






















Základy práce s ekonometrickým programom EViews

Martin Lukáčik, Adriana Lukáčiková, Karol Szomolányi

Program **EViews** je navrhnutý pomocou koncepcie objektov. Objekty sú súhrny informácií a operácií, ktoré sú spojené spolu v jednoducho použiteľnej jednotke. Celá práca v **EViews** zahŕňa manipuláciu s rôznymi objektmi. Najdôležitejší objekt v **EViews** je pracovný zošit a prvé kroky pri každej úlohe vytvárajú nový alebo zavádzajú existujúci pracovný zošit do pamäti.

Základne objekty v EViews sú pracovný zošit, časový rad, a rovnica. Sú tu aj iné objekty, ktoré slúžia pre špeciálne úlohy. Zoznam **EViews** objektov obsahuje: vektor koeficientov, databázy, rovnice, grafy, skupiny, modely, pool (časové rady/prierezové dáta), vzorky, časové rady, systémy, matice, tabuľky, text, vektorovú autoregresiu, vektory/riadkové, a vektor skalárov. Všetky objekty okrem pracovného zošita a databázy majú vlastné ikony, ktoré sú zobrazené v okne pracovného zošita (v prípade zapnutého filtra sú zobrazené zvolené typy objektov). Keď je vytvorený nový pracovný zošit, v pracovnom zošite sú zobrazené počiatočné dva objekty vektor koeficientov (samé nuly) a rad reziduálov (vyplnený zápisom NA).

 Alpha	 Model	 Sym
 Coefficient Vector	 Pool	 System
 Equation	 Rowvector	 Table
 Graph	 Sample	 Text
 Group	 Scalar	 Valmap
 Logl	 Series	 VAR
 Mantx	 Sspace	 Vector

Vytvorenie objektu: vyberte **Object/New object** zo základného menu alebo ponuky okna pracovného zošita, kliknite na typ objektu, ktorý chcete vytvoriť, pomenujte a kliknite na **OK**. Pre niektoré typy sa otvorí iné dialógové okno, kde treba vybrať z možností. Pre väčšinu objektov sa výsledné okno otvorí hneď. Pracovný zošit musí byť otvorený pred vytvorením objektu. Proces tvorby pracovného zošita a objektov budú vysvetlené neskôr.

Matematické výrazy

EViews obsahuje rozsiahlu množinu zabudovaných operátorov, ktorá dovolí uskutočňovať veľmi jednoducho matematické operácie s dátami. Mimo podporované štandardné matematické a štatistické operácie **EViews** poskytuje niekoľko špecializovaných funkcií pre automatické spracovanie diferencií a percentuálnych rozdielov, ktoré pomáhajú hľadať v časových radoch dôležité informácie. Všetky operátory popísané nižšie môžu byť použité v časových radoch aj pre skalárne hodnoty. Keď sú aplikované na časový rad, operácia je vykonaná pre každé pozorovanie v aktuálnom výbere.

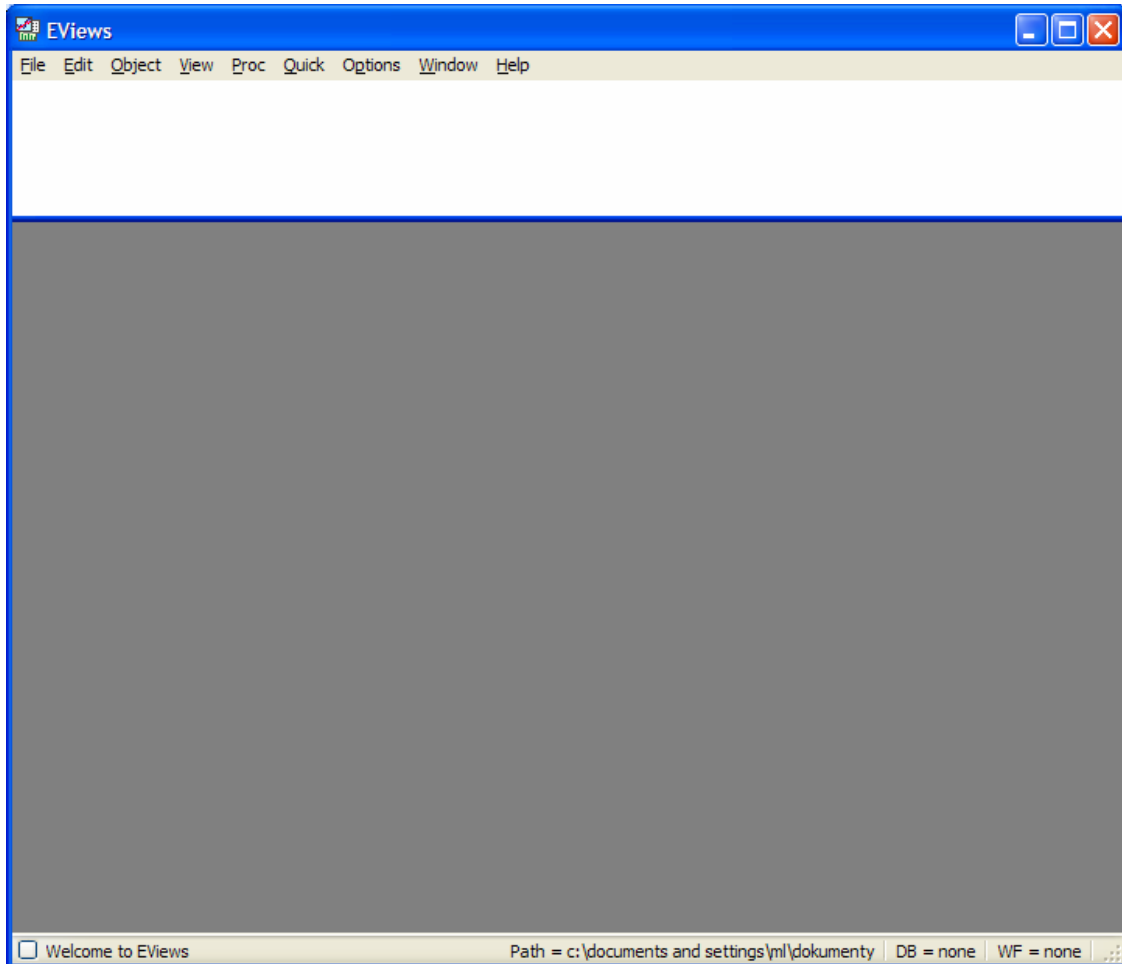
EViews dodržiava obvyklé poradie matematických operátorov zľava doprava s prioritou operátorov podľa úrovne od prvej až po piatu

1. ^
2. *, /
3. +, odčítanie (-)
4. <, >, <=, >=, =
5. a, alebo

Operátory:

+ sčítanie $x+y$, - odčítanie $x-y$, * násobenie $x*y$, / delenie x/y , ^ umocnenie x^y
> väčší ako - vráti hodnotu 1 ak $x>y$, inak vráti 0, < menší ako - vráti hodnotu 1 ak $x<y$, inak 0
= rovná sa, <> nerovná sa, <= menší alebo rovný, >= väčší alebo rovný vráti 1 alebo 0
and logické a, **or** logické alebo.

Popis EViews hlavného okna



Titulný pás: Titulný pás s názvom programu **EViews** je na vrchole hlavného okna – posun okna.

Základné menu: Pod titulným pásom je základné menu. Ak premiestnite ukazovateľ k zvolenej hodnote a kliknete na ľavé tlačidlo myši, objaví sa otvorená ponuka. V ponukách sa dajú zadať iba čierne položky, zatiaľ čo šedé položky nie sú práve použiteľné.

Riadiace okno: Pod pásom menu je oblasť nazvaná riadiace okno, kde sa zapisujú **EViews** príkazy. Príkaz je splnený okamžite po stlačení klávesy **Enter**.

Stavový riadok: Na spodku okna je stavový riadok, ktorý je rozdelený na štyri sekcie. Ľavá časť niekedy zobrazuje stavové hlásenia. Susedná časť ukazuje implicitný adresár, ktorý **EViews** používa na hľadanie súborov s údajmi. Posledné dve sekcie zobrazujú názvy implicitnej databázy a pracovného zošita.

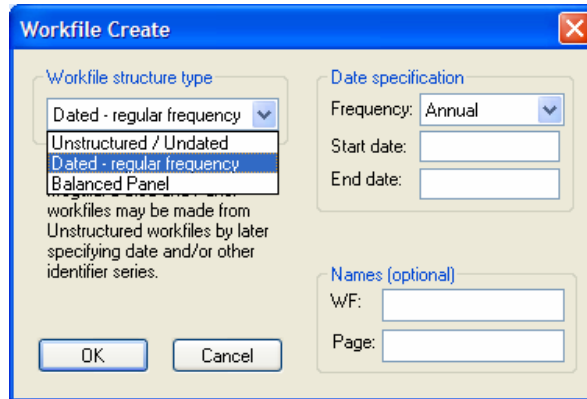
Pracovná plocha: Oblasť uprostred okna je pracovná plocha, kde **EViews** zobrazuje rôzne výsledné okná, ktoré vytvárate pri práci. Tieto okná sa podobajú na hárky uložené na stole, keď

pracujete. Okná sa prekrývajú, pričom najvrchnejšie okno je posledne spustené, a to sa snaží zobraziť vždy do stredu, alebo aktívne okno, ktoré ste si zvolili z už otvorených okien. Len aktívne okno má stmavnutý horný pás.

Jednoduchý príklad regresie

a) Vytvorenie pracovného zošita

Vyberte z ponuky základného menu **File/New/Workfile**



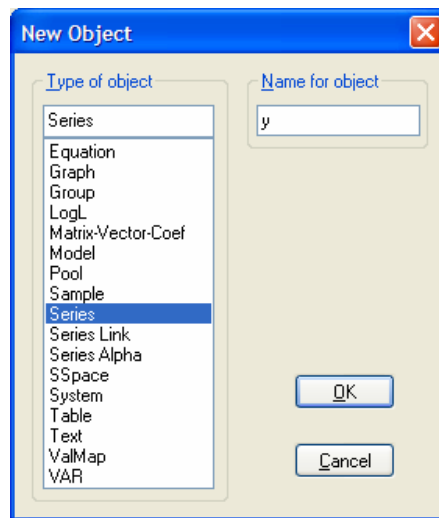
Na výber sú ponúknuté tri štruktúry dát: *neštruktúrované údaje* (dodatočne sa dajú previesť na ľubovoľný iný typ), *na časové rady s pravidelne opakovanou frekvenciou* (ročné, mesačné, štvrťročné, ...) pre ktoré je potrebné špecifikovať obdobie pre prvý (napríklad 1993Q1) a posledný údaj (dodatočne sa dajú pridať alebo ubrať pozorovania) a *vybilancované panelové dáta* (rovnaký počet časových radov pre každý prierezový údaj). Pomenovanie pracovného zošita je dôležité až pri uložení súboru na disk (a pomenovanie hárku - stránky iba pre prípad využitia viacerých hárkov).

Po vyplnení komunikačného okna zadáme **Ok**. Na pracovnej ploche sa otvorí okno pracovného zošita (ak ste nezadali meno, tak Untitled) s dvoma objektmi: vektorom koeficientov s rezervovaným pomenovaním "c" a rad reziduálov s rezervovaným pomenovaním "resid". Pod ponukou pracovného zošita sa nachádza informácia o dĺžke radov **Range** a výberu **Sample** s ktorou sa uskutočňujú odhady, kliknutím na túto informáciu sa dá táto dĺžka upraviť.

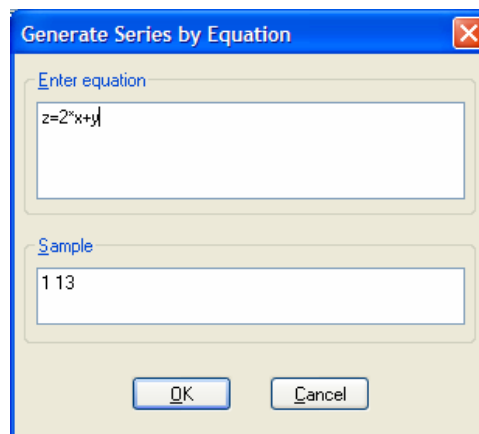
Ak by ste ukončili prácu bez uloženia pracovného zošita na disk, vytvorené informácie by sa stratili, preto je vhodné vytvorený pracovný zošit (rovnako aj po každej významnej zmene v ľubovoľnom objekte) uložiť pomocou základného menu **File/Save** respektíve **Save** z okna pracovného zošita. Pričom prvýkrát sa zadáva meno a umiestnenie (ako ponuka **File/Save as**).

b) Vloženie údajov do pracovného zošita

Pre vytvorenie nového časového radu v menu zvolíte **Object/New object/Series**, pomenujte rad a kliknite na **Ok**. V pracovnom zošite sa vytvorí nový objekt so zadaným menom. Dvojklikom na meno radu ho otvoríme v novom okne. Obsahuje hodnoty "NA". Po vytvorení je rad v móde prezerania. Na zadávanie resp. upravovanie existujúcich hodnôt ho musíme prepnúť tlačidlom **Edit+/-** v okne objektu. Po zadaní (po jednej hodnote resp. nakopírovaním z tabuľkového procesora Excel – *pričom desiatinným oddeľovačom je bodka*) nakoniec opäť stlačte Edit+/-, tým sa hodnoty zadajú. Nezabúdajte si prácu uložiť do súboru pomocou Save.



Ak chcete vytvoriť nový rad z už zadaných radov, napríklad dvojnásobné hodnoty pôvodného radu x plus rad y a pomenovať ho napríklad z , zadajte **Quick/Generate Series** z hlavného menu alebo **Genr** z menu pracovného zošita a zadajte výraz ako rovnicu do okna. Potvrdením **Ok** vytvoríte nový rad.



c) Vytvorenie skupiny (Group) z preskúvaných radov

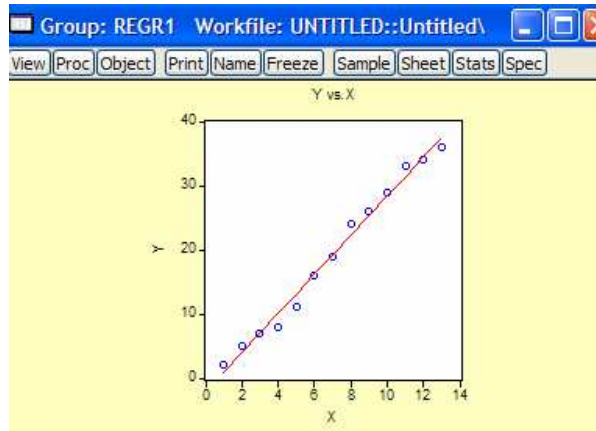
Predpokladajme, že máte vytvorené dva rady. Jeden rad tvorí závislá premenná y a druhý rad nezávislá premenná x . (Mená objektov - radov voľte krátke jednoslovné ale zrozumiteľné.) Označte najprv nezávislú premennú (1 klik v pracovnom zošite) stlačte a držte **Ctrl** a označte závislú premennú. (Ak tvoríte skupinu z viac ako dvoch radov, tiež dávajte najskôr nezávislé premenné.)

Po označení všetkých radov, ktoré chcete do skupiny vyberte **Show** z menu a **Ok** resp. pravé tlačidlo myši na vyznačenú modrú farbu radov a zvolte **Open as Group**. Otvorí sa nepomenovaný objekt Group. Meno nepomenovanej skupine zadáte kliknutím na **Name** pod ktorým sa ukáže v pracovnom zošite, ale do súboru sa zapíše až po zadaní **Save**.

d) Zobrazenie závislej premennej oproti nezávislej na grafe

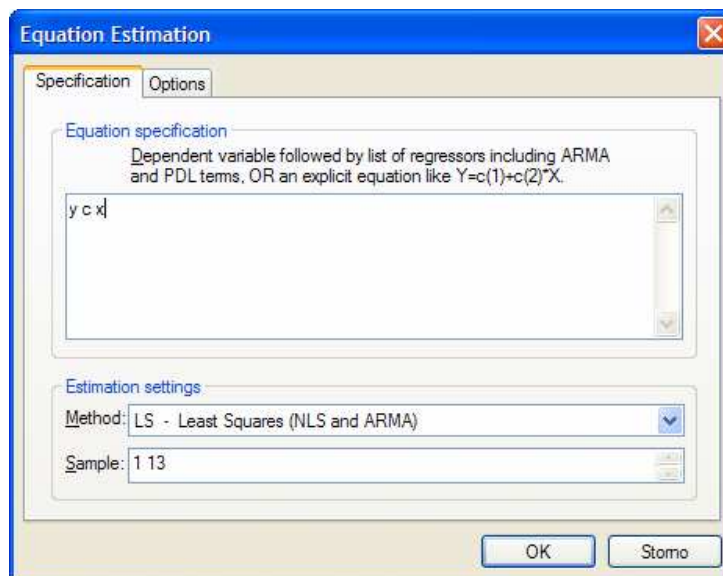
Dvojklikom na jej meno otvorte skupinu s radmi, kde prvá je nezávislá premenná. Vyberte z menu **View/Graph/Scatter/Scatter with Regression**. Na základe grafu sa rozhodnete, či je

medzi radmi významná závislosť. (Na obrázku je a silne pozitívna.) Ak chcete upraviť vlastnosti zobrazovaného grafu, tie sa menia v menu **Object/View Options/Options**.

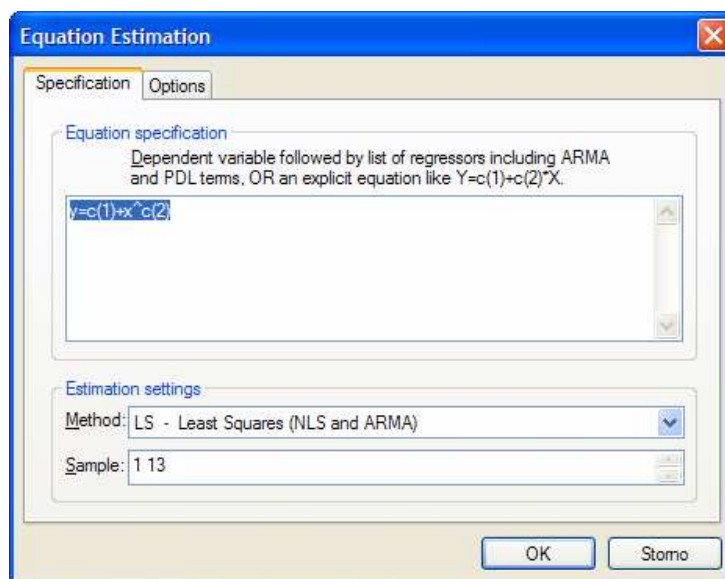


e) Odhad regresných koeficientov metódou najmenších štvorcov

Pre odhad koeficientov rovnice vyberáme z menu **Object/New object/Equation**, kde si rovnicu môžeme hneď pomenovať alebo jej meno zadáme až po zobrazení výsledku cez **Name**. Výsledok s odhadnutými koeficientami a sprievodnými štatistikami sa zobrazí po špecifikácii tvaru rovnice. *Lineárny model* stačí zadať vymenovaním premenných (oddelených medzerami), pričom prvá je vždy závislá premenná. Ak chceme odhadovať model aj s absolútnym členom do zoznamu pridáme aj rezervované písmenko "c". Pri špecifikácii si môžeme zvoliť metódu odhadu (LS je metóda najmenších štvorcov) a počet období, z ktorých robíme odhad – Sample, ten je nastavený podľa hodnoty pre celý pracovný zošit, ale dá sa pre každý odhad zmeniť.



Nelineárny model musíme zadať presným zápisom rovnice pričom pre odhadované parametre sú rezervované označenia postupne c(1), c(2), c(3), atď. Pozor c(0) nie je dovolené.



Uloženie výsledku do súboru sa realizuje až po zadaní **Save**.

Obsah okna objektu rovnica

Výsledkom odhadu sú nielen parametre, ale aj množstvo sprievodných štatistík.

Dependent Variable: Y				
Method: Least Squares				
Date: 10/25/05 Time: 08:53				
Sample: 1 13				
Included observations: 13				
Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-2.076923	0.792457	-2.620867	0.0238
X	3.043956	0.099840	30.48830	0.0000
R-squared	0.988305	Mean dependent var	19.23077	
Adjusted R-squared	0.987241	S.D. dependent var	11.92444	
S.E. of regression	1.346917	Akaike info criterion	3.574152	
Sum squared resid	19.95604	Schwarz criterion	3.661067	
Log likelihood	-21.23199	F-statistic	929.5363	
Durbin-Watson stat	1.012175	Prob(F-statistic)	0.000000	

Prvý riadok obsahuje názov závislej premennej.

Druhý riadok obsahuje názov metódy použitej pre odhad parametrov.

Tretí riadok obsahuje dátum a čas uskutočnenia regresie.

Štvrtý riadok obsahuje veľkosť výberu regresie.

Piaty riadok obsahuje počet zahrnutých pozorovaní (keď je menší ako výber, chýba pozorovanie)

Pre nelineárne modely rátané iteračne sa zobrazuje aj počet iterácií, kým nedošlo ku konvergencii hodnôt a takisto aj odhadovaný tvar rovnice. Pre iné metódy odhadu sa zobrazujú aj ďalšie informácie napríklad zoznam inštrumentov alebo typ použitého jadra (kernel) atď.

Výsledky odhadu parametrov rovnice sú zvýraznené žltou farbou

Prvý stĺpec identifikuje každú premennú, c označuje konštantu. Druhý stĺpec obsahuje hodnoty odhadnutých parametrov. V treťom stĺpci nájdeme odhady štandardných odchýlok parametrov. Štvrtý stĺpec obsahuje hodnoty t-štatistík a piaty pravdepodobnosť s ktorou platí nulová hypotéza o nevýznamnosti parametra.

Súhrnné štatistiky sa nachádzajú pod tabuľkou odhadnutých hodnôt parametrov.

R squared = R^2 – koeficient determinácie ukazuje podiel variancie závislej premennej vysvetlenej nezávislými premennými

Adjusted R squared = korigovaný koeficient determinácie (upravený o stupne voľnosti)

$S.E.$ of regression = štandardná odchýlka regresie

Sum squared resid = suma štvorcov reziduálov, metóda najmenších štvorcov ju minimalizuje

Log likelihood = hodnota logaritmu vierohodnosti

Durbin-Watson stat = štatistika testujúca autokoreláciu reziduálov

Mean dependent var = priemer závislej premennej

$S.D.$ dependent var = štandardná odchýlka závislej premennej

Akaike info criterion = informačné kritérium slúžiace pri výbere z viacerých modelov

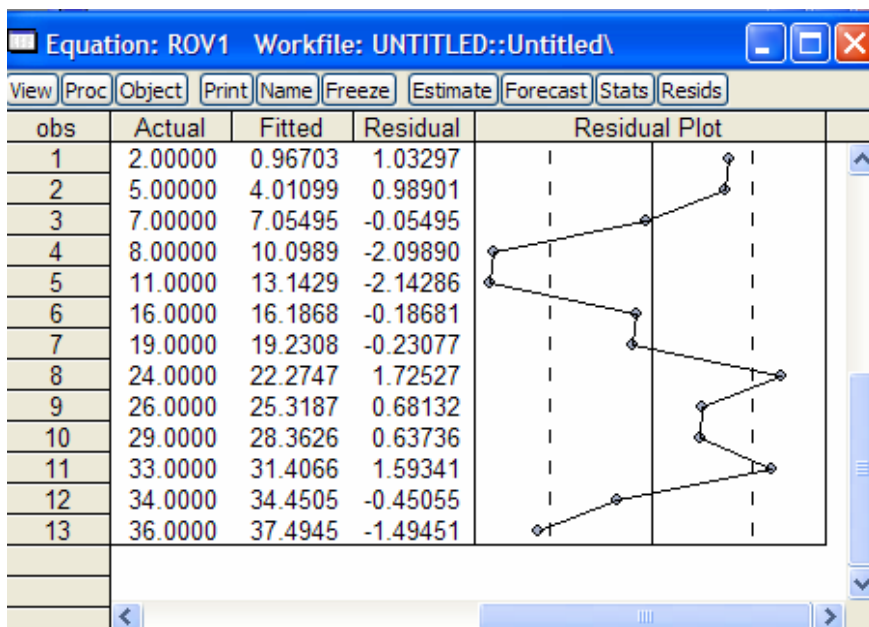
Schwarz criterion = ďalšie informačné kritérium (lepšia hodnota je menšia u oboch)

F -statistic = test hypotézy, že všetky parametre okrem konštanty sú nevýznamné

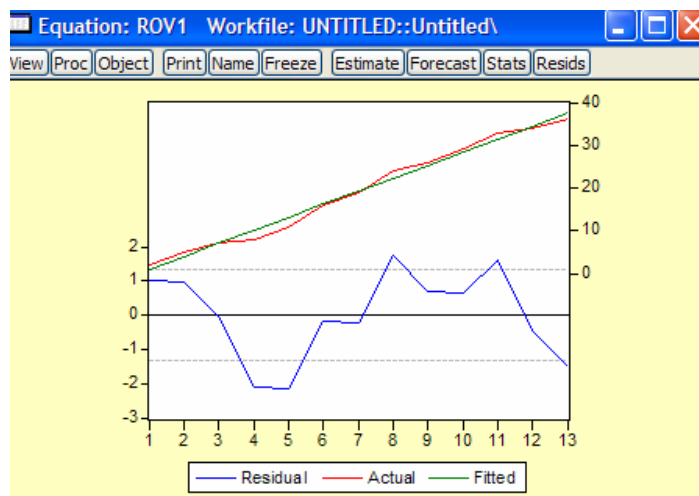
Prob(F -statistic) = pravdepodobnosť s ktorou platí nulová hypotéza o nevýznamnosti modelu

Nelineárny model neobsahuje informáciu o F štatistike a jej pravdepodobnostnej hodnote, lebo nemá v nich žiadnu výpovednú hodnotu. Pri iných typoch odhadov sa vypočítavajú aj ďalšie sprievodné štatistiky.

Zobrazenie skutočnej hodnoty závislej premennej, hodnoty závislej premennej vyrovnanej modelom a hodnoty reziduálu pre všetky pozorovania získame v okne rovnice cez menu **View/Actual,Fitted,Residual/Actual,Fitted,Residual Table**.



Tieto hodnoty môžeme zobrazit' aj pomocou grafu stlačením tlačidla **Resids**.



Späť k odhadnutým hodnotám sa vrátite stlačením **Stats**. Ak chcete upraviť odhadovaný tvar modelu, teda zmeniť jeho špecifikáciu, stlačte **Estimate** a zmeňte tvar rovnice či zoznam premenných. Menu **View** poskytuje množstvo štatistík pre testy koeficientov respektíve testy skúmajúce reziduály ako náhrady za náhodné zložky.

Ak chcete samostatne pracovať s reziduálmi, môžete si z nich vytvoriť nový rad pomocou menu **Proc/Make Residual Series ...**, kde zadáte meno tohto radu a potvrdíte Ok.

Ak chcete pracovať s vyrovnanými hodnotami, najprv si vytvorte rad reziduálov a následne vygenerujte rad [podľa návodu vytvorenia nového radu z už existujúceho radu](#) ako rozdiel skutočných hodnôt závislej premennej a reziduálov.

Zobrazenie tvaru odhadnutého modelu

Aby sa nám zobrazil tvar odhadnutého modelu, vyberáme z menu **View/Representations**. Zobrazia sa tri zápisy. *Substituted Coefficients* je zápis s odhadnutými hodnotami, *Estimation Equation* je všeobecný tvar modelu a *Estimation Command* je zápis príkazu, ktorý vykoná tento odhad, ak by sme ho zadali do Riadiaceho okna (tj. nemusíme ísť cez menu, ale môžeme priamo zadávať príkazy v riadiacom okne – možnosť programovania príkazov.)

```

Equation: ROV1  Workfile: UNTITLED::Untitled\
View Proc Object Print Name Freeze Estimate Forecast Stats Resids
Estimation Command:
=====
LS Y C X

Estimation Equation:
=====
Y = C(1) + C(2)*X

Substituted Coefficients:
=====
Y = -2.076923077 + 3.043956044*X

```


Testovanie hypotéz a konfidenčné intervaly

Všetky dôležité údaje potrebné pre testovanie hypotéz [sa nachádzajú v obsahu okna rovnica](#) – vyznačenom žltou farbou. Hodnoty t -štatistik sú v štvrtom stĺpci vypočítané ako podiel hodnôt odhadnutých parametrov v druhom stĺpci a odhadov ich štandardných odchýliek z tretieho stĺpca.

Pre vysvetlenie si práce so základnými vypočítanými testovacími hodnotami si vytvoríme pomocný vektor do ktorého si postupne uložíme všetky hodnoty. Predpokladajme, že skúmame regresiu, kde závislou premennou je y a nezávislými premennými sú x_1 a x_2 . Rovnicu odhadujúcu takýto model s konštantou ste pomenovali **rov1** (zadaná v poradí y a x_1 x_2).

Postup:

1. Vytvoríme si vektor so šiestimi prvkami s názvom **hyp**. Do riadiaceho okna zapíšete príkaz **vector(6) hyp** a zadajte **Enter**. V pracovnom zošíte pribudne šesťmiestny nulový vektor.
2. Pre určenie kritickej hodnoty t -štatistiky obojstranného testu na 5% hladine významnosti zapíšeme do riadiaceho okna príkaz **hyp(1)=@qtdist(.975,(rov1.@regobs-rov1.@ncoef))** a zadajte **Enter**. Do prvého prvku vektora **hyp** sa zapíše vypočítaná kritická hodnota. Výpočet využíva zabudované **klúčové slová** (tie začínajú @) **qtdist**, ktoré vráti po špecifikovaní dvoch hodnôt t -štatistiku. Tie hodnoty sú $(1-\alpha/2)=0,975$ – vyrátaná z hladiny významnosti α a počet stupňov voľnosti $(n-(k+1))$ určený z kľúčových slov **regobs** (počet pozorovaní) a **ncoef** (počet parametrov vrátane konštanty), ktorých spôsob zápisu je zrejмый zo zadaného príkazu. Volajú sa pre pomenovanú rovnicu **rov1**.
3. Pre určenie kritickej hodnoty t -štatistiky jednostranného testu na 5% hladine významnosti zapíšeme do riadiaceho okna príkaz **hyp(2)=@qtdist(.95,(rov1.@regobs-rov1.@ncoef))** a zadajte **Enter**. Do druhého prvku vektora **hyp** sa zapíše vypočítaná kritická hodnota.
4. Ak testujeme významnosť parametra (obojstranný test), tak porovnáваме hodnotu zo štvrtého stĺpca okna rovnica príslušného parametra s prvou hodnotou vektora **hyp** (toto rozhodnutie je zřejмый z pravdepodobnostnej hodnoty z piateho stĺpca). Ak testujeme pozitívnosť efektu skúmanej premennej na závislú premennú (jednostranný test), tak porovnáваме hodnotu zo štvrtého stĺpca okna rovnica príslušného parametra s druhou hodnotou vektora **hyp**.
5. Ak chceme vypočítať dolnú hranicu 90% konfidenčného intervalu pre premennú x_1 (je vo výpise druhá v poradí, lebo **rov1** sa odhadovala y a x_1 x_2) zapíšeme do riadiaceho okna príkaz **hyp(3)=rov1.@coefs(2)-(@qtdist(.95,(rov1.@regobs-rov1.@ncoef)))*rov1.@stderrs(2)** a zadáme **Enter**. Vo výpočte pribudli dve nové kľúčové slová: **coefs(2)** označujúce hodnotu parametra, ktorý bol zapísaný ako 2. v poradí a **stderrs(2)** označujúci hodnotu štandardnej odchýlky parametra zapísaného tiež ako 2. v poradí (c je prvé, x_1 je druhé a x_2 je tretie).
6. Ak chceme vypočítať hornú hranicu 90% konfidenčného intervalu pre premennú x_1 (je vo výpise druhá v poradí, teda v tabuľke výpisu je v druhom riadku) zapíšeme príkaz **hyp(4)=rov1.@coefs(2)+(@qtdist(.95,(rov1.@regobs-rov1.@ncoef)))*rov1.@stderrs(2)** a zadáme **Enter**. Hodnota sa vypočíta do štvrtého prvku vektora **hyp**.
7. Aby sa zapamätala vypočítaná hodnota F štatistiky modelu v našom vektore, stačí zapísať príkaz **hyp(5)=rov1.@f** a zadať **Enter**.
8. Pre určenie kritickej hodnoty F štatistiky na 5% hladine významnosti zapíšeme do riadiaceho okna príkaz **hyp(6)=@qfdist(.95,rov1.@ncoef-1,rov1.@regobs-rov1.@ncoef)** a **Enter**.
9. Dvojklíkom na vektor **hyp** v pracovnom zošíte si môžeme pozrieť kedykoľvek ľubovoľnú vypočítanú hodnotu.

Časovo posunuté premenné, diferencie a logaritmy premenných

Práca s časovo posunutými premennými je veľmi jednoduchá. Posunutú premennú do odhadu rovnice zadáte, ak hneď za pôvodné meno premennej vložíte do zátvorky číslo vyjadrujúce o aké obdobie sa pri odhade chcete posunúť. Ak dozadu tak so znamienkom mínus. Teda ak predpokladáte, že spotreba v súčasnom období (*spot*) závisí od disponibilného príjmu (*dp*) a spotreby v predchádzajúcom období, tak zápis rovnice bude vyzerat' *spot c dp spot(-1)*. Ak chcete zadať viac posunov tej istej premennej, tak zoznam *spot spot(-1) spot(-2) spot(-3)* nahradíme jednoduchým *spot(0 to -3)*

Čo sa týka nastavenia použitého časového obdobia, tak to automaticky za vás riadi program. Pričom sa snaží použiť čo najvyšší počet pozorovaní z výberu (ktoré existujú).

Podobne jednoduchá je práca s diferenciami časových radov. Tie sa zadávajú pomocou písmena *d* za ktorým je v zátvorke špecifikovaná premenná a spôsob diferencie. Ak chcete zadať rozdiel susedných pozorovaní spotreby teda *spot-spot(-1)* tak stačí keď zapíšete *d(spot)*. Ak chcete zadať diferenciu diferencie, teda rozdiel už diferencovaného radu resp. diferenciu druhého stupňa zadáte *d(spot,2)*. Ak chcete zadať rozdiel pozorovaní spotreby a spotreby spred štyroch období teda *spot-spot(-4)*, tak zapíšete *d(spot,0,4)*. Prvé číslo je teda stupeň diferencie a druhé, o koľko období posunutú hodnotu odrátavam.

Logaritmy sú analógiou predchádzajúceho. Logaritmus spotreby bude *log(spot)* a rozdiel susedných logaritmov spotreby bude *dlog(spot)*. Rozdiel logaritmov spotreby spred štyroch období *log(spot)-log(spot(-4))* sa dá zapísať *dlog(spot,0,4)*.

Trend, sezónne a umelé premenné

Práca so sezónnymi premennými je možná iba v *časových radoch s pravidelne opakovanou frekvenciou* menšou ako rok. Teda v polročných, štvrťročných, mesačných atď. Ak chceme napr. pridať do odhadu sezónnu premennú pre štvrtý štvrťrok, tak medzi premenné zaradíme *@seas(4)*.

Ak chceme pridať časový trend, tak zadáme *@trend*. Táto premenná priradí prvému pozorovaniu nulu a ďalšie hodnoty zvyšuje o jedna. Teda druhé pozorovanie bude 1, tretie 2 atď. Ak chcete zadať trend, ktorý začne jednotkou v prvom pozorovaní, tak ak máte údaje v pracovnom zošite napr. od prvého štvrťroku 1993 (teda 1993Q1), tak zadáte trend s o 1 nižším obdobím, teda pridáte *@trend(1992Q4)*.

Vytvorenie umelej premennej, ktorá všetkým obdobiam okrem zvoleného priradí nuly a vybranému obdobiu 1 (používame pre elimináciu výkyvu), sa dá uskutočniť sekvenciou dvoch príkazov. Najprv vytvoríme rad núl, pričom ho pomenujeme tak, aby sme vedeli, pre ktoré obdobie zadávame umelú premennú. Teda zadáme **series u94q1=0** a stlačíme **Enter**. Teda vytvárame umelú premennú pre prvý štvrťrok roku 1994. Ale zatiaľ máme všade nuly. Druhý príkaz, ktorý zadáme je **u94q1(@dtoo("1994q1"))=1** a opäť stlačíme **Enter**. Tento príkaz zvolenému radu a v ňom konkrétnemu pozorovaniu, ktoré získame konverziou zadaného dátumu na poradie v rade, priradí hodnotu 1.

Ak chceme vytvoriť umelú premennú rozdeľujúcu pozorovania na dve množiny, kde faktor nepôsobí (budú 0) a kde faktor pôsobí (budú 1), tak to vytvoríme kombináciou trendovej a logickej premennej. Nech nastal zvrät v prvom štvrťroku roku 1999, tak po zadaní príkazu **series u99q1end=(@trend(1998Q4)>0)** a **Enter** sa vytvorí požadovaný rad *u99q1end*. Logické premenné vrátia nulu, ak výraz je nepravda a jednotku, ak výraz platí.

Autokorelácia

Ak chceme zobrazíť na grafe reziduály, aby sme detekovali autokoreláciu, najprv [si vytvoríme rad reziduálov](#). Aby sme vedeli o akú premennú ide, pomenovanie volíme napríklad ey pre rad reziduálov z odhadu y . A následne si ho môžeme prezrieť v okne časového radu ey pomocou **View/Line Graph**. Ak nechceme vytvárať reziduály, tento graf sa dá otvoriť priamo v okne rovnice pomocou **View/Actual, Fitted, Residual/Residual Graph**. Výhodou prvého spôsobu je, že si takto môžeme priamo odhadnúť inicializačnú hodnotu koeficientu autokorelácie z regresie ey od $ey(-1)$, teda bez absolútneho člena.

Testovanie prítomnosti autokorelácie je možné realizovať štandardne vypočítavanou Durbinovou-Watsonovou štatistikou. Alebo *Godfreyovým-Breuschovým testom* autokorelácie, čo je test Lagrangeových multiplikátorov, ktorý má χ^2 rozdelenie s p stupňami voľnosti a jeho štatistika je definovaná vzťahom

$$GB = n \cdot R^2$$

kde R^2 je koeficient determinácie pomocnej regresie reziduálov metódy najmenších štvorcov na vysvetľujúcich premenných a postupne v čase posunutých reziduáloch (chýbajúce hodnoty sú nahradené nulami). Nulová hypotéza predpokladá neexistenciu prvých p autokorelácií.

Godfreyov-Breuschov test na rozdiel od Durbinovho-Watsonovho testu je nielen testom autokorelácie prvého rádu, ale súčasným testom prvých p autokorelácií a dá sa použiť aj pre modely, kde vysvetľujúcou premennou je aj posunutá vysvetľovaná premenná. Eviews ho vypočítava v menu **View/Residual Tests/Serial Correlation LM Test...**, kde si na začiatku zvolíte rád testovania p .

Eviews na odstránenie autokorelácie prvého rádu používa zabudovaný člen **AR(1)** (a analogicky aj pre vyššie rády **AR(p)**), ktorý pridáme do rovnice, ktorú odhadujeme. Výsledkom je, že parameter pri **AR(1)** je koeficientom autokorelácie zo všetkými zodpovedajúcimi štatistikami. Následne môžeme skúmať autokorelačnú a parciálnu autokorelačnú funkciu cez ponuku **View/ARMA Structure...**, kde špecifikujeme **Correlogram**. Táto metóda je nelineárna.

Ak chceme problém riešiť *Cochraneovou-Orcuttovou metódou*, ktorá využíva iba lineárnu metódu najmenších štvorcov, musíme zabezpečiť vykonávanie jednotlivých iterácií tejto metódy. Postup tejto metódy:

MNŠ sa odhadnú parametre pôvodného modelu.

1. Z reziduálov sa sformuluje autoregresný model $e_t = \rho e_{t-1} + \varepsilon_t$, z ktorého sa MNŠ odhadne $\hat{\rho}$.

2. Pomocou $\hat{\rho}$ sa vypočítajú zovšeobecnené diferencie a vytvorí sa transformovaný model

$$(y - \rho * y(-1)) = \beta_0 + \beta_1 * (x1 - \rho * x1(-1)) + \beta_2 * (x2 - \rho * x2(-1)).$$

3. MNŠ sa odhadnú parametre transformovaného modelu a pokračuje sa bodom 1.

Tento iteračný proces pokračuje, kým sa nesplní kritérium konvergenencie, napr. $|\hat{\rho}_r - \hat{\rho}_{r-1}| \leq \delta$.

Ako to zabezpečiť v Eviews?

Predpokladajme, že pôvodný odhad mal tvar y c $x1$ $x2$ a z neho vytvoríme rad reziduálov e .

1. Po vytvorení reziduálov e , odhadneme model e $e(-1)$ a túto rovnicu si pomenujeme **rove**.

2. Pre výpočet zovšeobecnených diferencií vytvoríme rovnicu s menom **rovzd** cez ponuku **Objects/New Object/Equation**, kde zadáme tvar:

$$y - rove.@coefs(1)*y(-1) \text{ c } x1 - rove.@coefs(1)*x1(-1) \text{ x2} - rove.@coefs(1)*x2(-1)$$

3. Prepíšeme pôvodné reziduály e novovytvoreným radom reziduálov pomocou príkazu **series e=y-(rovzd.@coefs(1)+rovzd.@coefs(2)*x1+rovzd.@coefs(3)*x2** a zadáme **Enter**.
4. Opakovane zadávavame kroky 1, 2 a 3 v uvedenom poradí, kým sa odhadnuté ρ , teda *rove.@coefs(1)* nemení na viac ako treťom desatinnom mieste. (Opakovanie kroku 1 a 2 je najjednoduchšie, ak v okne fóvnicke *rove* a následne *rovzd* zadáme **Estimate**. Opakovanie kroku 3 je najjednoduchšie, ak kliknete do riadku príkazu a zadáte **Enter**.)
5. Na konci nezabudnite previesť parameter absolútneho člena β_0 na pôvodný parameter β_0 pomocou príkazu **scalar beta0=rovzd.@coefs(1)/(1-rove.@coefs(1))**.

Heteroskedasticita

Ak chceme zobrazit' na grafe reziduály, aby sme detekovali heteroskedasticitu, najprv [si vytvoríme rad reziduálov](#). Následne skúmame priebeh reziduálov oproti premennej, ktorú podozrievame, že spôsobuje heteroskedasticitu (respektíve voči každej vysvetľujúcej premennej). Pripomínáme, že vytvoríme **Group**, kde prvá je označovaná podozrievaná premenná a druhý je označený reziduál a následne skúmame **View/Graph/Scatter/Simple Scatter**.

Všeobecný test nesplnenia predpokladu o konštantnosti rozptylov navrhol White. *Whiteov test* nepomáha určiť príčinu ani spôsob odstránenia heteroskedasticity, je určený len na jej detekciu. Test porovnáva variančno-kovariančnú maticu odhadu získaného klasickou metódou najmenších štvorcov a zovšeobecnenou metódou najmenších štvorcov.

Pri testovaní sa využíva pomocná regresia štvorcov reziduálov zo skúmaného modelu na všetkých vysvetľujúcich premenných v modeli (vrátane konštanty), na ich druhých mocninách a vzájomných párových súčinoch vysvetľujúcich premenných (respektíve bez nich). Z pomocnej regresie sa vypočíta testovacia štatistika, ktorá má χ^2 rozdelenie s k stupňami voľnosti a je definovaná vzťahom

$$W = n.R^2$$

kde k je počet premenných v pomocnej regresii (pozor bez konštanty) a R^2 je koeficient determinácie pomocnej regresie. Eviews tento test (rovnako ako ekvivalent bez párových súčinov) vypočítava v menu **View/Residual Tests/White Heteroscedasticity (cross term)**.

Parkom navrhnutý postup testovania predpokladá testovanie heteroskedasticity závislej od konkrétnej premennej x (premennej, ktorú podozrievame) vo všeobecnom tvare

$$\sigma_i^2 = cx_i^\alpha e^{v_i} \quad \text{kde } \ln e = 1$$

Parkov test spočíva v testovaní štatistickej významnosti parametra α odhadnutého zo zlogaritmovaného výrazu, v ktorom sa doporučuje nahradenie logaritmu neznámeho rozptylu ($\ln \sigma_i^2$) logaritmom štvorca reziduálu ($\ln e_i^2$), teda skúma sa významnosť α v modeli

$$\ln e_i^2 = \ln c + \alpha \ln x_i + v_i$$

Takže nech v pôvodnom modeli je závislou premennou y a nezávislými premennými sú x_1 a x_2 . Rovnica je zadaná ako $y = c x_1 x_2$.

1. Vytvoríme si rad reziduálov ey z tohto modelu.
2. Vytvoríme novú rovnicu, kde odhadujeme $\log(ey^2) = c \log(x_1)$, ak podozrievame x_1 .
3. Vo výslednom výpise odhadu sledujeme významnosť parametra pri premennej $\log(x_1)$.

Goldfeld a Quandt navrhli test, kde proti hypotéze o konštantnosti rozptylov je v alternatívnej hypotéze navrhnutý konkrétny tvar závislosti rozptylu od štvorca zvolenej „podozrivej“ premennej modelu. *Goldfeldov-Quandtov test* spočíva vo vypočítaní podielu odhadov rozptylu náhodných porúch dvoch samostatných skupín pozorovaní a jeho porovnaní s tabelovanou hodnotou F rozdelenia s $n_L - k$, $n_S - k$ stupňami voľnosti.

Separátne skupiny pozorovaní sa získajú rozdelením množiny pozorovaní, pri ich neklesajúcom zoradení podľa heteroskedasticitu spôsobujúcej premennej, do dvoch rovnako veľkých skupín s $n_L = n_S$ pozorovaniami (index L označuje vyššie hodnoty). Pričom sa doporučuje vynechanie pätiny centrálnych pozorovaní.

Testovacia štatistika Goldfeldovho-Quandtovo testu sa vypočíta podľa vzťahu

$$GQ = \frac{\mathbf{e}_L^T \mathbf{e}_L / (n_L - k)}{\mathbf{e}_S^T \mathbf{e}_S / (n_S - k)} = \frac{\mathbf{e}_L^T \mathbf{e}_L}{\mathbf{e}_S^T \mathbf{e}_S}$$

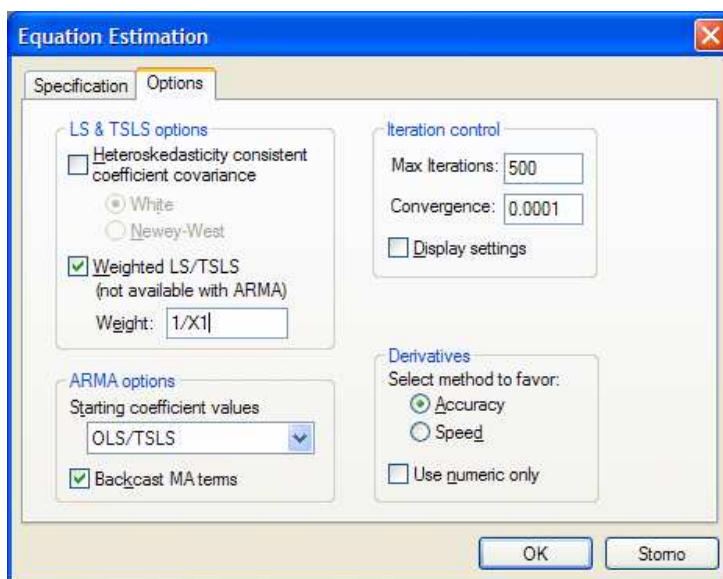
kde $\mathbf{e}_L^T \mathbf{e}_L$ je suma štvorcov reziduálov skupiny pozorovaní s vyššími hodnotami.

Postup v Eviews

1. Najprv si pomenujeme stránku s pôvodnými dátami v našom súbore napríklad menom *Original*, tak že kliknete na uško stránky s nápisom *Untitled* pravým tlačidlom myši. Vyberieme **Rename Workfile Page...** a zadáme meno a **Ok**.
2. Potom si skopírujeme existujúce dáta do nového listu pomocou **Proc/Copy-Extract from Current Page►By Value to New Page or Workfile...** a zadajte **Ok**.
3. Na novej stránke *Untitled* usporiadame údaje podľa premennej, ktorú podozrievame zo spôsobovania heteroskedasticity (napr. x_1) pomocou **Proc/Sort Current Page**, kde zadáme x_1 , vyberieme *Ascending* (teda vzostupne) a potvrdíme **Ok**.
4. Vytvoríme si trojprvkový vektor príkazom **vector(3) GQ**
5. Pri odhade pôvodnej rovnice, povedzme s menom **rov1**, zadáme výber (*Sample*) tak, že obsahuje skupinu *s väčšími hodnotami*. Po odhade si do prvého prvku vektora GQ zaznačíme hodnotu sumy štvorcov reziduálov príkazom **GQ(1)=rov1.@ssr**
(Ak máme napr. 30 pozorovaní, tak pätinu stredných vypúšťame, teda 1/5 z 30 je 6, teda neuvažujeme pozorovania 13 až 18 a skupinu s vyššími pozorovaniami tvoria pozorovania 19 až 30.)
6. Zopakujeme odhad pôvodnej rovnice, opäť s menom **rov1**, ale zadáme výber (*Sample*) tak, že obsahuje skupinu *s menšími hodnotami*. Po odhade si do druhého prvku vektora GQ zaznačíme hodnotu sumy štvorcov reziduálov príkazom **GQ(2)=rov1.@ssr**
(Ak máme napr. 31 pozorovaní, tak pätinu stredných vypúšťame, teda 1/5 z 31 je 6,2; ale vypustením iba 6 pozorovaní by sme dostali dve skupiny s nerovnakým počtom pozorovaní, teda neuvažujeme až 7 pozorovaní, čo sú pozorovania 13 až 19 a skupinu s vyššími pozorovaniami tvoria pozorovania 20 až 31.)
7. Ako posledný krok vypočítame hodnotu GQ štatistiky do tretieho prvku vektora ako podiel prvých dvoch hodnôt, teda príkazom **GQ(3)=GQ(1)/GQ(2)** a porovnáme s tabelovanou hodnotou F rozdelenia.

Vážená metóda najmenších štvorcov, ktorá sa používa na odstránenie problémov vyplývajúcich z prítomnosti heteroskedasticity, sa v Eviews dá zadať buď v definovaní odhadovaného tvaru alebo jednoduchšie ako voľba tejto možnosti na druhej záložke **Options** pri odhade rovnice.

Tak, ak problém v odhadovanom tvare $y = c + x_1 + x_2$ spôsobuje premenná x_1 , môžeme zadať váženú metódu pri odhade zadaním $y/x_1 \quad 1/x_1 \quad x_1/x_1 \quad x_2/x_1$ alebo jednoduchšie výberom tejto metódy **Weighted LS/TLS** na druhej záložke **Options** pri odhade rovnice a zadaním váhy v tomto prípade $1/x_1$.



Na uvedenej záložke je aj druhá zaujímavá metóda pre prípad heteroskedasticity a je ňou **Heteroskedasticity consistent coefficient covariance**. Výber tejto metódy a zvolenie možnosti White odhadne rovnaké hodnoty parametrov ako obyčajná metóda najmenších štvorcov, ale koriguje hodnoty štandardných odchýlok a tým aj t štatistík pre prípad prítomnosti heteroskedasticity.

Multikolinearita

V prípade dokonalej multikolinearity Eviews vypíše správu „Near singular matrix.“

Ak overujeme multikolinearitu pomocou *jednoduchých korelačných koeficientov* (ktorých vysoká hodnota je práve znakom prítomnosti silného stupňa multikolinearity), postup získania korelačnej matice môže byť nasledovný:

- Po odhade konkrétnej rovnice napríklad v tvare $y = c + x_1 + x_2 + x_3$ si vytvoríme skupinu (*Group*) zo všetkých premenných, ktoré sa v nej vyskytujú pomocou **Proc/Make Regressor Group**. Vychádzať z odhadu rovnice nie je nevyhnutné, stačí si klasicky vytvoriť skupinu.
- V takejto skupine, ktorá obsahuje premenné v poradí, v akom sú zadané v rovnici, získame korelačné koeficienty pomocou **View/Correlations►Common Sample** respektíve **Pairwise Sample** (*common* vyberá do výpočtu, len tie pozorovania, ktoré nechýbajú ani pre jednu premennú, *pairwise* vyberá maximálny počet pre každú dvojicu premenných).

Získaná korelačná matica sa dá pomocou **Freeze** previesť na tabuľku (táto funkcia funguje pre každé usporiadanie hodnôt či výsledkov a uľahčuje presun výsledkov naspäť do tabuľkového procesora respektíve do textového editora).

Iným spôsobom zisťovania významného stupňa multikolinearity je výpočet determinantu matice štandardizovaných premenných (od každého pozorovania premennej je odpočítaný jej priemer a tento rozdiel je podelený jej štandardnou odchýlkou). Ak prvky takejto matice podelíme počtom pozorovaní získame maticu korelačných koeficientov medzi premennými. Determinant takejto matice je z intervalu od 0 po 1. Čím bližší je nule, tým bližšie sme k dokonalej multikolinearite. Pre testovanie vytvorili *Farrar a Glauber* testovaciu štatistiku, ktorá má χ^2 rozdelenie s $\frac{1}{2}k(k-1)$ stupňami voľnosti a je definovaná vzťahom

$$FG = -\left(n-1 - \frac{(2k+5)}{6}\right) * \ln |\mathbf{X}_s^T \mathbf{X}_s|$$

kde \mathbf{X}_s označuje maticu štandardizovaných premenných, k je počet vysvetľujúcich premenných, n je počet pozorovaní a ak je testovaná štatistika FG väčšia ako zodpovedajúca tabuľková hodnota, medzi premennými je významný stupeň multikolinearity.

Postup testovania:

1. Vytvoríme si skupinu (Group) z vysvetľujúcich premenných pomenovanú napr. **aa** príkazom **group aa x1 x2 x3**
2. Vytvoríme priamo maticu korelačných koeficientov pomenovanú napr. **kor** pomocou príkazu **matrix kor=@cor(aa)**
3. Vytvoríme si dvojprvkový vektor **fg** pre uchovanie výsledku determinantu sledovanej matice a hodnoty Farrarovej-Glauberovej štatistiky príkazom **vector(2) fg**
4. Do prvého prvku vektora **fg** vypočítame hodnotu determinantu príkazom **fg(1)=@det(kor)**
5. Do druhého prvku vektora **fg** vypočítame hodnotu Farrarovej-Glauberovej štatistiky príkazom **fg(2)=-((30-1)-((2*3+5)/6))*log(@det(kor))**, kde číslo 30 je v našom ukázkovom prípade počet pozorovaní a 3 je počet vysvetľujúcich premenných $x_1 x_2 x_3$. Ostatné čísla sú zo vzťahu.

Prognózovanie

Predpokladajme, že máte k dispozícii údaje pre tri premenné $y x_1 x_2$ za obdobie od prvého štvrťroku roku 1993 po posledný štvrťrok roku 2004. Teda v pracovnom zošite máte nastavený **Range: 1993Q1 2004Q4**. A chcete prognózovať hodnoty y na roky 2005 a 2006. Postup:

1. Dvojkliknete v pracovnom zošite na nápis **Range**, pričom sa zjaví okno *Workfile structure*, v ktorom zmeníte **End date** z hodnoty **2004Q4** na **2006Q4** a potvrdíte **Yes**.
2. Krok 1 zmení aj nastavenie hodnoty **Sample**, ak ste mali predtým **Sample** nastavený na všetky hodnoty z rozsahu (**@all**). Ak ste mali **Sample** zmenený, nastavte ho na rovnakú hodnotu ako **Range**.
3. Otvorte rady vysvetľujúcich premenných $x_1 x_2$ ako skupinu a pomocou **edit+/-** doplňte chýbajúce údaje pre štvrťroky z rokov 2005 a 2006 (expertný odhad).
4. Odhadnite rovnicu y c $x_1 x_2$ s nastaveným výberom **Sample: 1993Q1 2004Q4**.
5. V okne objektu rovnica zvolte **Forecast** vyberte meno pre rad prognózovaných hodnôt **Forecast name** (prednastavené je **yf**) a nastavte **Forecast sample** na **1993Q1 2006Q4**.
6. Výsledkom je rad **yf** (ak ste nezmenili meno), kde hodnoty v období 1993Q1 2004Q4 sú vyrovnané hodnoty z odhadu a hodnoty v období 2005Q1 2006Q4 sú vypočítané prognózy.

Zoznam literatúry:

- [1] EViews 5 User's Guide. Manuál k programu.od spoločnosti Quantitative Micro Software.
- [2] <http://www.feweb.vu.nl/econometriclinks/orientatie/Eviews4introduction.pdf>
- [3] http://wps.aw.com/wps/media/objects/3254/3332253/tutorials/eviews_tutorial_4_1.pdf
- [4] http://www.acc.ncku.edu.tw/chinese/faculty/wangt/About%20Class/Class/eviews_handout.pdf