ASIL SHAAR

FINN3302

النمذجة المالية

CHAPTER 4: Further development and analysis of the classical linear regression model

Chapter 4

Multiple regression model

$$Y_t = B_1 + B_2 X_{2t} + B_3 X_{3t} + B_4 X_{4t} + \cdots + B_K X_{Kt} + U_t$$

interest constant term

(Cetris paribus effect)

slope parameters → represent the marginal impact of changing one of the explanatory variables associated with the parameter while holding other variables constant.

$$Y_1 = B_1 + B_2 X_{21} + B_3 X_{31} + \cdots + B_K X_{K1} + U_1$$
$$Y_2 = B_1 + B_2 X_{22} + B_3 X_{32} + \cdots + B_K X_{K2} + U_2$$
$$\vdots$$
$$Y_t = B_1 + B_2 X_{2t} + B_3 X_{3t} + \cdots + B_K X_{Kt} + U_t$$

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_T \end{bmatrix}_{T*1} = \begin{bmatrix} 1 & X_{21} & X_{31} & \cdots & X_{K1} \\ 1 & X_{22} & X_{32} & \cdots & \\ 1 & \vdots & \vdots & & \\ 1 & X_{2t} & X_{3t} & \cdots & X_{Kt} \end{bmatrix}_{T*K} * \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_K \end{bmatrix}_{K*1} + \begin{bmatrix} U_1 \\ U_2 \\ \vdots \\ U_t \end{bmatrix}_{T*1}$$

Matrix dimensions: number of rows * number of columns

$$\hat{u} = \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ \hat{u}_T \end{bmatrix}_{T*1}$$

transpose

$$\hat{u}^{*} = [\hat{u}_1\ \hat{u}_2 \cdots \hat{u}_T]_{1*T}$$

$$\hat{u}^{*} * \hat{u} = [\hat{u}_1 \hat{u}_2 \cdots \hat{u}_T]_{1*T} * \begin{bmatrix} \hat{u}_1 \\ \hat{u}_2 \\ \vdots \\ u_T \end{bmatrix}_{T*1} = \sum \hat{u}^2 = RSS$$

⇒

$$\hat{B} = (x'x)^{-1} x'y \qquad \begin{bmatrix} \hat{B}_1 \\ \hat{B}_2 \\ \vdots \\ \hat{B}_k \end{bmatrix}_{k*1}$$

Ols estimator in matrix notation.

$$S^2 = \frac{\hat{u}'\hat{u}}{T-K}$$

K → number of parameters to be estimated in the multiple regression Case including the Constant term.

$$S^2 (x'x)^{-1}$$

var of the estimate

→ Variance Covariance matrix

Example:

$K = 3 \quad T = 15$

$$Y_t = B_1 + B_2 X_{2t} + B_3 X_{3t} + u_t$$

$$(x'x)^{-1} = \begin{bmatrix} 2 & 3.5 & -1 \\ 3.5 & 1 & 6.5 \\ -1 & 6.5 & 4.3 \end{bmatrix}_{3\times3} \qquad x'y = \begin{bmatrix} -3 \\ 2.2 \\ 0.6 \end{bmatrix}_{3\times1} \qquad \hat{u}'\hat{u} = RSS = 10.96$$

$$\hat{B} = (x'x)^{-1} x'y = \begin{bmatrix} 2 & 3.5 & -1 \\ 3.5 & 1 & 6.5 \\ -1 & 6.5 & 4.3 \end{bmatrix}_{3\times3} \begin{bmatrix} -3 \\ 2.2 \\ 0.6 \end{bmatrix}_{3\times1}$$

$$= \begin{bmatrix} 2*-3 + 3.5*2.2 + -1*0.6 \\ 3.5*-3 + 1*2.2 + 6.5*0.6 \\ -1*-3 + 6.5*2.2 + 4.3*0.6 \end{bmatrix}_{3*1} = \begin{bmatrix} 1.1 \\ -4.4 \\ 19.88 \end{bmatrix}_{3*1} \begin{matrix} \to \hat{B}_1 \\ \to \hat{B}_2 \\ \to \hat{B}_3 \end{matrix}$$

$$\hat{Y}_t = \underset{(1.35)}{1.1} - \underset{(0.96)}{4.4}\,X_{2t} + \underset{(1.98)}{19.88}\,X_{3t}$$

$$\hat{S}^2 = \frac{\hat{u}'\hat{u}}{T-K} = \frac{RSS}{T-K} = \frac{10.96}{15-3} = 0.91$$

$$(X'X)^{-1} = \begin{bmatrix} 2 & 3.5 & -1 \\ 3.5 & 1 & 6.5 \\ -1 & 6.5 & 4.3 \end{bmatrix}$$

$$0.91 * \begin{bmatrix} 2 & 3.5 & -1 \\ 3.5 & 1 & 6.5 \\ -1 & 6.5 & 4.3 \end{bmatrix} = \begin{bmatrix} 1.83 & 3.2 & -0.91 \\ 3.2 & 0.91 & 5.91 \\ -0.91 & 5.91 & 3.91 \end{bmatrix}$$

$$\left.\begin{aligned} var(\hat{B}_1) &= 1.83 \rightarrow SE(\hat{B}_1) = \sqrt{1.83} = 1.35 \\ var(\hat{B}_2) &= 0.91 \rightarrow SE(\hat{B}_2) = \sqrt{0.91} = 0.96 \\ var(\hat{B}_3) &= 3.91 \rightarrow SE(\hat{B}_3) = \sqrt{3.91} = 1.98 \end{aligned}\right]$$

نرجع نكتبهم في المعادلة

$0 \leq R^2 \leq 100\%$

↓ how well our model (SRF) fits the data

$TSS = \sum (Y_t - \bar{Y}_t)^2$

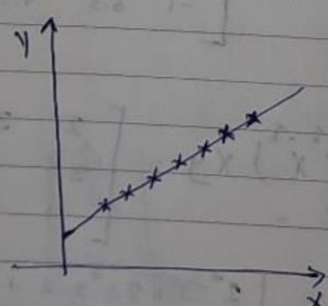
explained sum of squares — Residual sum of squares

$TSS = \hat{ESS} + RSS$

$\sum(Y_t - \bar{Y})^2 = \sum(\hat{Y}_t - \bar{Y})^2 + \sum(Y_t - \hat{Y}_t)^2$

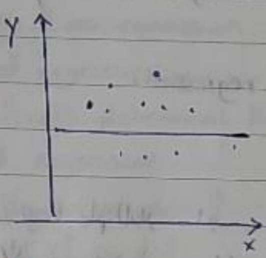
إذا $R^2 = 1$ يعني $RSS = 0$

$TSS = ESS$



→

اذا $R^2 = 0$ يعني ما بقدر اوضح العلاقة



$ESS = 0$

$TSS = RSS$

$$\boxed{R^2 = \frac{ESS}{TSS}}$$

$$\frac{TSS}{TSS} = \frac{ESS}{TSS} + \frac{RSS}{TSS} \qquad \frac{ESS}{TSS} = 1 - \frac{RSS}{TSS}$$

$$1 = \frac{ESS}{TSS} + \frac{RSS}{TSS} \qquad \boxed{R^2 = 1 - \frac{RSS}{TSS}}$$

problems with R^2:

هاي المشكلة مالها حل

(1) we cannot use R^2 to compare between two models if the dependent variable was different.

Example:

model 1 — بس بقدر افسر R^2

$$y_t = B_1 + B_2 X_{2t} + u_t$$

$R^2 = 0.8$ — 80% of the variation in Y was explained by the variation in X.

model 2

$$Z_t = B_1 + B_2 X_{2t} + U_t$$

$R^2 = 0.6$ — 60% of the variation in Z was explained by the variation in X

ما بقدر احكي انه model 1 افضل من model 2

$\rightarrow$

(2) R^2 will never fall if we add explanatory variables → وجدوا انه حلها باستخدام adjusted R^2

(3) R^2 takes high values in time series regression.

Data mining
is searching many series for statistical relationships without theoretical justification

بمعنى الـ Multiple Regression احسن من الـ Simple Regression to add variables
↓ بس ما نزيد

ممكن يكون علاقات بدون Theory و ما يكون إلها معنى

في الـ Multiple Regression انفسر الـ slop بحط هاي الجملة
holding other variables constant.

F-test

Cost of shipment $= B_1 + \underline{B_2} * \text{Package weight}_t + \underline{B_3} * \text{distance shipped}_t + U_t$

Example:

للي انون effect الـ package weight على Cost of shipment هل يساوي distance shipped effect على Cost of shipment

$$H_0: B_2 = B_3$$
$$H_1: B_2 \neq B_3$$

two sided test دائمًا

unrestricted regression → بنخلي ols يحدد prameter

restricted regression → القيود مفروضة على Cofficent

→

F-test

1. We estimate two regression equations which are:-
   a) unrestricted regression.
      ↳ OLS freely determines the value of the parameters.
   b) restricted regression.
      ↳ take the values under the null hypothesis, impose them and then rearrange the model if needed.

Example:

$$Y_t = B_1 + B_2 X_{2t} + B_3 X_{3t} + B_4 X_{4t} + u_t$$

$H_0: B_3 + B_4 = 1$ F-test

$H_1: B_3 + B_4 \neq 1$

- unrestricted regression:

$$Y_t = B_1 + B_2 X_{2t} + B_2 X_{3t} + B_4 X_{4t} + u_t$$

- restricted regression:

مش فارق شو بستخدم بطلع نفس النتيجة.

$$\begin{cases} B_3 = 1 - B_4 \\ B_4 = 1 - B_3 \end{cases}$$

$$Y_t = B_1 + B_2 X_{2t} + B_3 X_{3t} + (1 - B_3) X_{4t} + u_t$$

$$Y_t = B_1 + B_2 X_{2t} + B_3 X_{3t} + \underline{X_{4t}} - B_3 X_{4t} + u_t$$

بدون coeffecient

لازم أنقلها للجانب الآخر

$$Y_t - X_{4t} = B_1 + B_2 X_{2t} + B_3 X_{3t} - B_3 X_{4t} + u_t$$

$$Y_t - X_{4t} = B_1 + B_2 X_{2t} + B_3 (X_{3t} - X_{4t}) + u_t$$

let $P_t = Y_t - X_{4t}$

let $Q_t = X_{3t} - X_{4t}$

$$\boxed{P_t = B_1 + B_2 X_{2t} + B_3 Q_t + u_t} \rightarrow$$

[2] Calculate F - statistic :

$$\text{F-stat} = \frac{RRSS - URSS}{URSS} \times \frac{T - K}{m}$$

RSS from the restricted regression

RSS from the unrestricted regression

number of prameters to be estimated in the unrestricted regression including the costant term

number of restriction

$H_0: B_3 + B_4 = 1$

$H_1: B_3 + B_4 \neq 1$, $\boxed{B_4 = 1 - B_3}$

بنعد عدد الاشارات (=) في H_0 عشان نطلع عدد restriction.

F-stat takes positive values

F-stat ~ F-distribution.

[3] get Critical values

[4] perform the test: if F-stat > F-critical then Reject H_0

Hypothesis of interst:

Example:

$$Y_t = B_1 + B_2 X_{2t} + B_3 X_{3t} + B_4 X_{4t} + B_5 X_{5t} + U_t$$

joint significance:

$H_0: B_2 = 0$ and $B_3 = 0$ and $B_4 = 0$ and $B_5 = 0$

$H_1: B_2 \neq 0$ or $B_3 \neq 0$ or $B_4 \neq 0$ or $B_5 \neq 0$.

اي Hypothesis بنستخدمها t-test بقدر استخدمها F-test (والعكس غير صحيح)

→

$H_0: \beta_2\beta_3 = 2$ or $H_0: \beta_2^2 = 1$ Cannot be tested. * ما نقدر نستخدمهم لل t-test و F-test

Example:

$$Y_t = B_1 + B_2 X_{2t} + B_3 X_{3t} + B_4 X_{4t} + u_t \qquad T = 144$$

$H_0: B_2 = 1$ and $B_3 = 1$

$H_1: B_2 \neq 1$ or $B_3 \neq 1$

Q.1

unrestrected regression:

$$Y_t = B_1 + B_2 X_{2t} + B_3 X_{3t} + B_4 X_{4t} + u_t$$

Restrected regression:

$$Y_t = B_1 + X_2 + X_3 + B_4 X_{4t} + u_t$$

$$Y_t - X_{2t} - X_{3t} = B_1 + B_4 X_{4t} + u_t$$

let $Z_t = Y_t - X_{2t} - X_{3t}$

$$- Z_t = B_1 + B_4 X_{4t} + u_t$$

الأكبر دائماً RRSS

Q.2 If two RSS are 436.1 and 397.2 respectively, perform the test.

RRSS = 436.1 URSS = 397.2

$$F\text{-stat} = \frac{RRSS - URSS}{URSS} \times \frac{T - K}{m}$$

$$\frac{436.1 - 397.2}{397.2} \times \frac{144 - 4}{2} = \boxed{6.68}$$

$F\text{-critical} = \boxed{3.07}$

F-stat > F-critical then Reject H_0.

$\Rightarrow$

problem 2, 3, 5, 6, 8

* problem 2 Page 176

ⓐ $H_0: B_3 = 2$

t-test / F-test

$m = 1$

ⓑ $H_0: B_3 + B_4 = 1$

F-test

$m = 1$

ⓒ $H_0: B_3 + B_4 = 1$ and $B_5 = 1$

F-test

$m = 2$

ⓓ $H_0: B_2 = 0$ and $B_3 = 0$ and $B_4 = 0$ and $B_5 = 0$

Joint significant

F-test

$m = 4$

ⓔ $H_0: B_2 B_3 = 1$

neither t-test nor F-test

* problem 3 Page 176.

$$y_t = B_1 + B_2 X_{2t} + B_3 X_{3t} + B_4 X_{4t} + B_5 X_{5t} + U_t$$

اي وحدة من دول بتبقى Joint significant

Hypothesis d

$H_0: B_2 = 0$ and $B_3 = 0$ and $B_4 = 0$ and $B_5 = 0$

$H_1: B_2 \neq 0$ or $B_3 \neq 0$ or $B_4 \neq 0$ or $B_5 \neq 0$

Joint significance test.

* problem 4 page 177

$RRSS \geq URSS$

OLS method Choose the Coefficients when RSS is at it's minimum.

* problem 5 page 177

$H_0: B_3 + B_4 = 1$ and $\boxed{B_5 = 1}$

$\boxed{B_4 = 1 - B_3}$

$$Y_t = B_1 + B_2 X_{2t} + B_3 X_{3t} + (1 - B_3) X_{4t} + X_{5t} + U_t$$

$$Y_t = B_1 + B_2 X_{2t} + B_3 X_{3t} + X_{4t} - B_3 X_{4t} + X_{5t} + U_t$$

$\Rightarrow$

$$Y_t - X_{4t} - X_{5t} = B_1 + B_2 X_{2t} + B_3 X_{3t} - B_3 X_{4t} + U_t$$
$$Y_t - X_{4t} - X_{5t} = B_1 + B_2 X_{2t} + B_3 (X_{3t} - X_{4t}) + U_t$$

let $Z_t = Y_t - X_{4t} - X_{5t}$

let $Q_t = X_{3t} - X_{4t}$

restricted model

$$Z_t = B_1 + B_2 X_{2t} + B_3 Q + U_t$$

$$F\text{-stat} = \frac{RRSS - URSS}{URSS} * \frac{T-k}{m}$$

$$\frac{102.87 - 91.41}{91.41} * \frac{96-5}{2} = \boxed{5.7}$$

using excel F-critical $= 3.09$

$5.7 > 3.09$ then Reject H_0

* problem 6 page 177

$$t\text{-stat} = \frac{\hat{B}}{SE(\hat{B})}$$

T = 200
d.f. = 195
α = 5%

market to book ratio $= \frac{\text{market price / share}}{\text{BV / share}}$

BV / share $= \frac{\text{Total common equity}}{\text{number of common shares outstanding}}$

P/E ratio $= \frac{\text{market price / share}}{\text{EPS}}$

t-critical = 1.97

* $\hat{B}_1$ t-ratio $= \frac{0.08}{0.064} = 1.25$

$|1.25| < 1.97$ then fail the Reject H_0

* $\hat{B}_2$ t-ratio $= \frac{0.801}{0.147} = 5.4$

$|5.4| > 1.97$ then Reject H_0

There is a relationship between size of the firm and the stock Return.

* $\hat{B}_3$ t-ratio $= \frac{0.321}{0.136} = 2.3$

$|2.3| > 1.97$ then Reject H_0

There is a relationship between MB ratio and stock Return.

$\Rightarrow$

* $\hat{B}_4$ t-ratio $= \frac{0.164}{0.42} = 0.39$

$|0.39| < 1.97$ then fail to Reject H_0

There is no arelationship between P/E ratio and stock Return.

* $\hat{B}_5$ t-ratio $= \frac{-0.084}{0.120} = -0.7$

$|-0.7| = 0.7 < 1.97$ then fail to Reject H_0

There is no a relationship between beta and Stock Return.

* Delete P/E, Beta.

$B = \frac{\Delta y}{\Delta x}$ $\quad -0.084 = \frac{\Delta y}{1.2-1}$ $\quad = -0.084(0.2) = \Delta y$

$\boxed{\Delta y = -1.68\%}$

Problem 8 page 177

Second model → R^2 is higher (extra explanatory variable)

second model → adjusted R^2 is expected to be higher but we should consider degrees of freedom.