Introduction to econometrics

IV. Multiple linear regression model

Content

Introduction

- Multiple LRM discussed in more detail.
- Some proofs as an illustration.
- Hypothesis testing – extended methods.

Content

Classical assumptions

1. $E(\epsilon_i) = 0$.
2. $var(\epsilon_i) = E(\epsilon_i^2) = \sigma^2$.
3. $cov(\epsilon_i, \epsilon_j) = 0$ for $i \neq j$.
4. ϵ_i is Normally distributed.
5. $X_{1i}, \ldots, X_{ki}$ are fixed (non-random) variables.

Parameters estimates – two explaining variables

- Model:

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i.$$

- Minimize SSR:

$$\widehat{\beta}_1 = \frac{\left(\sum x_{1i} y_i\right)\left(\sum x_{2i}^2\right) - \left(\sum x_{2i} y_i\right)\left(\sum x_{1i} \sum x_{2i}\right)}{\left(\sum x_{1i}^2\right)\left(\sum x_{2i}^2\right) - \left(\sum x_{1i} x_{2i}\right)^2},$$

$$\widehat{\beta}_2 = \frac{\left(\sum x_{2i} y_i\right)\left(\sum x_{1i}^2\right) - \left(\sum x_{1i} y_i\right)\left(\sum x_{1i} \sum x_{2i}\right)}{\left(\sum x_{1i}^2\right)\left(\sum x_{2i}^2\right) - \left(\sum x_{1i} x_{2i}\right)^2},$$

$$\widehat{\alpha} = \overline{Y} - \widehat{\beta}_1 \overline{X}_1 - \widehat{\beta}_2 \overline{X}_2,$$

where

$$y_i = Y_i - \overline{Y},$$

$$x_{1i} = X_{1i} - \overline{X}_1,$$

$$x_{2i} = X_{2i} - \overline{X}_2.$$

OLS estimate – error terms variance

- Unbiased estimator, σ^2:

$$s^2 = \frac{\sum \widehat{\epsilon}_i^2}{N-k-1},$$

where

$$\widehat{\epsilon}_i = Y_i - \widehat{\alpha} - \widehat{\beta}_1 X_{1i} - \ldots - \widehat{\beta}_k X_{ki}$$

are OLS residuals.

Estimating variance of parameters estimates – two regressors

- Case $k = 2$:

$$var\left(\widehat{\beta}_1\right) = \frac{\sigma^2}{(1 - r^2)\sum x_{1i}^2},$$
$$var\left(\widehat{\beta}_2\right) = \frac{\sigma^2}{(1 - r^2)\sum x_{2i}^2},$$

  where r is (sample) correlation coefficient between X_1 and X_2.

- In practice – estimates of σ^2.
- Useful for hypothesis testing.

Test of parameter significance (assuming σ^2 is known)

- Koop (2008), p. 94.
- In practice σ^2 is not known $\Rightarrow$ t-test.

Measure of model fit

- Coefficient of determination:

$$R^2 = 1 - \frac{SSR}{TSS} = 1 - \frac{\sum \widehat{\epsilon}_i^2}{\sum \left(Y_i - \overline{Y}\right)^2}.$$

- Adding new explanatory variables will always increase R^2.
- Adjusted R^2, $\overline{R}^2$:

$$\overline{R}^2 = 1 - \frac{\frac{SSR}{N-k-1}}{\frac{TSS}{N-1}} = 1 - \frac{s^2}{\frac{1}{N-1}\sum \left(Y_i - \overline{Y}\right)^2}.$$

Content

Omitted variable bias I

- True model:

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i.$$

- OLS estimates:

$$\widehat{\beta}_1 = \frac{\left(\sum x_{1i} y_i\right)\left(\sum x_{2i}^2\right) - \left(\sum x_{2i} y_i\right)\left(\sum x_{1i} \sum x_{2i}\right)}{\left(\sum x_{1i}^2\right)\left(\sum x_{2i}^2\right) - \left(\sum x_{1i} x_{2i}\right)^2}.$$

- Lower-case letters – deviations from means.

Omitted variable bias II

- Model:

$$Y_i = \alpha + \beta_1 X_{1i} + \epsilon_i.$$

- Parameter estimate of β_1:

$$\tilde{\beta}_1 = \frac{\sum x_{1i} y_i}{\sum x_{1i}^2},$$

- $\tilde{\beta}_1$ is biased.

Omitted variable bias – proof

- May be shown:

$$\begin{aligned} E\left(\tilde{\beta}_1\right) &= E\left(\beta_1 + \frac{\beta_2 \sum x_{1i}x_{2i}}{\sum x_{1i}^2} + \frac{\sum x_{1i}\left(\epsilon_i - \overline{\epsilon}\right)}{\sum x_{1i}^2}\right) \\ &= \beta_1 + \frac{\beta_2 \sum x_{1i}x_{2i}}{\sum x_{1i}^2}. \end{aligned}$$

- $\tilde{\beta}_1$ is biased.

Omitted variable bias – comments

- Bias does not exist in case $\beta_2 = 0$ or $\frac{\sum x_{1i}x_{2i}}{\sum x_{1i}^2}$.
- If $\beta_2 = 0$ then X_2 is not omitted.
- $\frac{\sum x_{1i}x_{2i}}{\sum x_{1i}^2}$ connected with correlation between X_1 and X_2 (denoted by r).
- Bias does not arise if omitted variable is uncorrelated with included variabe.

Inclusion of irrelevant explanatory variables

- True model: $Y_i = \alpha + \beta_1 X_{1i} + \epsilon_i$.
- Incorrect specification: $Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \epsilon_i$.
- Wrong estimator:

$$\tilde{\beta}_1 = \frac{\left(\sum x_{1i} y_i\right)\left(\sum x_{2i}^2\right) - \left(\sum x_{2i} y_i\right)\left(\sum x_{1i} \sum x_{2i}\right)}{\left(\sum x_{1i}^2\right)\left(\sum x_{2i}^2\right) - \left(\sum x_{1i} x_{2i}\right)^2}.$$

- Correct estimator:

$$\widehat{\beta}_1 = \frac{\sum x_{1i} y_i}{\sum x_{1i}^2}.$$

- If $\tilde{\beta}_1$ unbiased, then using Gauss-Markov theorem $var\left(\tilde{\beta}_1\right) > var\left(\widehat{\beta}_1\right)$.
- Including irreevant variables leads to less precise estimates.

Multicollinearity

- High or perfect correlation among explanatory variables.
- OLS estimator has problem estimating separate marginal effects.
- Two explanatory variables:

$$var\left(\widehat{\beta}_1\right) = \frac{\sigma^2}{(1-r^2)\sum x_{1i}^2},$$
$$var\left(\widehat{\beta}_2\right) = \frac{\sigma^2}{(1-r^2)\sum x_{2i}^2}.$$

- Used in hypothesis testing. High multicolinearity $\rightarrow$ small t-statistic, wide confidence intervals.

Content

Introduction

- General model:

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots + \beta_k X_{ki} + \epsilon_i.$$

- Hypothesis including more parameters.
- F-tests and likelihod ratio tests.

Content

Basic test

- Test $R^2 = 0$ meets the hypothesis:

$$H_0 : \beta_1 = \ldots = \beta_k = 0.$$

- not the same as k individual hypothesis $H_0 : \beta_1 = 0$, $H_0 : \beta_2 = 0$ až $H_0 : \beta_k = 0$.
- F-statistics for a model with k explanatory variabes and an intercept:

$$F = \frac{R^2}{1 - R^2} \frac{N - k - 1}{k}.$$

- Assuming null hypothesis is true, F-statistics is distributed as $F_{k,N-k-1}$.

General tests

- Unrestricted model:

$$Y_i = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 X_{3i} + \epsilon_i.$$

- For example:

$$H_0 : \beta_1 = \beta_2 = 0.$$

- Able to include any linear restrictions: $a\beta_1 + b\beta_2 + c\beta_3 = d$ for some constants a, b, c a d.
- Restricted model:

$$Y_i = \alpha + \beta_3 X_{3i} + \epsilon_i.$$

General tests — examples

- Hypothesis:

$$H_0 : \beta_1 = 0, \quad \beta_2 + \beta_3 = 1.$$

- Second restriction ay be written as $\beta_2 = 1 - \beta_3$.
- Resticted model:

$$Y_i - X_{2i} = \alpha + \beta_3 (X_{3i} - X_{2i}) + \epsilon_i.$$

- Simple LRM with dependent variable $Y - X_2$, with an intercept and the explaining variable $(X_{3i} - X_{2i})$.

General tests – F-test

- For inear restrictions; test statistics:

$$F = \frac{\left(SSR_R - SSR_{UR}\right)/q}{SSR_{UR}/\left(N-k-1\right)}.$$

- SSR is sum of squared residuals, subscripts UR (unresticted model) and R (restricted model).
- Number of restrictions is q.
- Intuition: „big“ values of F suggest H_0 is not correct.
- F is distributed as $F_{q,N-k-1}$.
- F-statistics using R^2 (only for the same dependent variables in both models):

$$F = \frac{\left(R^2_{UR} - R^2_{R}\right)/q}{\left(1-R^2_{UR}\right)/\left(N-k-1\right)}.$$

Content

Motivace

- More complicated than F-test $\times$ wider variety of applications.
- Likelihood function:

$$
\begin{aligned}
& L\left(\alpha, \beta_1, \ldots, \beta_k, \sigma^2\right) \\
& = \prod_{i=1}^{N} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{1}{2\sigma^2}\left(Y_i - \alpha - \beta_1 X_{1i} - \ldots - \beta_k X_{ki}\right)^2\right] \\
& = \frac{1}{\left(2\pi\sigma^2\right)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\left(Y_i - \alpha - \beta_1 X_{1i} - \ldots - \beta_k X_{ki}\right)^2\right].
\end{aligned}
$$

Some basic results

- ML estimates correspond to OLS estimates: $\widehat{\alpha}$, $\widehat{\beta}_1, \ldots, \widehat{\beta}_k$.
- ML estimate of the error terms variance is biased:

$$\widehat{\sigma}^2 = \frac{\sum \left(Y_i - \widehat{\alpha} - \widehat{\beta}_1 X_{1i}, \ldots, \widehat{\beta}_k X_{ki} \right)^2}{N} = \frac{\sum \widehat{\epsilon}_i^2}{N}.$$

- Likelihood for unrestricted MLEs:

$$L\left(\widehat{\alpha}^U, \widehat{\beta}_1^U, \ldots, \widehat{\beta}_k^U, \widehat{\sigma}^{2U}\right).$$

- Likelihood evaluated at restricted MLEs:

$$L\left(\widehat{\alpha}^R, \widehat{\beta}_1^R, \ldots, \widehat{\beta}_k^R, \widehat{\sigma}^{2R}\right).$$

Ilustration

- Three explaining variables.

$$H_0 : \beta_1 = 0, \quad \beta_2 + \beta_3 = 1.$$

- Restricted model

$$Y_i - X_{2i} = \alpha + \beta_3 \left(X_{3i} - X_{2i} \right) + \epsilon_i .$$

- OLS estimates $\rightarrow \widehat{\alpha}^R$ a $\widehat{\beta}_3^R$.
- Values of $\widehat{\beta}_1^R$ and $\widehat{\beta}_2^R$? $\rightarrow$ restrictions from H_0, $\widehat{\beta}_1^R$ and $\widehat{\beta}_2^R = 1 - \widehat{\beta}_3^R$.
- Possible noninear restrictions, e.g.: $H_0 : \beta_1 = \beta_2^3,, \beta_3 = \frac{1}{\beta_2} \rightarrow$ in general $H_0 : g(\beta_1, \ldots, \beta_k) = 0$, where $g(\cdot)$ is a set of k noninear functions.
- Non-linear estimates using econometric software.

Likelihood ratio test

- Likelihood ratio:

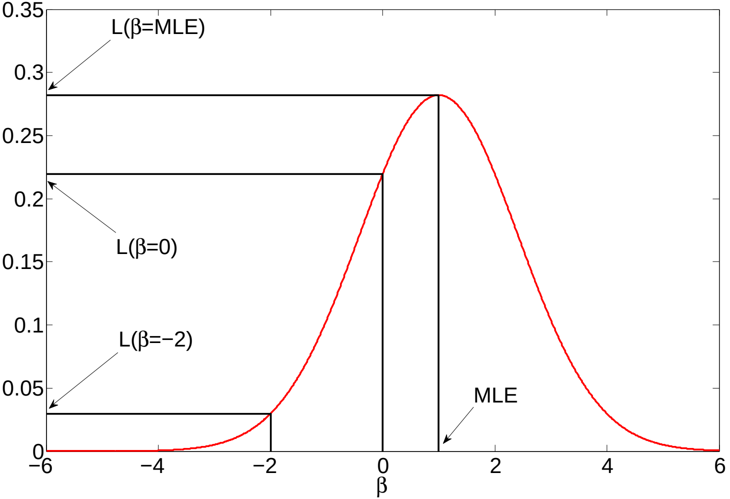
$$\lambda = \frac{L\left(\widehat{\alpha}^R, \widehat{\beta}_1^R, \ldots, \widehat{\beta}_k^R, \widehat{\sigma}^{2R}\right)}{L\left(\widehat{\alpha}^U, \widehat{\beta}_1^U, \ldots, \widehat{\beta}_k^U, \widehat{\sigma}^{2U}\right)}.$$

- Test statistics $-2\ln(\lambda)$.
- Statistics is distributed (approximately) as χ^2: $-2\ln(\lambda) \sim \chi^2_q$ (q is a number of restrictions in H_0).
- Intuition: including restrictions leads to a lower likelihood.
- Platí: $L\left(\widehat{\alpha}^R, \widehat{\beta}_1^R, \ldots, \widehat{\beta}_k^R, \widehat{\sigma}^{2R}\right) \leq L\left(\widehat{\alpha}^U, \widehat{\beta}_1^U, \ldots, \widehat{\beta}_k^U, \widehat{\sigma}^{2U}\right)$ a tedy $0 \leq \lambda \leq 1$.
- H_0 is true $\Rightarrow$ λ should be near 1 $\Rightarrow$ test statistics $-2\ln(\lambda)$ should be small.

Examples

- Koop (2008), pp. 107–108.
- Figure of $N(1,2)$.

Likelihood function



Alternative for LRM

- Likelihood function:

$$L\left(\widehat{\alpha}, \widehat{\beta}_1, \ldots, \widehat{\beta}_k, \widehat{\sigma}^2\right) = \frac{1}{(2\pi\widehat{\sigma}^2)^{\frac{N}{2}}} \exp\left[-\frac{1}{2\widehat{\sigma}^2}\sum_{i=1}^{N}\left(Y_i - \widehat{\alpha} - \widehat{\beta}_1 X_{1i} - \ldots - \widehat{\beta}_k X_{ki}\right)^2\right].$$

- Using variance estimates: $\widehat{\sigma^2}$:

$$L\left(\widehat{\alpha}, \widehat{\beta}_1, \ldots, \widehat{\beta}_k, \widehat{\sigma}^2\right) \propto \frac{1}{(\widehat{\sigma}^2)^{\frac{N}{2}}} \propto \frac{1}{(SSR)^{\frac{N}{2}}},$$

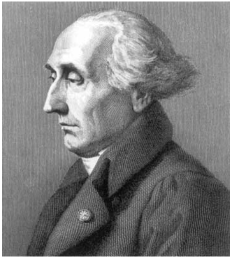
where $SSR = \sum \widehat{\epsilon}_i^2$.

- Likelihood ratio: $\lambda = \dfrac{\frac{1}{(SSR^R)^{\frac{N}{2}}}}{\frac{1}{(SSR^U)^{\frac{N}{2}}}} = \left(\dfrac{SSR^U}{SSR^R}\right)^{\frac{N}{2}}.$

Wald and Lagrange multiplier tests

- Approximations of LR test.

Abraham Wald
(1902–1950)



Joseph-Louis Lagrange
(1736–1813)

Wald test

- Only unrestricted estimates.
- Example: hypothesis $H_0 : g(\alpha, \beta_1, \beta_2, \ldots, \beta_k) = c$
- ML estimates $\widehat{\alpha}^U, \widehat{\beta}_1^U, \ldots, \widehat{\beta}_k^U$.
- Idea: if H_0 is true then unrestricted estimates should meet the restrictions (approximately).
- $g(\widehat{\alpha}^U, \widehat{\beta}_1^U, \ldots, \widehat{\beta}_k^U)$ near c.

Wald test – statistics

- Wald statistics:

$$W = \frac{\left[g\left(\widehat{\alpha}^U, \widehat{\beta}_1^U, \ldots, \widehat{\beta}_k^U\right) - c\right]^2}{var\left[g\left(\widehat{\alpha}^U, \widehat{\beta}_1^U, \ldots, \widehat{\beta}_k^U\right)\right]}.$$

- In some cases denominator easy to compute, e.g. for $g(\widehat{\alpha}^U, \widehat{\beta}_1^U, \ldots, \widehat{\beta}_k^U) = \widehat{\beta}_1^U + \widehat{\beta}_2^U$:

$$var\left(\widehat{\beta}_1^U + \widehat{\beta}_2^U\right) = var\left(\widehat{\beta}_1^U\right) + var\left(\widehat{\beta}_2^U\right) + 2cov\left(\widehat{\beta}_1^U, \widehat{\beta}_2^U\right).$$

- Non-linear restrictions $\rightarrow$ ekonometric software.
- Distribution of the test statistics:

$$W \sim \chi_q^2,$$

where q is number of restrictions.

Lagrange multiplier test

- Only restricted estimates.
- Example: unrestricted model, simple LRM, β; restricted model for $H_0 : \beta = c$.
- $\widehat{\beta}^R = c$.
- Motivation: if H_0 true, then MLE of restricted model should be close to unrestricted MLE (in our case, c should be near $\widehat{\beta}$ (OLS or ML estimate).
- Basic calculus: maximum of likelihood function, first derivative equals zero (slope).
- If H_0 true, then derivative of likelihood function evaluated at $\widehat{\beta}^R$ should be close to zero.

Lagrange multiplier test – statistics

- Test statistics:

$$LM = \frac{\left[d \ln L\left(\widehat{\beta}^R\right)\right]^2}{I\left(\widehat{\beta}^R\right)}.$$

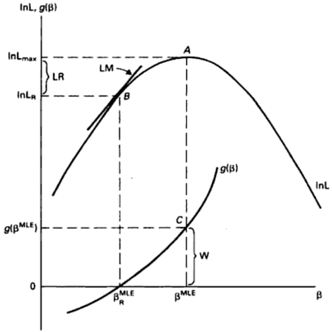
- Intuition: how far away from zero does the slope of the likelihood function become if we impose the restrictions?
- Numerator is the direct measure of its size $\times$ relative to its uncertainty.
- Denominator LM is related to the variance of the first derivative of the likelihood function: $I(\cdot)$ (*information matrix*).
- LM statistics is distributed approximately (assymptotically) as:

$$LM \sim \chi^2_q,$$

where q is the number of restrictions in H_0.

Comparing tests

- Likelihood ratio (LR) test, Wald test (W), Lagrange multiplier test (LM).
- Log-likelihood ($\ln L$) as a function of β; β^{MLE} maximum; restriction $g(\beta) = 0$; restricted value β_R^{MLE}.



Zdroj: Kennedy (2008) – A Guide to Econometrics.

Content

Choice of functional form

- Koop (2008), pp. 109–115 (including examples).
- Non-linearity in regression.
- Logarithms of the variables and interpretation of the parameters.
- Interaction terms and power of the variables– changing marginal effects.
- How to decide which non-linear form?
- Changing the measure of variables – any changes in estimates and appropriate statistics?