \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2 + u \\ \Longrightarrow \quad \frac{\Delta y}{\Delta x_2} &= \beta_2 + \beta_3 x_1 \end{aligned}$$

Note that slope coefficients may no longer make sense! Theoretically, β_2 means the effect of x_2 on y, only when $x_1 = 0$; when $x_1 \neq 0$, β_2 does not fully cover the effect of x_2 alone; in some cases, $x_1 = 0$ is even impossible!

Hence, we do the reparametrization and center the data:

$$y = \alpha_0 + \delta_1 x_1 + \delta_2 x_2 + \beta_3 (x_1 - \mu_1)(x_2 - \mu_2) + u$$

Note that μ_1, μ_2 stand for population means. They may be replaced by sample means when conducting the regression. Here, δ_2 means the effect of x_2 if all variables take on their mean values, which is more reasonable. If necessary, interaction may be centered at other interesting values. Another advantage of reparametrization is that, it will make standard errors for *partial* effects at the mean (as in the example above) values available!

More generally speaking, in models with quadratics, interactions, and other nonlinear functional forms, the partial effect depends on the values of one or more explanatory variables. Average Partial Effect (APE) is a summary measure to describe the relationship between dependent variable and each explanatory variable. The computation is clear, after getting the partial effect and plugging in the estimated parameters, average partial effects for each unit across the sample. For quadratics or interactions, centering the explanatory variable will ensure the coefficient is APE itself.

5.2 Goodness of Fit

General remarks on R^2:

- A high R^2 does not imply that there is a causal interpretation.
- A low R^2 does not preclude the precise estimation of partial effects.

Recap what we have learned for R^2:

$$R^2 = 1 - \frac{SSR}{SST} = 1 - \frac{SSR/n}{SST/n} \to 1 - \frac{\sigma_u^2}{\sigma_y^2}$$

which is an estimate for $1 - \frac{\sigma_u^2}{\sigma_y^2}$.

A better estimate should take into account degrees of freedom. The adjusted R^2 is then

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

Clearly, the adjusted R-squared imposes a penalty for adding new regressors. The adjusted R-squared increases if and only if, the t statistic of a newly-added regressor is greater than 1 in absolute value.

The relationship between R^2 and adjusted R^2 is

$$\bar{R}^2 = 1 - \frac{n-1}{n-k-1} \cdot (1 - R^2)$$

Note that the adjusted R^2 may even get negative.

One advantageous application of adjusted R^2 is that, $\bar{R}^2$ can be used to compare and choose between nonnested models. In contrast, a comparison between the ordinary R^2 of both nonnested models would be unfair, because they may differ in the number of parameters, forms of parameters. However, neither ordinary R^2 or adjusted R^2 can be used to compare models which differ in their definition of the dependent variable.