Lecture 2: Restrictions. Stability

Econometric Methods

Andrzej Torój

SGH Warsaw School of Economics – Institute of Econometrics

Outline

1. Restrictions in models: testing and estimation
   - Restrictions
   - Functional form
2. Testing model stability
   - Chow test
   - CUSUM test
3. Multicollinearity

Outline

1. Restrictions in models: testing and estimation
2. Testing model stability
3. Multicollinearity

Example: constant returns to scale (1)

Constant returns to scale in Cobb-Douglas type production function

$$Y_t = \alpha_0 L_t^{\alpha_1} C_t^{\alpha_2} \varepsilon_t$$

$$ln(Y_t) = ln(\alpha_0) + \alpha_1 ln(L_t) + \alpha_2 ln(C_t) + ln(\varepsilon_t)$$

subject to:

$$\alpha_1 + \alpha_2 = 1$$

that is:

$$\alpha_2 = 1 - \alpha_1$$

Restricted Least Squares (1)

OLS objective:

$$S=\sum_{t=1}^{T}\varepsilon_t^2=\sum_{t=1}^{T}\left(y_t-\beta_0-\beta_1x_{1,t}-\beta_2x_{2,t}-\ldots-\beta_kx_{k,t}\right)\rightarrow\min_{\beta_0,\beta_1,\ldots}$$

RLS objective:

$$S=\sum_{t=1}^{T}\varepsilon_t^2=\sum_{t=1}^{T}\left(y_t-\beta_0-\beta_1x_{1,t}-\beta_2x_{2,t}-\ldots-\beta_kx_{k,t}\right)\rightarrow\min_{\beta_0,\beta_1,\ldots}\ \text{s.t. } f\left(\beta_o,\beta_1,\ldots\right)=0$$

If restrictions are linear, then:

$$\underbrace{\begin{array}{ccccccccc}
r_{11}\beta_0 & + & r_{12}\beta_1 & + & \ldots & + & r_{1k}\beta_k & = & q_1\\
r_{21}\beta_0 & + & r_{22}\beta_1 & + & \ldots & + & r_{2k}\beta_k & = & q_2\\
\vdots & & & & & & & & \\
r_{m1}\beta_0 & + & r_{m2}\beta_1 & + & \ldots & + & r_{mk}\beta_k & = & q_m
\end{array}}_{\mathbf{R}\boldsymbol{\beta}=\mathbf{q}}$$

Restricted Least Squares (2)

In such a case, the solution to the constrained minimization problem leads to the following modification of OLS estimator, referred to as RLS:

$$\hat{\boldsymbol{\beta}}^{*} = \hat{\boldsymbol{\beta}} - \left(\mathbf{X}^{T}\mathbf{X}\right)^{-1}\mathbf{R}^{T}\left[\mathbf{R}\left(\mathbf{X}^{T}\mathbf{X}\right)^{-1}\mathbf{R}^{T}\right]^{-1}\left(\mathbf{R}\hat{\boldsymbol{\beta}} - \mathbf{q}\right)$$

Let us define **residuals' sum of squares (RSS)** in the **restricted (R)** and **unrestricted (U)** model, respectively:

$$RRSS = \sum_{t=1}^{T}\hat{\varepsilon}_t^{*2} = \sum_{t=1}^{T}\left(y_t - \mathbf{x_t}\hat{\boldsymbol{\beta}}^{*}\right)^2$$

$$URSS = \sum_{t=1}^{T}\hat{\varepsilon}_t^{2} = \sum_{t=1}^{T}\left(y_i - \mathbf{x_t}\hat{\boldsymbol{\beta}}\right)^2$$

Example: constant returns to scale (2)

Estimation results under restriction (function **ORGLS**):

```
Coefficients:
(Intercept)           l_L          l_C
      1.069         0.637        0.363

Unconstrained solution:
(Intercept)           l_L          l_C
     1.1706        0.6030       0.3757
```

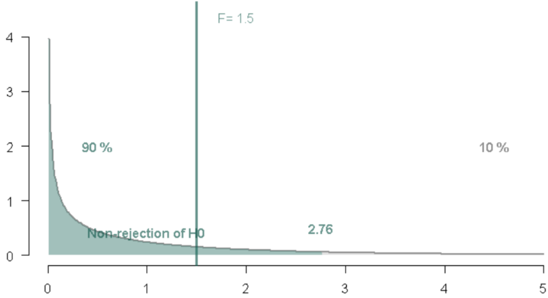
Wald test of restrictions' validity (1)

Wald Test

$H_0 : \mathbf{R}\beta = \mathbf{q}$, i.e. all imposed restrictions are true
$H_1 : \mathbf{R}\beta \neq \mathbf{q}$, i.e. at least one imposed restriction is false
Test statistic: $F = \frac{(RRSS-URSS)/m}{URSS/(T-k)}$ distributed as $F(m, T-k)$.

Wald test of restrictions' validity (2)

Example graphical interpretation (10% significance level implies the critical value of 2.76; if the computed value of test statistic is 1.5, this means a non-rejection of H0):



Example: constant returns to scale (3)

```
#Testing constraints and RLS regression
install.packages("car")
library(car)
linearHypothesis(f_produkcji,"l_L+l_C=1")
Linear hypothesis test

Hypothesis:
l_L  + l_C = 1

Model 1: restricted model
Model 2: l_Y ~ l_L + l_C

  Res.Df     RSS Df Sum of Sq      F Pr(>F)
1     25 0.85574
2     24 0.85163  1 0.0041075 0.1158 0.7366
```

High *p-value* – non-rejection of the null hypothesis that the restrictions are valid.

Excluded and included variables testing

Test of excluded variables

$H_0: \beta_{k-h}, \dots \beta_k = 0$: variables $x_{k-h}, \dots x_k$ are correctly excluded from the model
H_1 : these variables should be included

Test of included variables

$H_0: \beta_{k-h}, \dots \beta_k = 0$: variables $x_{k-h}, \dots x_k$ are incorrectly included in the model
H_1 : these variables are correctly included

Examples

- Cobb-Douglas production function:
  $ln\left(Y_t\right) = ln\left(\alpha_0\right) + \alpha_1 ln\left(L_t\right) + \alpha_2 ln\left(C_t\right) + ln\left(\varepsilon_t\right)$
- Alternatives from the literature: translogarithmic form + inclusion of trend:
  $ln\left(Y_t\right) = ln\left(\alpha_0\right) + \alpha_1 ln\left(L_t\right) + \alpha_2 ln\left(C_t\right) + \alpha_3 t + \alpha_4 ln\left(L_t\right) ln\left(C_t\right)$
- Was it correct to omit $ln\left(L_t\right) ln\left(C_t\right)$ and t (i.e. $H_0 : \alpha_3 = \alpha_4 = 0$)?

RESET test

RESET test

H_0 : no omitted variables, correct functional form, correct dynamic specification; H_1 : incorrect specification

Test regression: powers of theoretical values included in the model equation as regressors

$$\text{e.g.}\quad \begin{array}{l} 1 \rightarrow \hat{y}^2 \\ 2 \rightarrow \hat{y}^2\ \hat{y}^3 \\ 3 \rightarrow \hat{y}^2\ \hat{y}^3\ \hat{y}^4 \end{array}$$

```
#Test RESET
install.packages("lmtest")
library(lmtest)
resettest(f_produkcji, power = 2:3, data=f_produkcji_dane)
```

Outline

Chow test

Special case of the Wald test:

- split the sample into 2 subsamples;
  - baseline model treated as **restricted**: in assumes invariant parameters in both subsamples ($k+1$ constraints – as many as parameters);
  - in the unrestricted model, parameters can change between subsamples (in fact, 2 models estimated for both subsamples).

Chow test

$H_0 : \beta_A = \beta_B$, i.e. parameters equal in both subsamples (A and B)
$H_1 : \beta_A \neq \beta_B$, i.e. parameters different in both subsamples
Test statistic: $F = \frac{(RRSS-URSS)/(k+1)}{URSS/[T-2(k+1)]}$ distributed as $F\left[k+1, T-2\left(k+1\right)\right]$.

Chow test

Special case of the Wald test:

- split the sample into 2 subsamples;
    - baseline model treated as **restricted**: in assumes invariant parameters in both subsamples ($k + 1$ constraints – as many as parameters);
    - in the unrestricted model, parameters can change between subsamples (in fact, 2 models estimated for both subsamples).

Chow test

$H_0 : \beta_A = \beta_B$, i.e. parameters equal in both subsamples (A and B)
$H_1 : \beta_A \neq \beta_B$, i.e. parameters different in both subsamples
Test statistic: $F = \frac{(RRSS - URSS)/(k+1)}{URSS/[T-2(k+1)]}$ distributed as $F\left[k+1, T-2\left(k+1\right)\right]$.

Chow test

Special case of the Wald test:

- split the sample into 2 subsamples;
  - baseline model treated as **restricted**: in assumes invariant parameters in both subsamples ($k+1$ constraints – as many as parameters);
  - in the unrestricted model, parameters can change between subsamples (in fact, 2 models estimated for both subsamples).

Chow test

$H_0 : \beta_{\boldsymbol{A}} = \beta_{\boldsymbol{B}}$, i.e. parameters equal in both subsamples (A and B)
$H_1 : \beta_{\boldsymbol{A}} \neq \beta_{\boldsymbol{B}}$, i.e. parameters different in both subsamples
Test statistic: $F = \frac{(RRSS-URSS)/(k+1)}{URSS/[T-2(k+1)]}$ distributed as $F\left[k+1, T-2\left(k+1\right)\right]$.

Chow test

Special case of the Wald test:

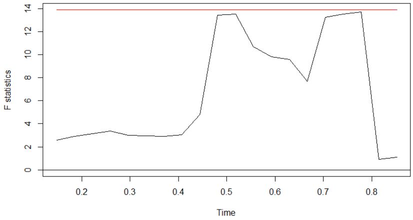
- split the sample into 2 subsamples;
  - baseline model treated as **restricted**: in assumes invariant parameters in both subsamples ($k+1$ constraints – as many as parameters);
  - in the unrestricted model, parameters can change between subsamples (in fact, 2 models estimated for both subsamples).

Chow test

$H_0 : \boldsymbol{\beta_A} = \boldsymbol{\beta_B}$, i.e. parameters equal in both subsamples (A and B)
$H_1 : \boldsymbol{\beta_A} \neq \boldsymbol{\beta_B}$, i.e. parameters different in both subsamples
Test statistic: $F = \frac{(RRSS - URSS)/(k+1)}{URSS/[T-2(k+1)]}$ distributed as $F\left[k+1, T-2\left(k+1\right)\right]$.

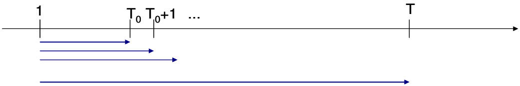
Chow test

In R, one can test for various breakpoints -- command ***Fstat*** from the package ***strucchange***:



F statistic values (black line) situated below the red line (critical value) indicate that no structural change can be confirmed for a given breakpoint.

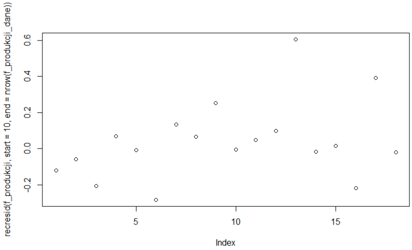
Recursive stability diagnostics (1)

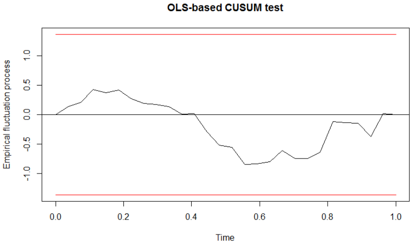


Sample: 1,...,T. Start at a given $T_0 \in (1;T)$, such that the estimation is feasible for the (long enough) sample 1,...,T_0. Models are estimated with samples 1,...,T_0; 1,...,$T_0 + 1$; ...; 1,...,T.

Recursive stability diagnostics (2)

- coefficient evolution after sample extension;
  - one-period-ahead forecast errors (graphical evaluation – big errors indicate a moment of potential break)
  - CUSUM (when the black line exits the red "corridor", this means a potential break)





Outline

Multicollinearity

Regressors are not independent:

- some are a linear combination of others (extreme case), then...
  - $\mathbf{X}^T\mathbf{X}$ is a singular matrix and we cannot compute $\beta = \left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\mathbf{y}$
- some are highly correlated (usual case), then...
  - $\mathbf{X}^T\mathbf{X}$ may not be singular, but its diagonal elements still close to zero...
  - ... then the diagonal elements of $\left(\mathbf{X}^T\mathbf{X}\right)^{-1}$ extremely high (and so the standard errors)

Multicollinearity – diagnostics

1. correlation matrix
   - only bilateral relationships
   - no cut-off value above which the problem can be considered serious
2. inflation variance factor (VIF) for regressor j
   - $VIF_j = \frac{1}{1-R_j^2}$
   - where R_j^2 is R-squared from the regression of variable j on the rest of the explanatory variables
   - limit value: 10, above – multicollinearity
3. condition index
   - $\kappa = \sqrt{\frac{\lambda_{max}}{\lambda_{min}}}$
   - where λ denotes eigenvalues of the matrix derived from $X^T X$ by division of every cell (i, j) by the product of diagonal elements (i, i) and (j, j)
   - limit value: 20-30, above – multicollinearity

Correlation matrix and VIF

```
summary(f_produkcji_translog)
cor(data.frame(l_L,l_C,f_produkcji_dane$time,l_L*l_C))
install.packages("corrgram")
library(corrgram)
corrgram(data.frame(l_L,l_C,f_produkcji_dane$time,l_L*l_C),lower.panel=panel.shade,upper.panel=NULL)
vif(f_produkcji_translog)
```

Conditional indices

- **Conditional indices** of regressor matrix

```
install.packages("perturb")
library(perturb)
colldiag(data.frame(l_L,l_C,f_produkcji_dane$time,l_L*l_C))
```

- In addition, the proportions in variance decomposition (>50%) indicate the source more specifically.

```
> colldiag(data.frame(l_L,l_C,f_produkcji_dane$time,l_L*l_C))
Condition
Index    Variance Decomposition Proportions
           intercept l_L   l_C   f_produkcji_dane.time l_L...l_C
1    1.000 0.000     0.000 0.000 0.007                 0.000
2    5.030 0.000     0.000 0.000 0.879                 0.000
3   13.192 0.003     0.000 0.000 0.000                 0.003
4   59.937 0.000     0.043 0.083 0.112                 0.001
5  337.663 0.997     0.957 0.917 0.001                 0.996
```

Multicollinearity – solutions

- strengthen the precision of estimation by **expanding sample size, removing a variable, imposing restrictions or calibrating the parameter**
- "manually" increase the diagonal values in $\mathbf{X}^T\mathbf{X}$ (***ridge regression***)
- "squeeze" the common variance of the collinear variables into a lower number of new, independent ones (***principal components***)

Multicollinearity – solutions

- strengthen the precision of estimation by **expanding sample size, removing a variable, imposing restrictions or calibrating the parameter**
- "manually" increase the diagonal values in $\mathbf{X}^T\mathbf{X}$ (***ridge regression***)
- "squeeze" the common variance of the collinear variables into a lower number of new, independent ones (***principal components***)

Multicollinearity – solutions

- strengthen the precision of estimation by **expanding sample size, removing a variable, imposing restrictions or calibrating the parameter**
- "manually" increase the diagonal values in $\mathbf{X}^T\mathbf{X}$ (***ridge regression***)
- "squeeze" the common variance of the collinear variables into a lower number of new, independent ones (***principal components***)

Example

Student satisfaction survey

Investigate the issue of multicollinearity in the student satisfaction model from lecture 1.