Section Note 07: Prediction and Prediction Interval

JJ Chen*

October 2, 2014

Bugs might exist. Below are a summary of today's session:

1 BLP

1. The best linear predictor (BLP) of y^0 given $\boldsymbol{x}^0$ is

$$\hat{y}^0 = (\boldsymbol{x}^0)'\hat{\boldsymbol{\beta}},$$

where $\hat{\boldsymbol{\beta}}$ is the OLS estimator $\hat{\boldsymbol{\beta}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y}$.

2 BLP with CIs

2. The variance of BLP is

$$\mathrm{V}(\hat{y}^0) = \sigma^2(\boldsymbol{x}^0)'(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{x}^0),$$

and the estimated variance of BLP is

$$\hat{\mathrm{V}}(\hat{y}^0) = \hat{\sigma}^2(\boldsymbol{x}^0)'(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{x}^0).$$

*jchen215@uic.edu

The squared root of $\hat{\mathrm{V}}(\hat{y}^0)$ is the standard error of prediction (`stdp` in Stata) $se(\hat{y}^0)$, which can be used to construct confidence intervals for predicting of a "typical" observation $\mathrm{E}(y^0|\boldsymbol{x} = \boldsymbol{x}^0)$:

$$[\hat{y}^0 - t_{\alpha/2} * se(\hat{y}^0), \hat{y}^0 + t_{\alpha/2} * se(\hat{y}^0)].$$

See the Stata code for these type of prediction plots with CIs, and note that in general the interval is smallest at the mean values of $\boldsymbol{x}^0$.

3. The variance of prediction error (or forecast error in some context) using BLP is

$$\mathrm{V}(e^0) = \mathrm{V}(y^0) + \mathrm{V}(\hat{y}^0) = \sigma^2 + \sigma^2(\boldsymbol{x}^0)'(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{x}^0),$$

and the estimated variance of prediction error (or forecast error) is

$$\hat{\mathrm{V}}(e^0) = \hat{\mathrm{V}}(y^0) + \hat{\mathrm{V}}(\hat{y}^0) = \hat{\sigma}^2 + \hat{\sigma}^2(\boldsymbol{x}^0)'(\boldsymbol{X}'\boldsymbol{X})^{-1}(\boldsymbol{x}^0).$$

The squared root of $\hat{\mathrm{V}}(e^0)$ is the standard error of prediction error (`stdf` in Stata, which stands for standard error of the forecast) $se(\hat{e}^0)$. It can be used to construct confidence intervals for predicting of a "specific" observation y^0:

$$[\hat{y}^0 - t_{\alpha/2} * se(\hat{e}^0), \hat{y}^0 + t_{\alpha/2} * se(\hat{e}^0)].$$

Note that this confidence interval would be larger and one intuition is that we need to take into account of the uncertainty of "unobservables" associated with a "specific" observation, which is captured by the term $\mathrm{V}(y^0)$. See the Stata code for these type of prediction plots with CIs.