

Øvelsestime 5

Flere restriktioner (F-test) og MATA

Agenda

- 1) Introduktion til PS5
- 2) Gruppearbejde: Diskussion
- 3) Opsamling
- 4) Gruppearbejde: Stata/MATA øvelser
- 5) Opsamling

Tobias Gabel Christiansen
tgc@econ.ku.dk
22-03-2021

KØBENHAVNS UNIVERSITET



1) Planen for semesteret

Dato	Opgave
Uge 7 - 15/02	PS1
Uge 8 - 22/02	PS2
Uge 9 - 01/03	PS3
Uge 10 - 08/03	OO1
Uge 11 - 15/03	PS4
Uge 12 - 22/03	PS5 + gennemgang af OO1
Uge 13 - 29/03	PS6
Uge 14 - 05/04	Påskeferie
Uge 15 - 12/04	OO2
Uge 15 - 19/04	PS7
Uge 16 - 26/04	PS8 + gennemgang af OO2
Uge 17 - 03/05	OO3
Uge 18 - 10/05	PS9
Uge 19 - 17/05	PS10 + gennemgang af OO3
	Spørgetime

← Her er vi :-)

1) Introduktion til problem set 5

- I Problem set 3 brugte vi data fra Makro I til at estimere hvorvidt absolut og betinget konvergens fandt sted:

$$gy_i = \beta_0 + \beta_1 \ln(y60_i) + u_i \quad (\text{PS3-1})$$

$$gy_i = \beta_0 + \beta_1 \ln(y60_i) + \beta_2 \text{struc}K_i + v_i \quad (\text{PS3-2})$$

- Vi udvider nu modellen for betinget konvergens ved også at inkludere humankapital:

$$gy_i = \beta_0 + \beta_1 \ln(y60_i) + \beta_2 \text{struc}K_i + \beta_3 \text{struc}H_i + \epsilon_i \quad (\text{PS5-1})$$

- Hvor $\text{struc}K_i = \ln(sk_i) - \ln(n_i + 0,075)$ og $\text{struc}H_i = \ln(sh_i) - \ln(n_i + 0,075)$

1) Introduktion til Problem set 5

Absolut konvergens:

- På langt sigt vil BNP pr. arbejder konvergere mod én og samme vækstrate for alle lande, så alle lande konvergere mod samme niveau for BNP pr. arbejder.
- Lande med relativt lav BNP pr. arbejder ved et initialt tidspunkt vil opleve relativt højere vækst.

Betinget konvergens:

- På langt sigt vil BNP pr. arbejder konvergere mod en lande-specifik vækstrate som er bestemt ud fra strukturelle egenskaber (fx investeringsrater for fysisk og human kapital).
- BNP pr. arbejder konvergerer mod samme niveau for lande der strukturelt set er ens.
- Lande med relativt lav BNP pr. arbejder ved et initialt tidspunkt oplever højere vækst for givne strukturelle niveauer.

2) Gruppearbejde - Diskussion

To-do:

- Diskutér opgave 1. og 2. i PS5
- Vi samler op om 10 minutter

1. The extended Solow model (1) includes the simple model for absolute convergence as a special case. Write up the parameter restrictions to be imposed as a null hypothesis on model (1) to make it possible to
2. What assumptions must be met in order for the proposed test to be valid?

Brug for hjælp? Skriv Jeres gruppenummer i Padlet (link i chatten)

3) Opsamling

1. The extended Solow model (1) includes the simple model for absolute convergence as a special case. Write up the parameter restrictions to be imposed as a null hypothesis on model (1) to make it possible to

3) Opsamling - Diskussion

1. The extended Solow model (1) includes the simple model for absolute convergence as a special case. Write up the parameter restrictions to be imposed as a null hypothesis on model (1) to make it possible to

De to modeller vi ønsker at sammenligne er givet ved:

$$gy_i = \beta_0 + \beta_1 \ln(y60_i) + u_i \quad (0)$$

$$gy_i = \beta_0 + \beta_1 \ln(y60_i) + \beta_2 strucK_i + \beta_3 strucH_i + \epsilon_i \quad (1)$$

Model (0) er **nested** i (1). Med andre ord er (0) et specialtilfælde af (1), hvor $\beta_2 = \beta_3 = 0$. Vi kan derfor teste om model (1) er bedre til at forklare variationen i gy_i end model (0) med følgende nulhypotese:

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

Alternativhypotesen er at én eller begge parametre er forskellig fra nul:

$$H_a: \beta_1 \neq 0 \vee \beta_2 \neq 0$$

Dvs. at (1) forklarer signifikant mere af variationen i gy_i end (0)

Vi kan benytte en **F-test** med to restriktioner til at teste H_0 . Hvis vi kan forkaste H_0 så forkaster vi modellen for absolut konvergens (0) til fordel for modellen for betinget konvergens (1).

3) Opsamling - Diskussion

F-statistikken er givet ved:

$$F = \frac{\frac{(SSR_R - SSR_{UR})}{q}}{\frac{SSR_{UR}}{n - k - 1}} \geq 0$$

SSR_R = Sum of squared residuals for den restrikerede model (dvs. modellen hvor vi sætter $\beta_2 = \beta_3 = 0$)

q = Angiver antallet af restriktioner (dvs. to)

SSR_{UR} = Sum of squared residuals for den urestrikerede model ($\beta_2 \neq 0 \vee \beta_3 \neq 0$)

$n - k - 1$ = Angiver antallet af frihedsgrader i den urestrikerede model

- F-statistikken undersøger om den urestrikerede (UR) model er signifikant bedre til at forklare variationen i den afhængige variabel sammenlignet med den restrikerede model (R).
- Den urestrikerede model (UR) vil altid forklare mindst lige så meget af variationen som den restrikerede model (R). Med andre ord er $SSR_R \geq SSR_{UR}$ (samme årsag som at R^2 aldrig falder når vi inkluderer flere kontrolvariable).
- Vi vil dog gerne vide om ændringen i SSR er signifikant eller blot skyldes at inklusionen af flere variable tilfældigvis fanger noget af støjen i vores sample.
- Store ændringer i SSR når vi går fra den urestrikerede model (UR) til den restrikerede model (R) indikerer at det ikke blot skyldes at vi fanger tilfældig støj men derimod at den urestrikerede model (UR) rent faktisk er bedre til at forklare variationen i den afhængige variabel. Dette slår ud som en høj værdi for F-statistikken (bevis imod H_0)

3) Opsamling - Diskussion

2. What assumptions must be met in order for the proposed test to be valid?

3) Opsamling - Diskussion

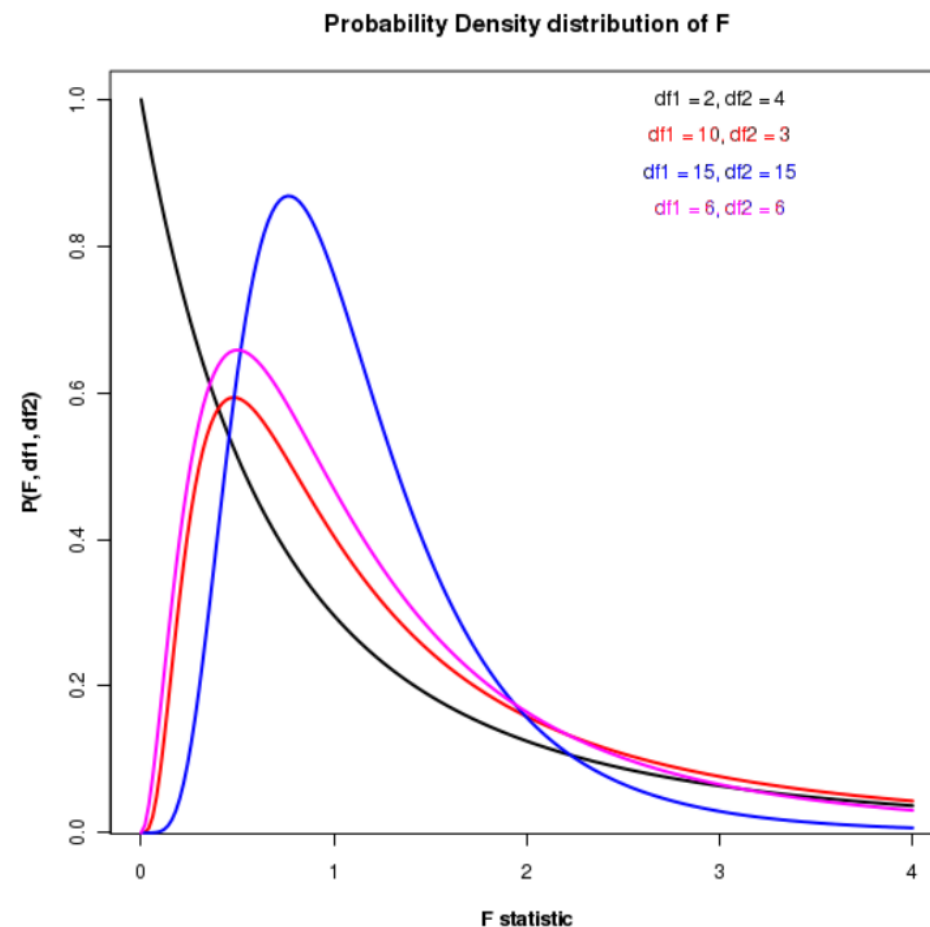
2. What assumptions must be met in order for the proposed test to be valid?

Under MLR.1-6 følger F-statistikken (eksakt) en F-fordeling under H_0 :

$$F = \frac{\frac{(SSR_R - SSR_{UR})}{q}}{\frac{SSR_{UR}}{n - k - 1}} \sim F_{q, n-k-1}$$

F-statistikken følger dog også asymptotisk en F-fordeling når MLR.6 er brudt (dvs. uden normalfordelte fejllæd). Der gælder derfor også at under MLR.1-5 og $n \rightarrow \infty$, så følger F-statistikken en F-fordeling under H_0 :

$$F = \frac{\frac{(SSR_R - SSR_{UR})}{q}}{\frac{SSR_{UR}}{n - k - 1}} \sim^a F_{q, n-k-1} \quad \text{for } n \rightarrow \infty$$



4) Opsamling Stata

To-do:

- Løs Stataøvelserne
- Vi samler op om 1 time

Hints:

1.a: Manglende værdier i Stata gemmes som "." – brug denne viden til at droppe disse observationer

1.c: Stata rapporterer automatisk en F-test der tester om alle parametre (bortset fra konstantleddet) er lig 0. Findes i øverste højre hjørne af regressionsoutputtet.

Generelt: Stata kan udregne kritiske værdier for både t- og F-statistikker:

display invttail($n - k - 1, \alpha$)

display invFtail($q, n - k - 1, \alpha$)

Brug for hjælp? Skriv Jeres gruppenummer i Padlet (link i chatten)

5) Stata - opsamling

1.a Eliminate the observations with missing values for the variable, **sh**, and data quality "D". After this sample restriction, there should be 77 countries available in **pwt.dta**.

- Overvej altid hvilke observationer I frasorterer når I filtrerer data.
- Hvis vi systematisk fjerner en type lande kan det påvirke hvor repræsentativt vores sample er for populationen (sample selection bias).
- I PS3 så vi fx at lande med grade D havde både lavere initial BNP og lavere vækstrate.
- Ikke noget tydeligt mønster når vi frasorterer manglende værdier. Men måske er manglende værdier tegn på andre systematiske forskelle som vi umiddelbart ikke observerer?
- Svært at gøre noget ved. Det er trods alt kun det data vi har tilgængeligt. Men overvej altid om det kan have implikationer for den efterfølgende analyse eller ej, og gør læseren opmærksom på det!

```
generate ly60=log(y60)
generate strucK=log(sk)-log(n+0.075)
generate strucH=log(sh)-log(n+0.075)
list country if strucH==.
```

country
Luxembourg
Iceland
Barbados
Gabon
Iran
Romania
China
Guinea
Gambia. The

```
drop if grade=="D"
drop if strucH==.
describe
```

```
obs:      77
vars:     14
```

5) Stata - opsamling

1.b Estimate the extended model (1) using **regress** and comment on the estimated parameters. What is the value of R^2 for the regression — and what does it mean?

$\widehat{\beta}_1 < 0$: Når initial BNP pr. arbejder stiger med 1% så er det forventede fald i den gns. vækstrate (approx.) — 0,007 pp. Fortegnet taler for betinget konvergens (statistisk signifikant), men omvendt er effektens størrelse meget lav (lav økonomisk/praktisk signifikans).

$\widehat{\beta}_2, \widehat{\beta}_3 > 0$: Højere strukturelle niveauer (målt ved investeringsrater i fysisk og human kapital fratrullet befolkningsvækst og deprecierungsrate) implicerer højere vækstrater.

$R^2 = 0,42$: Modellen forklarer 42% af variationen i gennemsnitlige vækstrater over perioden.

`regress gy ly60 strucK strucH`

Source	SS	df	MS	Number of obs	=	77
Model	.004825571	3	.001608524	F(3, 73)	=	17.50
Residual	.006711089	73	.000091933	Prob > F	=	0.0000
				R-squared	=	0.4183
Total	.011536659	76	.000151798	Adj R-squared	=	0.3944
				Root MSE	=	.00959

gy	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ly60	-.0074633	.0016806	-4.44	0.000	-.0108126	-.0041139
strucK	.0123714	.0026831	4.61	0.000	.0070239	.0177188
strucH	.0053726	.0027139	1.98	0.052	-.0000362	.0107814
_cons	.0028857	.0028927	1.00	0.322	-.0028795	.0086509

5) Stata - opsamling

1.c What is the value of the F -test for the significance of the overall regression? Are the individual coefficients of the explanatory variables statistically significant? Explain what hypotheses you test using the F and t tests. Discuss what assumptions are needed for the test procedures to work.

- Stata rapporterer automatisk to-sidede t -tests for alle variable, som tester:

$$H_0: \beta_j = 0$$

$$H_a: \beta_j \neq 0$$

- Samt en F -test for hele modellen, som tester:

$$H_0: \beta_1 = \beta_2 = \beta_3 = 0$$

$$H_a: \text{mindst en } \beta_j \neq 0 \quad j = 1, 2, 3 \dots$$

- Bemærk F -testen inkluderer ikke $j = 0$. Med andre ord sammenligner vi vores model med en model der kun består af konstantleddet (som blot prædikerer $\hat{y}_i = \bar{y}$ for alle i)

```
regress gy ly60 strucK strucH
```

Source	SS	df	MS	Number of obs =	77
Model	.004825571	3	.001608524	F(3, 73)	= 17.50
Residual	.006711089	73	.000091933	Prob > F	= 0.0000
				R-squared	= 0.4183
				Adj R-squared	= 0.3944
Total	.011536659	76	.000151798	Root MSE	= .00959

gy	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ly60	-.0074633	.0016806	-4.44	0.000	-.0108126	-.0041139
strucK	.0123714	.0026831	4.61	0.000	.0070239	.0177188
strucH	.0053726	.0027139	1.98	0.052	-.0000362	.0107814
_cons	.0028857	.0028927	1.00	0.322	-.0028795	.0086509

5) Stata - opsamling

- 1.c What is the value of the F -test for the significance of the overall regression? Are the individual coefficients of the explanatory variables statistically significant? Explain what hypotheses you test using the F and t tests. Discuss what assumptions are needed for the test procedures to work.
- Stata rapporterer $F = 17,50$ med tilhørende p-værdi på 0,0000, så vi kan på et 0,01% signifikansniveau ($\alpha = 0,0001$) forkaste at alle $\beta_j = 0$. Med andre ord forklarer vores model signifikant mere af variationen i gy end hvis vi blot prædikterede $\hat{y}_i = \bar{y}$ for alle.

De individuelle t-statistikker viser at:

- $\log(y60)$ har en p-værdi på 0,000 og derfor er meget signifikant forskellig fra 0 (kan forkaste H_0 selv på 0,1% signifikansniveau [$\alpha = 0,001$])
- $struck$ har en p-værdi på 0,000 og er derfor meget signifikant forskellige fra 0 (kan forkaste H_0 selv på 0,1% signifikansniveau [$\alpha = 0,001$])
- $struch$ har en p-værdi på 0,052 og er derfor insignifikant på et 5% signifikansniveau [$\alpha = 0,05$] men signifikant på fx 5,3% signifikansniveau [$\alpha = 0,053$].

`regress gy ly60 struck struch`

Source	SS	df	MS	Number of obs =	77
Model	.004825571	3	.001608524	F(3, 73)	= 17.50
Residual	.006711089	73	.000091933	Prob > F	= 0.0000
Total	.011536659	76	.000151798	R-squared	= 0.4183
				Adj R-squared	= 0.3944
				Root MSE	= .00959

gy	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ly60	-.0074633	.0016806	-4.44	0.000	-.0108126	-.0041139
struck	.0123714	.0026831	4.61	0.000	.0070239	.0177188
struch	.0053726	.0027139	1.98	0.052	-.0000362	.0107814
_cons	.0028857	.0028927	1.00	0.322	-.0028795	.0086509

Både F- og t-statistikker er asymptotisk valide under MLR.1-5

5) Stata - opsamling

1.d Provide an appropriate alternative hypothesis for the t -test of $\beta_1 = 0$ in model (1) and perform the test. [Hint: What does the hypothesis of conditional convergence imply for β_1 ?]

- Stata tester automatisk mod et to-sidet alternativ for t -testen:

$$\begin{aligned}H_0: \beta_1 &= 0 \\ H_a: \beta_1 &\neq 0\end{aligned}$$

- Da vi kun er interesseret i at undersøge om konvergens ($\beta_1 < 0$) finder sted og er ligeglade med muligheden for divergens ($\beta_1 > 0$) kan vi argumentere for at bruge én-sidet test i stedet:

$$\begin{aligned}H_0: \beta_1 &= 0 \\ H_a: \beta_1 &< 0\end{aligned}$$

- Ved at gøre dette siger vi altså på forhånd at vi kun fokuserer på negative værdier for β_1 og ser fuldstændig bort fra positive værdier for β_1 . Med andre ord bruger vi vores egen viden til at indsnævre det vi undersøger (se slides til PS3).
- **Bemærk:** I modsætning til t -test kan vi ikke bruge F -testen overfor en-sidede alternativer.

5) Stata - opsamling

1.e Investigate using an F -test whether the structural characteristics together may be omitted from the regression model. Formulate the relevant null hypothesis.

- i. Calculate the F -test by hand. This can be done using the results from question 1.b and a regression of the restricted model where the null hypothesis has been imposed.

- Den relevante hypotese er:

$$H_0: \beta_2 = \beta_3 = 0$$

$$H_a: \text{mindst en } \beta_j \neq 0, j = 2, 3$$

- Vi estimerer den restriktede og urestriktede model og beregner F -statistikken:

$$F = \frac{\frac{(SSR_R - SSR_{UR})}{q}}{\frac{SSR_{UR}}{n - k - 1}} = \frac{\frac{0,01153 - 0,00671}{2}}{\frac{0,00671}{73}} = 26,2$$

- På et 5% signifikansniveau [$\alpha = 0,05$] er den kritiske værdi for $F_{2,73}$ givet ved:

```
. display invFtail(2,77-3-1,0.05)
3.1221029
```

Så vi forkaster H_0 .

Urestriktede model:

```
regress gy ly60 strucK strucH
```

Source	SS	df	MS	Number of obs	=	77
Model	.004825571	3	.001608524	F(3, 73)	=	17.50
Residual	.006711089	73	.000091933	Prob > F	=	0.0000
				R-squared	=	0.4183
				Adj R-squared	=	0.3944
Total	.011536659	76	.000151798	Root MSE	=	.00959

gy	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ly60	-.0074633	.0016806	-4.44	0.000	-.0108126	-.0041139
strucK	.0123714	.0026831	4.61	0.000	.0070239	.0177188
strucH	.0053726	.0027139	1.98	0.052	-.0000362	.0107814
_cons	.0028857	.0028927	1.00	0.322	-.0028795	.0086509

Restriktede model:

```
regress gy ly60
```

Source	SS	df	MS	Number of obs	=	77
Model	2.1268e-06	1	2.1268e-06	F(1, 75)	=	0.01
Residual	.011534533	75	.000153794	Prob > F	=	0.9067
				R-squared	=	0.0002
				Adj R-squared	=	-0.0131
Total	.011536659	76	.000151798	Root MSE	=	.0124

gy	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
ly60	.0001604	.0013641	0.12	0.907	-.0025571	.0028779
_cons	.0164728	.0025477	6.47	0.000	.0113975	.0215481

5) Stata - opsamling

- ii. Let STATA calculate the F -test. This can be done via the **test** command following **regress**. The following code may be of help:

```
regress gy ly60 strucK strucH  
test strucK=strucH=0
```

- Alternativt kan bruge Stata:

```
regress gy ly60 strucK strucH  
test strucK=strucH=0
```

- Hvilket giver os samme resultat som ovenfor:

```
F( 2, 73) = 26.23  
Prob > F = 0.0000
```

- Med andre ord betyder det at vi forkaster at $\beta_2 = \beta_3 = 0$ hvilket betyder at en eller begge de strukturelle variable (investeringsrater i hhv. fysisk og humankapital) forklarer en signifikant del af variationen i de gennemsnitlige vækstrater og derfor ikke bør udelades. Det taler altså til fordel for hypotesen om betinget konvergens.

5) Stata - opsamling

- **MATA** er et program i Stata hvor vi eksplicit kan regne med matricer
- I øvelsen her bruger vi det til at regne efter og se om vi kan finde frem til samme estimerer og standardfejl som Stata giver os når vi bruger regress.

i. The first step is to call **mata** in STATA:

```
mata  
end
```

Når man kører MATA kode er det vigtigt at vi altid kører hele blokken dvs. fra *mata* til *end* på samme tid for at undgå fejl.

5) Stata - opsamling

ii. The next step is to read the relevant data into matrices.

```
mata
// INPUT DATA //
y = st_data(., "gy")
x = st_data(., "ly60 strucK strucH")
x
end
```

Vi gemmer data i to matricer:

- `st_data(a,b)` loader data fra datasættet i Stata, og gemmer det som en matrix
- Første argument (a) angiver rækker. Ved at skrive "." tager vi alle rækker med (alle n observationer)
- Andet argument (b) angiver kolonner. Disse angives ved "var1" og "var1 var2 var3"

Ønsker vi at se vores matricer kan vi blot kalde dem ved at skrive deres navn:

x		1	2	3
1	-1.19209e-07	.7348176837	.2361927629	
2	-.8596246243	.9025245905	.3871903419	
3	-.2726309001	1.179592371	.173224315	

5) Stata - opsamlina

iii. The third step is to create a vector corresponding to the constant term. This is done by creating a matrix of ones:

- Først laver vi en vektor der kun består af 1-taller

```
mata
// INPUT DATA //
y = st_data(., "gy")
x = st_data(., "ly60 strucK strucH")
x

// DEFINE CONSTANT AND X-MATRIX //
cons = J(rows(x), 1, 1)
end
```

- $J(a,b,c)$ danner en matrix af dimension $a \times b$ hvor alle rækker/kolonner fyldes med konstanten angivet som c

5) Stata - opsamling

iv. The matrix of all regressors (including intercept) is created as follows:

- Så merger vi denne vector med vores matrix x, og får derved vores velkendte X-matrix som har dimensionerne $(n \times k + 1)$ og består af et konstantled og de forklarende variable

```

mata
// INPUT DATA //
y = st_data(., "gy")
x = st_data(., "ly60 strucK strucH")
x

// DEFINE CONSTANT AND X-MATRIX //
cons = J(rows(x), 1, 1)
X = (x, cons)
X
end

```

- (x, z) merger to matricer (kræver at dimensioner passer)

	X			
	1	2	3	4
1	-1.19209e-07	.7348176837	.2361927629	1
2	-.8596246243	.9025245905	.3871903419	1
3	-.2726309001	1.179592371	.173224315	1

5) Stata - opsamling

v. Calculate the OLS estimator, $\beta = (X'X)^{-1}X'y$,

```

mata
// INPUT DATA //
y = st_data(., "gy")
x = st_data(., "ly60 strucK strucH")
x

// DEFINE CONSTANT AND X-MATRIX //
cons = J(rows(x), 1, 1)
X = (x, cons)
X

// CALCULATE OLS ESTIMATES //
beta_hat = (invsym(X' * X)) * (X' * y)
beta_hat
end

```

beta_hat

1

1	-.0074632722
2	.0123713856
3	.0053725833
4	.0028856984

- Matricer ganges sammen ved brug af * (kræver at dimensioner passer)
- Matricer transponeres ved brug af '
- Matricer inverteres ved brug af invsym(a) hvor a angiver matrixen (kræver den ikke er singular)

5) Stata - opsamling

vi. Calculate the OLS standard errors using $var(\hat{\beta}|X) = \hat{\sigma}^2(X'X)^{-1}$.

- Først finder vi residualerne ved:

$$\hat{u} = y - X\hat{\beta}$$

- Så beregner vi et estimat for variansen af fejlledet:

$$\hat{\sigma}^2 = \frac{\hat{u}'\hat{u}}{n - k - 1}$$

- Så kan vi beregne variansen for vores OLS estimat som under MLR.5 er:

$$var(\hat{\beta}|X) = \hat{\sigma}^2(X'X)^{-1}$$

- $var(\hat{\beta}|X)$ er en $(k + 1) \times (k + 1)$ matrix. Elementerne i diagonalen angiver variansen for hvert estimat, mens elementerne udenfor diagonalen angiver kovarianserne.
- Vi kan derfor finde standardfejlen for hvert estimat ved at tage kvadratroden af diagonalen

```
mata
// INPUT DATA //
y = st_data(., "gy")
x = st_data(., "ly60 strucK strucH")
x

// DEFINE CONSTANT AND X-MATRIX //
cons = J(rows(x), 1, 1)
X = (x, cons)
X

// CALCULATE OLS ESTIMATES //
beta_hat=(invsym(X'*X))*(X'*y)

// CALCULATE OLS STD ERRORS //
u_hat=y-X*beta_hat
s2=(1/(rows(X)-cols(X)))*(u_hat'*u_hat)
V_ols=s2*invsym(X'*X)
se_ols=sqrt(diagonal(V_ols))

end
```


5) Stata - opsamling

vii. Calculate t test statistics and print all results

- Da vi nu har beregnet estimer og standardfejl kan beregne t -statistikken for hypotesen at populations parametrene er lig 0:

$$t = \frac{\hat{\beta} - 0}{se(\hat{\beta})}$$

- Nedenstående tabel viser estimer, standardfejl og t -statistikker for konstantleddet samt de tre variable. Resultaterne stemmer overens med resultaterne fra tidligere.

(beta_hat, se_ols, ttest)

	1	2	3
1	-.0074632722	.0016805604	-4.440942606
2	.0123713856	.0026831221	4.610817164
3	.0053725833	.0027139005	1.979653739
4	.0028856984	.0028927061	.9975774649

```

mata
// INPUT DATA //
y = st_data(., "gy")
x = st_data(., "ly60 strucK strucH")
x

// DEFINE CONSTANT AND X-MATRIX //
cons = J(rows(x), 1, 1)
X = (x, cons)
X

// CALCULATE OLS ESTIMATES //
beta_hat=(invsym(X'*X))*(X'*y)

// CALCULATE OLS STD ERRORS //
u_hat=y-X*beta_hat
s2=(1/(rows(X)-cols(X)))*(u_hat'*u_hat)
V_ols=s2*invsym(X'*X)
se_ols=sqrt(diagonal(V_ols))

// CALCULATE T TEST STATISTICS //
ttest=beta_hat:/se_ols

// PRINT RESULTS //
(beta_hat, se_ols, ttest)
end

```