The multiple regression model (II)

Ragnar Nymoen

University of Oslo

14 February 2013

This lecture:

Based on the references and the model specification in Lecture 9:

- Statistical properties of estimators
- t-tests for the multivariate case

OLS estimates (expressions) I

For

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i \;\; i = 1, 2, \ldots, n \tag{1}$$

we have the following sample estimates

$$\hat{\beta}_1 = \frac{\hat{\sigma}^2_{X_2}\hat{\sigma}_{Y,X_1} - \hat{\sigma}_{Y,X_2}\hat{\sigma}_{X_1,X_2}}{\hat{\sigma}^2_{X_1}\hat{\sigma}^2_{X_2} - \hat{\sigma}^2_{X_1,X_2}} \tag{2}$$

$$\hat{\beta}_2 = \frac{\hat{\sigma}^2_{X_1}\hat{\sigma}_{Y,X_2} - \hat{\sigma}_{Y,X_1}\hat{\sigma}_{X_1,X_2}}{\hat{\sigma}^2_{X_1}\hat{\sigma}^2_{X_2} - \hat{\sigma}^2_{X_1,X_2}} \tag{3}$$

where $\hat{\sigma}^2_{X_j}(j = 1, 2)$, $\hat{\sigma}_{Y,X_j}$ $(j = 1, 2)$ and $\hat{\sigma}_{X_1,X_2}$ are empirical variances and covariances.

OLS estimates (expressions) II

Estimates for the two versions of the intercepts:

$$\hat{\beta}_0 = \bar{Y} + \hat{\beta}_1 \overline{X}_1 + \hat{\beta}_2 \overline{X}_2$$
$$\hat{\alpha} = \bar{Y}$$

Absence of perfect sample collinearity I

It is clear that (2) for $\hat{\beta}_1$ and (3) for $\hat{\beta}_2$ require

$$M := \hat{\sigma}^2_{X_1}\hat{\sigma}^2_{X_2} - \hat{\sigma}^2_{X_1,X_2} = \hat{\sigma}^2_{X_1}\hat{\sigma}^2_{X_2}(1 - r^2_{X_1X_2}) > 0$$

Cannot have perfect empirical correlation between the two regressors. Must have:

$$\hat{\sigma}^2_{X_1} > 0, \text{ and } \hat{\sigma}^2_{X_2} > 0 \text{ and } r^2_{X_1X_2} < 1 \iff -1 < r_{X_1X_2} < 1$$

- If any one of these conditions should fail, we have what the textbooks call **exact (or perfect) collinearity**.
- Absence of perfect collinearity is a requirement about the nature of the sample.

Absence of perfect sample collinearity II

- The case of $r_{X_1X_2} = 0$ also has a name. It is called **perfect orthogonality**. It does not create any problems in (2) or (3).
- In practice, the relevant case is $-1 < r_{X_1X_2} < 1$, i.e. a **degree of collinearity** (not perfect)

Expectation I

- Conditional on the values of X_1 and X_2, $\hat{\beta}_1$ is still a random variable because ε_i and Y_i are random variables.
- In that interpretation $\hat{\beta}_1$, $\hat{\beta}_2$, and $\hat{\beta}_0$ are **estimators** and we want to know their expectation, variance, and whether they are consistent or not.
- Start by considering $E(\hat{\beta}_1 \mid X_1, X_2)$, i.e., conditional on all the values of the two regressors.

Expectation II

- Write $\hat{\beta}_1$ as

$$\hat{\beta}_1 = \frac{\left(\hat{\sigma}^2_{X_2}\hat{\sigma}_{Y,X_1} - \hat{\sigma}_{Y,X_2}\hat{\sigma}_{X_1,X_2}\right)}{M}$$

then $E(\hat{\beta}_1 \mid X_1, X_2)$ becomes

$$E(\hat{\beta}_1 \mid X_1, X_2) = \frac{\hat{\sigma}^2_{X_2}}{M}E(\hat{\sigma}_{Y,X_1} \mid X_1, X_2) - \frac{\hat{\sigma}_{X_1,X_2}}{M}E(\hat{\sigma}_{Y,X_2} \mid X_1, X_2) \tag{4}$$

- Evaluate this in class, in order to show that

$$E(\hat{\beta}_j) = E\left[E(\hat{\beta}_j \mid X_1, X_2)\right] = \beta_j,\ j = 1, 2 \tag{5}$$

since $E(\varepsilon_i \mid X_1, X_2) = 0\ \forall\ i$ is generic for the regression model.

Variance I

Find that (under the classical assumptions of the model):

$$Var(\hat{\beta}_j \mid X_1, X_2) = \frac{\sigma^2}{n\hat{\sigma}^2_{X_j}\left[1 - r^2_{X_1,X_2}\right]}, j = 1, 2 \qquad (6)$$

and this also holds unconditionally.

- The BLUE property of the OLS estimators extends to the multivariate case (will no show)
- The variance (6) is low in samples that are informative about the "separate contributions" from X_1 and X_2:
  - $\hat{\sigma}^2_{X_j}$ high
  - $r^2_{X_1,X_2}$ low

Variance II

- $Var(\hat{\beta}_j)$ is lowest when $r^2_{X_1,X_2} = 0$, the regressors are orthogonal.
- Do not become tempted to say that "in order to estimate the marginal effect of X_2 on Y very precisely we should drop X_1 from the model". That will give a variance expression

$$\frac{\sigma'^2}{n\hat{\sigma}^2_{X_j}}$$

but $\sigma'^2 > \sigma^2$ in most cases!. (And there will be other problems as well).

Covariance I

In many applications weed to know $Cov(\hat{\beta}_1, \hat{\beta}_2)$.
It is easiest to find by starting from the second normal equation

$$\hat{\beta}_1 \hat{\sigma}^2_{X_1} + \hat{\beta}_2 \hat{\sigma}_{X_1,X_2} = \hat{\sigma}_{YX_1}$$

When we take (conditional) variance on both sides, we get

$$\hat{\sigma}^4_{X_1} Var(\hat{\beta}_1) + \hat{\sigma}^2_{X_1X_2} Var(\hat{\beta}_2) + 2\hat{\sigma}^2_{X_1} \hat{\sigma}_{X_1,X_2} Cov\left(\hat{\beta}_1, \hat{\beta}_2\right) = \frac{1}{n^2} Var(\hat{\sigma}_{YX_1})$$

The rhs we have from before:

$$n^{-2} Var(\hat{\sigma}_{YX_1}) = n^{-2} \frac{\sigma^2}{n} \hat{\sigma}^2_{X_1} = n^{-1} \sigma^2 \hat{\sigma}^2_{X_1}$$

Covariance II

Insertion of expressions for $Var(\hat{\beta}_1)$ and $Var(\hat{\beta}_2)$, solving for $Cov\left(\hat{\beta}_1, \hat{\beta}_2\right)$ gives

$$Cov\left(\hat{\beta}_1, \hat{\beta}_2\right) = -\frac{\sigma^2}{n}\frac{\hat{\sigma}_{X_1 X_2}}{M}$$

Algebra details in note on web-page.

Consistency of estimators I

Show for $\hat{\beta}_1$

$$\text{plim}\left(\hat{\beta}_1\right) = \text{plim}\left(\frac{\left(\hat{\sigma}^2_{X_2}\hat{\sigma}_{Y,X_1} - \hat{\sigma}_{Y,X_2}\hat{\sigma}_{X_1,X_2}\right)}{M}\right)$$

$$= \frac{\text{plim}\left(\hat{\sigma}^2_{X_2}\right)\text{plim}(\hat{\sigma}_{Y,X_1}) - \text{plim}(\hat{\sigma}_{Y,X_2})\,\text{plim}(\hat{\sigma}_{X_1,X_2})}{\text{plim}\, M}$$

Based on the assumptions of the regression model:

$$\text{plim}(\hat{\sigma}^2_{X_j}) = \sigma^2_{X_j}\ j = 1, 2$$

$$\text{plim}(\hat{\sigma}_{X_1,X_2}) = \sigma_{X_1 X_2}$$

$$\text{plim}\, M = \sigma^2_{X_1}\sigma^2_{X_2} - \sigma^2_{X_1,X_2}$$

Consistency of estimators II

$$\text{plim}(\hat{\sigma}_{Y,X_1}) = \beta_1\sigma^2_{X_1} + \beta_2\sigma_{X_1X_2} + \text{plim}\left[\frac{1}{n}\sum_{i=1}^{n}\varepsilon_i(X_{1i} - \bar{X}_1)\right]$$
$$= \beta_1\sigma^2_{X_1} + \beta_2\sigma_{X_1X_2}$$
$$\text{plim}(\hat{\sigma}_{Y,X_2}) = \beta_1\sigma_{X_1X_2} + \beta_2\sigma^2_{X_2}$$

$$\text{plim}\left(\hat{\beta}_1\right) = \frac{\sigma^2_{X_2}\left[\beta_1\sigma^2_{X_1} + \beta_2\sigma_{X_1X_2}\right] - \left[\beta_1\sigma_{X_1X_2} + \beta_2\sigma^2_{X_2}\right]\sigma_{X_1,X_2}}{\sigma^2_{X_1}\sigma^2_{X_2} - \sigma^2_{X_1,X_2}}$$
$$= \frac{\beta_1(\sigma^2_{X_2}\sigma^2_{X_1} - \sigma^2_{X_1X_2}) + \beta_2\sigma^2_{X_2}\sigma_{X_1X_2} - \beta_2\sigma^2_{X_2}\sigma_{X_1,X_2}}{\sigma^2_{X_1}\sigma^2_{X_2} - \sigma^2_{X_1,X_2}}$$
$$= \beta_1$$

The OLS estimators $\hat{\beta}_0$ and $\hat{\beta}_2$ are also consistent

Estimated standard errors and t-values I

- Just like in simple regression we need to replace $\sqrt{Var(\hat{\beta}_j)}$ from (6) by

$$\widehat{se}(\hat{\beta}_j) = \sqrt{\frac{\hat{\sigma}^2}{n\hat{\sigma}^2_{X_j}\left[1 - r^2_{X_1,X_2}\right]}}$$

where $\hat{\sigma}^2$ is an estimator.

- In the same way as in simple regression we make the **normality assumption** about the disturbances.

Estimated standard errors and t-values II

- Also, by the same logic as before we choose the unbiased estimator

$$\hat{\sigma}^2 = \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^2}{n-3}. \tag{7}$$

  where $\hat{\varepsilon}_i$ are the OLS residuals from the bivariate regression model.

- Note $n-3$ instead of $n-2$ since we have now 3 exact relationships between the n residuals.
- Again, in direct parallel to single regressor model we now have

$$T = \frac{\hat{\beta}_j - E(\hat{\beta}_j)}{\widehat{se}(\hat{\beta}_j)} \sim t(n-3),\ j = 1, 2. \tag{8}$$

Estimated standard errors and t-values III

- which is used in hypotheses testing an in the different forms of interval estimation.
- Some examples of null hypotheses that can be tested with t-tests:
- H_0: $\beta_1 = \beta_1^0$
- H_0: $\beta_2 = \beta_2^0$
- H_0: $\beta_1 + \beta_2 = a^0$

In class: Back to Andy's