EC708 Discussion 13
Trinity

Yan Liu

Department of Economics
Boston University

April 22, 2022

Outline

1. Trinity of Tests in Classical Linear Model
2. LM Test for Model Misspecification
   - Testing for Heteroskedasticity: Breush-Pagan LM
   - LM Test for $AR(p)$ Errors

Classical Linear Model

Consider the linear model

$$y = X\beta + u, \quad u|X \sim N(0, \sigma^2 I)$$

where $\dim(\beta) = k$. We are interested in testing linear restrictions

$$R\beta - r = 0,$$

where R is $q \times k$ and r is $q \times 1$.

Wald Test in Classical Linear Model

- Let $(\hat{\beta}_T, \hat{\sigma}_T^2)$ be the unrestricted MLE of (β, σ^2). Then,

$$\hat{\beta}_T = (X'X)^{-1}X'y \quad \text{and} \quad \hat{\sigma}_T^2 = \frac{1}{T}(y - X\hat{\beta}_T)'(y - X\hat{\beta}_T).$$

  We can show $\sqrt{T}(\hat{\beta}_T - \beta) \xrightarrow{d} N(0, \sigma^2 E[X_t X_t']^{-1})$ and $\hat{\sigma}_T^2 \xrightarrow{p} \sigma^2$.

- Under $H_0 : R\beta - r = 0$, $\sqrt{T}(R\hat{\beta}_T - r) \xrightarrow{d} N(0, \sigma^2 RE[X_t X_t']^{-1}R')$. Then, the Wald statistic is

$$W_T = (R\hat{\beta}_T - r)'[\hat{\sigma}_T^2 R(X'X)^{-1}R']^{-1}(R\hat{\beta}_T - r) \xrightarrow{d} \chi_q^2 \quad \text{under } H_0.$$

LM Test in Classical Linear Model

- Let $(\tilde{\beta}_T, \tilde{\sigma}_T^2)$ be the restricted MLE of (β, σ^2) and $\tilde{\lambda}_T$ be the vector of associated Lagrange multipliers.
- The FOCs for $\tilde{\beta}_T$ and $\hat{\beta}_T$ are respectively

$$\frac{1}{\tilde{\sigma}_T^2} X'(y - X\tilde{\beta}_T) - R'\tilde{\lambda}_T = 0, \tag{1}$$

$$\frac{1}{\hat{\sigma}_T^2} X'(y - X\hat{\beta}_T) = 0. \tag{2}$$

Plugging (2) into the RHS of (1),

$$\begin{aligned} & \frac{1}{\tilde{\sigma}_T^2} X'X(\hat{\beta}_T - \tilde{\beta}_T) = R'\tilde{\lambda}_T \\ \Rightarrow\ & \frac{1}{\tilde{\sigma}_T^2} R(\hat{\beta}_T - \tilde{\beta}_T) = R(X'X)^{-1}R'\tilde{\lambda}_T \\ \Rightarrow\ & \tilde{\lambda}_T = \frac{1}{\tilde{\sigma}_T^2}[R(X'X)^{-1}R']^{-1}(R\hat{\beta}_T - r). \end{aligned}$$

LM Test in Classical Linear Model

- Plugging the expression for $\tilde{\lambda}_T$ into (1), we have

$$\tilde{\beta}_T = \hat{\beta}_T - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(R\hat{\beta}_T - r) \xrightarrow{p} \beta$$

  under H_0 and thus $\tilde{\sigma}_T^2 = \frac{1}{T}(y - X\tilde{\beta}_T)'(y - X\tilde{\beta}_T) \xrightarrow{p} \sigma^2$.

- Hence, $T^{-1/2}\tilde{\lambda}_T \xrightarrow{d} N(0, (\sigma^2 RE[X_t X_t']^{-1}R')^{-1})$.
- Then, the LM statistic is

$$LM = (R\hat{\beta}_T - r)'[\tilde{\sigma}_T^2 R(X'X)^{-1}R']^{-1}(R\hat{\beta}_T - r) \xrightarrow{d} \chi_q^2 \quad \text{under } H_0.$$

Remark:

- Recall Wald test:

$$W_T = (R\hat{\beta}_T - r)'[\hat{\sigma}_T^2 R(X'X)^{-1}R']^{-1}(R\hat{\beta}_T - r).$$

  Differ in terms of estimator of σ^2.

LR Test in Classical Linear Model

- The log-likelihood function evaluated at $(\hat{\beta}_T, \hat{\sigma}_T^2)$ is

$$\begin{aligned}\bar{\ell}_T(\hat{\beta}_T, \hat{\sigma}_T^2) &= -\frac{T}{2}\ln 2\pi - \frac{T}{2}\ln \hat{\sigma}_T^2 - \frac{1}{2\hat{\sigma}_T^2}(y - X\hat{\beta}_T)'(y - X\hat{\beta}_T)\\ &= \text{cons} - \frac{T}{2}\ln \hat{\sigma}_T^2.\end{aligned}$$

- Similarly, the log-likelihood function evaluated at $(\tilde{\beta}_T, \tilde{\sigma}_T^2)$ is

$$\bar{\ell}_T(\tilde{\beta}_T, \tilde{\sigma}_T^2) = \text{cons} - \frac{T}{2}\ln \tilde{\sigma}_T^2.$$

- Then, the LR statistic is

$$LR = 2(\bar{\ell}_T(\hat{\beta}_T, \hat{\sigma}_T^2) - \bar{\ell}_T(\tilde{\beta}_T, \tilde{\sigma}_T^2)) = T\ln\left(\frac{\tilde{\sigma}_T^2}{\hat{\sigma}_T^2}\right) \xrightarrow{d} \chi_q^2 \quad \text{under } H_0.$$

Inequality $LM \leq LR \leq W$ in Classical Linear Model

For simplicity, normalize $R\beta = 0$. We start with two lemmas.

Lemma 1

Define $\tilde{u} = y - X\tilde{\beta}_T$ *and* $\hat{u} = y - X\hat{\beta}_T$*, then*

$$W_T = \frac{\tilde{u}'\tilde{u} - \hat{u}'\hat{u}}{\hat{\sigma}_T^2}, \quad LM_T = \frac{\tilde{u}'\tilde{u} - \hat{u}'\hat{u}}{\tilde{\sigma}_T^2}.$$

Proof: Recall that $\tilde{\beta}_T = \hat{\beta}_T - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(R\hat{\beta}_T - r)$, $X'\hat{u} = 0$, and $r = 0$. Then,

$$\begin{aligned}
\tilde{u}'\tilde{u} - \hat{u}'\hat{u} &= [\hat{u} - X(\tilde{\beta}_T - \hat{\beta}_T)]'[\hat{u} - X(\tilde{\beta}_T - \hat{\beta}_T)] - \hat{u}'\hat{u} \\
&= (\tilde{\beta}_T - \hat{\beta}_T)'X'X(\tilde{\beta}_T - \hat{\beta}_T) \\
&= \hat{\beta}_T'R'[R(X'X)^{-1}R']^{-1}R\hat{\beta}_T.
\end{aligned}$$

Inequality $LM \leq LR \leq W$ in Classical Linear Model

Lemma 2

The Wald, LR, and LM satisfy the following relations:

- $W_T = 2[\bar{\ell}_T(\hat{\beta}_T, \hat{\sigma}_T^2) - \bar{\ell}_T(\tilde{\beta}_T, \hat{\sigma}_T^2)]$;
- $LM_T = 2[\bar{\ell}_T(\hat{\beta}_T, \tilde{\sigma}_T^2) - \bar{\ell}_T(\tilde{\beta}_T, \tilde{\sigma}_T^2)]$;
- $LR_T = 2[\bar{\ell}_T(\hat{\beta}_T, \hat{\sigma}_T^2) - \bar{\ell}_T(\tilde{\beta}_T, \tilde{\sigma}_T^2)]$.

Proof: For any σ^2,

$$\bar{\ell}_T(\hat{\beta}_T, \sigma^2) = -\frac{T}{2}\ln\sigma^2 - \frac{T}{2}\ln 2\pi - \frac{1}{2}\frac{\hat{u}'\hat{u}}{\sigma^2},$$
$$\bar{\ell}_T(\tilde{\beta}_T, \sigma^2) = -\frac{T}{2}\ln\sigma^2 - \frac{T}{2}\ln 2\pi - \frac{1}{2}\frac{\tilde{u}'\tilde{u}}{\sigma^2}.$$

Plug in the expressions of Wald and LM statistics in Lemma 1.

Inequality $LM \leq LR \leq W$ in Classical Linear Model

Theorem 1

$LM_T \leq LR_T \leq W_T$ for any T.

Proof:

- $LR_T \geq LM_T$ if

$$\bar{\ell}_T(\hat{\beta}_T, \hat{\sigma}_T^2) \geq \bar{\ell}_T(\hat{\beta}_T, \tilde{\sigma}_T^2).$$

- $W_T \geq LR_T$ if

$$\bar{\ell}_T(\tilde{\beta}_T, \hat{\sigma}_T^2) \leq \bar{\ell}_T(\tilde{\beta}_T, \tilde{\sigma}_T^2).$$

These inequalities hold by the definition of $\hat{\sigma}_T^2$ and $\tilde{\sigma}_T^2$.

Remark: Theorem 1 applies to the GLS model:

$$y|X \sim N(X\beta, \sigma\Omega), \quad \Omega = \Omega(\omega),$$

where ω is a finite estimable parameter vector.

When Are LM, LR, and Wald the Same?

- Suppose σ^2 is known. Then by Lemma 2, the three are identical.
- More generally, if the log-likelihood has the following form

$$\bar{\ell}_T(\theta) = b - \frac{1}{2}(\theta - \hat{\theta}_T)' A (\theta - \hat{\theta}_T)$$

  where A is a symmetric positive definite matrix and $\hat{\theta}_T$ is a function of data, then $LM_T = LR_T = W_T$.
- $LM_T = LR_T = W_T = (\theta_0 - \hat{\theta}_T)' A (\theta_0 - \hat{\theta}_T)$, where θ_0 is subject to the constraint of H_0.

Remark: In general, whenever the true value of θ is close to θ_0, the log-likelihood function in the neighborhood of θ_0 is approximately quadratic for large T, hence asymptotic equivalence of the three tests.

Outline

1. Trinity of Tests in Classical Linear Model
2. LM Test for Model Misspecification
   - Testing for Heteroskedasticity: Breush-Pagan LM
   - LM Test for $AR(p)$ Errors

LM Test as A Diagnostic

- We can use hypothesis testing for specification search: null is a specification in favor and alternative is a more general specification.
- Test for this purpose is **diagnostic**: check if data are well represented by the specification.
- LM test is based on parameter fit under the null, usually expressed as residuals from the estimates under the null.
- Each alternative considered indicates a particular type of non-randomness.

Outline

1. Trinity of Tests in Classical Linear Model

2. LM Test for Model Misspecification
   - Testing for Heteroskedasticity: Breush-Pagan LM
   - LM Test for $AR(p)$ Errors

Testing for Heteroskedasticity: Breush-Pagan LM

Consider the following model:

$$y_t = x_t'\beta + u_t, \quad u_t|x_t, z_t \sim N(0, \sigma_t^2), \quad t = 1, \ldots, T.$$

Null and alternative hypothesis:

$$H_0 : \sigma_t^2 = \sigma^2 \; \forall t, \quad H_1 : \sigma_t^2 = h(z_t'\alpha),$$

where $z_t = (1, z_{1t}, \ldots, z_{qt})' \in \mathbb{R}^{q+1}$ and $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_q)'$. The null can be rewritten as

$$H_0 : \alpha_1 = \cdots = \alpha_q = 0.$$

- z_t can be a vector function of x_t, e.g. $z_t'\alpha = \alpha_0 + \alpha_1 x_t'\beta$.
- Under both the null and alternative, assume no serial correlation: $E[u_t u_s] = 0$ for $t \neq s$.

Breush-Pagan Test Statistic

Log-likelihood function:

$$\bar{\ell}_T(\beta, \alpha) = -\frac{T}{2}\ln(2\pi) - \frac{1}{2}\sum_{t=1}^{T}\ln(h(z_t'\alpha)) - \frac{1}{2}\sum_{t=1}^{T}\frac{(y_t - x_t'\beta)^2}{h(z_t'\alpha)}.$$

Note that

- FOC for $\tilde{\beta}$ implies $\frac{\partial \bar{\ell}_T}{\partial \beta}\Big|_{\alpha=\tilde{\alpha}, \beta=\tilde{\beta}} = 0$;
- Information matrix is block diagonal: $E[\frac{\partial^2 \ell_t(\beta,\alpha)}{\partial \beta \partial \alpha'}] = 0$.

We can calculate that the score is

$$\frac{\partial \ell_t(\beta, \alpha)}{\partial \alpha} = \frac{1}{2} z_t \Big(\frac{u_t^2}{h(z_t'\alpha)} - 1\Big)\frac{h'(z_t'\alpha)}{h(z_t'\alpha)}$$

and the information matrix is

$$I(\beta, \alpha) = -E\Big[\frac{\partial^2 \ell_t(\beta, \alpha)}{\partial \alpha \partial \alpha'}\Big] = \frac{1}{2}E\Big[z_t z_t' \frac{h'(z_t'\alpha)}{h(z_t'\alpha)}\Big].$$

Breush-Pagan Test Statistic

Evaluated under H_0, LM test statistic simplifies to ($\frac{h'(\tilde{\alpha}_0)}{h(\tilde{\alpha}_0)}$) is cancelled out)

$$LM_T = \frac{1}{2}\Big[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_t\Big(\frac{\tilde{u}_t^2}{h(\tilde{\alpha}_0)} - 1\Big)\Big]'\Big(\frac{1}{T}\sum_{t=1}^{T} z_t z_t'\Big)^{-1}\Big[\frac{1}{\sqrt{T}}\sum_{t=1}^{T} z_t\Big(\frac{\tilde{u}_t^2}{h(\tilde{\alpha}_0)} - 1\Big)\Big].$$

Let $f = (f_1, \dots, f_T)'$ with $f_t = \frac{\tilde{u}_t^2}{\tilde{\sigma}^2} - 1$, where $\tilde{u}_t$ and $\tilde{\sigma}^2$ are residuals and variance estimates under H_0. Then,

$$LM_T = \frac{1}{2} f' Z(Z'Z)^{-1}Z'f = \frac{1}{2} f' P_Z f.$$

Breush-Pagan Test Statistic: Procedures

1. Apply OLS to $y = X\beta + u$ and obtain residuals $\hat{u}$.
2. Compute $f_t = \frac{\hat{u}_t^2}{\hat{\sigma}^2} - 1$ where $\hat{\sigma}^2 = \frac{1}{T}\sum_{t=1}^{T} \hat{u}_t^2$.
3. Run OLS of f on Z and compute the LM statistic

$$LM_T = \frac{1}{2} f' P_Z f = \frac{1}{2} ESS.$$

Under the null, $LM_T \xrightarrow{d} \chi_q^2$.

Remark:

- Since $\text{plim}_{T\to\infty} f'f/T = 2$ under H_0 and H_1, an asymptotically equivalent test statistic is TR^2 from regressing f on Z.
- As long as Z has an intercept, the statistic can be computed by regressing $\hat{u}$ on Z and calculating TR^2.

Outline

1. Trinity of Tests in Classical Linear Model
2. LM Test for Model Misspecification
   - Testing for Heteroskedasticity: Breush-Pagan LM
   - LM Test for $AR(p)$ Errors

LM Test for $AR(p)$ Errors

Consider the model with k regressors in X and p lags in error:

$$
\begin{aligned}
y &= X\beta + u \\
u_t &= \psi_1 u_{t-1} + \cdots + \psi_p u_{t-p} + \varepsilon_t \\
\varepsilon_t &\sim \text{i.i.d. } N(0, \sigma^2)
\end{aligned}
$$

- The null is $H_0 : \psi = (\psi_1, \ldots, \psi_p) = \mathbf{0}$.
- Ignore the first p observations and rewrite the model as

$$
y_t = x_t'\beta + \sum_{j=1}^{p} \psi_j (y_{t-j} - x_{t-j}'\beta) + \varepsilon_t.
$$

LM Test for $AR(p)$ Errors

- Denote $\theta = (\beta, \psi)$. The log-likelihood after concentrating out σ^2 is

$$\bar{\ell}_T(\theta) = \text{cons} - \frac{T}{2} \ln \frac{1}{T} \varepsilon' \varepsilon = \text{cons} - \frac{T}{2} \ln \frac{1}{T} \sum_{t=1}^{T} \Big(u_t - \sum_{j=1}^{p} \psi_j u_{t-j} \Big)^2.$$

- Evaluated at the restricted MLE $\tilde{\theta}$, the score and estimated information matrix are

$$\frac{\partial \bar{\ell}_T(\tilde{\theta})}{\partial \theta} = \frac{F' \tilde{u}}{\tilde{\sigma}^2}, \quad \hat{I}_T(\tilde{\theta}) = \frac{1}{T} \frac{F'F}{\tilde{\sigma}^2},$$

where $F = (X, U)$ with U having rows $U_t = (\tilde{u}_{t-1}, \ldots, \tilde{u}_{t-p})$ and $\tilde{\sigma}^2 = \frac{1}{T}(\tilde{u} - F\tilde{\theta})'(\tilde{u} - F\tilde{\theta})$.

LM Test for $AR(p)$ Errors

- Construct the LM statistic

$$LM_T = \frac{1}{T}\frac{\partial \bar{\ell}_T(\tilde{\theta})}{\partial \theta'}\hat{I}_T(\tilde{\theta})^{-1}\frac{\partial \bar{\ell}_T(\tilde{\theta})}{\partial \theta} = \frac{1}{\tilde{\sigma}^2}\tilde{u}'F(F'F)^{-1}F'\tilde{u}.$$

- $LM_T = TR^2$ where R^2 is calculated from regressing $\tilde{u}$ on F.

LM Test for $AR(p)$ Errors: Procedures

1. Run OLS on y against X and get $\hat{u}$.
2. Run OLS on the auxiliary regression

$$\hat{u} = X\tau + U\delta + v.$$

3. Compute R^2 from the auxiliary regression and construct LM statistic

$$LM_T = TR^2.$$

Under the null, $LM_T \xrightarrow{d} \chi_p^2$.

Remarks:

- If X includes no lagged dependent variables, then $\text{plim}_{T\to\infty} \frac{X'U}{T} = 0$ and the auxiliary regression will be unaffected by leaving out the X's.
- For $p = 1$, this test is asymptotically equivalent to the Durbin-Watson statistic.

Testing for $AR(1)$ Errors: Durbin-Watson

- Consider the model with k regressors

$$y = X\beta + u$$

$$u_t = \rho u_{t-1} + \varepsilon_t$$

$$\varepsilon_t \sim \text{i.i.d.} N(0, \sigma^2)$$

- Consider the following test statistic

$$d = \frac{\sum_{t=2}^{T} (\hat{u}_t - \hat{u}_{t-1})^2}{\sum_{t=1}^{T} \hat{u}_t^2},$$

where $\hat{u}_t$ are OLS residuals.

- Why this statistic? Note that

$$d \approx \frac{2\sum_{t=2}^{T} (\hat{u}_t^2 - \hat{u}_t \hat{u}_{t-1})}{\sum_{t=2}^{T} \hat{u}_{t-1}^2} = 2 - 2\frac{\sum_{t=2}^{T} \hat{u}_t \hat{u}_{t-1}}{\sum_{t=2}^{T} \hat{u}_{t-1}^2} = 2(1 - \hat{\rho}).$$

Testing for $AR(1)$ Errors: Durbin-Watson

- Exact distribution of d depends on X. Can bound this dependence as a function of k and T. Hence critical values depend on k, T, and size α.
- Adding lagged dependent variables in X biases $\hat{\rho}$ downward.
- Durbin's h test provides a correction for the first order case.
  - Denote $\hat{\alpha}_1$ as OLS coefficient on y_{t-1}. Consider the following statistic

    $$h = \hat{\rho}\sqrt{\frac{T}{1 - T\hat{V}(\hat{\alpha}_1)}} \approx \big(1 - \frac{2}{d}\big)\sqrt{\frac{T}{1 - T\hat{V}(\hat{\alpha}_1)}},$$

    where $\hat{V}(\hat{\alpha}_1)$ is the variance estimator of $\hat{\alpha}_1$.
  - **Caveat:** It may happen that $T\hat{V}(\hat{\alpha}_1) > 1$ because of sampling fluctuation. In this case the test statistic undefined.