3. SPECIJALNI PROBLEMI LINEARNIH REGRESIONIH MODELA

3.1. Uvod

U dosadašnjem razmatranju linearnih regresionih modela, ili tačnije u ocenjivanju parametara linearnih regresionih modela metodom najmanjih kvadrata, istaknuta je važnost nekih osnovnih pretpostavki. Te takozvane pretpostavke najmanjih kvadrata se sastoje od sledećeg skupa ograničenja:

1) Član ε_j , $j = 1, 2, \ldots N$ koji opisuje grešku u linearnom regresionom modelu iz dela 3. je slučajna promenljiva sa konačnim vrednostima srednje vrednosti, varijanse i kovarijanse.

2) Očekivana vrednost ε_j je jednaka nuli i svaka od nezavisnih promenljivih vrednost X_{ij} , $i = 1, 2, \ldots, k$; $j = 1, 2, \ldots, N$ je nezavisna od ε_j , ili

$E(\varepsilon) = 0$ i $E(X_i\, \varepsilon^T) = 0$

gde je X_i vektor $X_{ij} = \{ X_{i1}, X_{i2}, \ldots, X_{iN} \}$.

3) Varijansa svakog ε_j , $j = 1, 2, \ldots, N$ je ista bez obzira na j, a kovarijansa jednaka nuli, tj.:

$$E\left(\varepsilon\varepsilon^T\right) = 0$$

što čini takozvanu pretpostavku homoskedastičnosti.

4) Nezavisna promenljiva je data bez greške i ima konačnu srednju vrednost i varijansu.

5) Matrica X linearnog modela sa više nezavisnih promenjljivih ima rang $k < N$, što znači da ne postoji linearna veza u skupu nezavisnih promenljivih kao i da broj observacija prevazilazi broj parametara koji se ocenjuju.

Navedeni skup pretpostavki, obezbeđuje važne osobine ocena parametara i to nepristrasnost, konzistentnost i asimptitsku efikasnost.

Međutim, u analizi praktičnih ekonometrijskih problema neke od navedenih pretpostavki ne važe. Na primer, u studijama tražnje u zavisnosti od dohotka, pretpostavka o konstantnosti varijanse, tj. homoskedastičnosti slučajne greske je nerealna, jer se varijansa obično povećava sa dohotkom.

U ovom delu će se razmotriti neka karakteristična odstupanja od navedenog skupa pretpostavki kao i način njihovog tretmana.

3.2. Multikolinearnost

Ako posmatramo matricu podataka X reda n x k sa rangom k, jedna od osnovnih pretpostavki linearnog regresionog modela je da ne postoji linearna zavisnost između nezavisnih promenljivih. Ta pretpostavka se zasniva na činjenici da ocena $\hat{\beta} = \left(X^T X\right)^{-1} X^T Y$ zahteva inverziju $X^T X$ što je nemoguće ostvariti ako je rang matrice podataka X, a samim tim i $X^T X$, manji od k. Za ovaj slučaj se kaže, da se radi o ekstremnoj multikolinearnosti, kada su neke ili sve nezavisne promenljive potpuno (perfektno) kolinearne. Potpuna kolinearnost nezavisnih promenljivih X_{ij} se izražava relacijom:

$$c_1 X_{1j} + c_2 X_{2j} + \ldots + c_k X_{kj} = 0, \qquad j = 1, 2, \ldots, n$$

koja važi za neki skup vrednosti c_i , koje sve nisu jednake nuli. Drugim rečima jedna od nezavisnih promenljivih se može prikazati kao linearna kombinacija ostalih, tj. rang matrice X se smanjuje za 1.

Primer: Posmatrajmo model

$$P_t = \beta_0 + \beta_1 L_t + \beta_2 D + \beta_3 X_3 + \beta_4 X_4 + \varepsilon_t$$

gde su:
- P - prihod od prodaje
- L - broj prodatih levih cipela
- D - broj prodatih desnih cipela
- X_3 i X_4 - ostali proizvodi

Prihod dolazi od prodaje obe, desne i leve cipele, tako da svaka od njih ima pravo u objašnjenju događaja pri ostvarenju prihoda, ali tada neki od parametara nemaju interpretaciju sa punim značenjem. Na pr. β_1 je parcijalni izvod od p po levoj cipeli, smatrajući sve ostale promenljive uključujući i desnu cipelu konstantnim, što nema smisla, jer se cipele uvek prodaju u paru.

Tako, iako vrednosti parametara β_1 i β_2 mogu biti dobijene, one ne mogu biti interpretirane. Problem se može rešiti predefinisanjem promenljivih, naime, treba definisati novu promenljivu : "par cipela".

Ovaj primer sa cipelama je bio slučaj ekstremne multikolinearnosti, zato što uključuje postojanje tačne linearne relacije između nezavisnih promenljivih.

Osnovne posledice multikolinearnosti su sledeće:

1) Tačnost statističkih ocena parametara drastično opada, pa postaje neobično složeno razdvojiti veze uticaja različitih nezavisnih promenljivih.

2) Ponekad se izbacuju promenljive iz analize, jer recimo njihovi koeficijenti nisu dovoljno različiti od nule, te se smatra da te promenljive nisu od uticaja, što može da bude pogrešan zaključak.

3) Ocenjivanje koeficijenata postaje vrlo osetljivo i npr. kada se na izvestan skup podataka pridodaju neki dodatni, može doći do promene posmatranih koeficijenata.
Kada je promenljiva uvedena u jednačinu i kada je kolinearna sa ostalim promenljivim u jednačini, postavlja se pitanje šta se uobičajeno dešava? Odgovor bi bio, da

1) Standardna greška koeficijenata raste i da
2) Koeficijent determinacije opada.

Pretpostavimo sledeće modele:

$$y_i = \hat{\beta}_1 X_{1i} + \varepsilon_{1i}\ , \qquad i = 1, 2, \dots, N$$
$$y_t = \hat{\beta}_1^{'} X_{1i} + \hat{\beta}_2^{'} X_{2i} + \varepsilon_i\ , \qquad i = 1, 2, \dots, N$$

Ocene $\hat{\beta}_1$ i $\hat{\beta}_1^{'}$ imaju varijanse:

$$V\left(\hat{\beta}_1\right) = \frac{\sigma^2}{\sum x_i^2}$$
$$V\left(\hat{\beta}_1^{'}\right) = \frac{\sigma^2}{\sum x_i^2\left(1 - r_{x_1x_2}^2\right)}$$

Kada postoji korelacija između dve nezavisne promenljive, $r_{x_1x_2}^2$ se razlikuje od nule pa je tada:

$$V\left(\hat{\beta}_1^{'}\right) > V\left(\hat{\beta}_1\right)$$

U nekim slučajevima korelacija između X_1 i X_2 može biti tako bliska da varijansa ocena postane ekstremno velika. U takvim slučajevima ocene nisu pouzdane i mogu biti odbačene, mada se pri tom mora biti izuzetno oprezan.

Kako izbeći multikolinearnost? Svakako, korisno je upotrebiti neku od metoda koje postoje za detekciju prisustva multikolinearnosti u datoj regresiji.

Najjednostavnija metoda je izračunavanje koeficijenata korelacija $r^2_{x_1x_2}$ između parova nezavisnih promenljivih. Ukoliko se pretpostavi da postoji visoki stepen korelacije X_i i X_j tad postoji i odgovarajući stepen kolinearnosti u datom modelu.

Takođe, u slučaju da se sa visokim nivoom značajnosti može usvojiti hipoteza da dati koefinijent b_i nije jednak nuli, a istovremeno F test tvrdi da su svi parametri b_2, b_3, . . . , b_k jednaki nuli, može se sumnjati da je prisutna multikolinearnost.

Dalje, ukoliko se dobije da je znak parametra suprotan od onog što kaže teorija, opet se može posumnjati da je prisutna multikolinearnost.

Formanlniji test detekcije prisustva multikolinearnosti je izračunavanje faktora inflacije varijanse za svaki parametar b_i . Naime, varijansa parametra b_i se može izraziti sa:

$$V\left(\hat{b}_i\right) = \frac{\sigma_e^2}{1 - R_i^2}$$

gde je R_i^2 višestruki koeficijent determinacije modela u kome se nezavisna promenljiva X_i regresirana na preostale nezavisne promenljive.

Faktor:

$$VIF = \frac{1}{1 - R_i^2}$$

se naziva faktorom inflacije varijanse i ukoliko je veći od 10 odnosno R_i^2 veće od 0,9 tad je promenljiva X_i kolinearna sa ostalim i treba je izostaviti.

Faktor:

$$TOL = \frac{1}{VIF} = 1 - R_i^2$$

se naziva faktorom tolerancije.

Primer multikolinearnosti:

Y - veličina prodaje

X_2 - cena artikla

X_3 - veličina ekonomske propagande

$$Y = \beta_1 + \beta_2 X_2 + \beta_3 X_3 + \varepsilon$$

Često se X_3 određuje kao procenat prihoda od prodaje, a prihod (broj komada x cena) i veličina ekonomske propagande su u zavisnosti:

$$X_3 = KX_2$$

Za sledeće podatke imamo primer perfektno korelisanih nezavisnih promenljivih

i	X_2	X_3	Y
1	2	6	23
2	8	9	83
3	6	8	63
4	10	10	103

$$X_3 = 5 + 0.5X_2$$

Međurezultati:

$\overline{X}_2 = 6.5$ $\overline{X}_3 = 8.25$ $\overline{Y} = 68$ $\sum x_{2i}^2 = 35$

$\sum x_{3i}^2 = 8.75$ $\sum x_{2i} y_i = 350$ $\sum x_{3i} y_i = 175$ $\sum x_{2i} x_{3i} = 17.5$

$$b_2 = \frac{350 * 8.75 - 175 * 17.5}{35 * 8.75 - (17.5)^2} = \frac{3062.5 - 3062.5}{306.25 - 306.25} = \frac{0}{0}$$

$$b_2 = \frac{175 * 35 - 350 * 17.5}{0} = \frac{6125 - 6125}{0} = \frac{0}{0}$$

samim tim je i b_i neodređena veličina.

3.3. Lažne (veštačke) promenljive

U dosadašnjim razmatranjima glavne promenljive X, iz opšteg modela opisanog linearnom jednačinom $Y = X\beta + \varepsilon$, su pretpostavljale promenljive koje su imale svoje ekonomsko značenje, odnosno mogle su se kvantitativno iskazati. Linearni model se može proširiti i sa takozvanim lažnim (veštačkim) promenljivim, čije korišćenje ima svrhu da predstavi uticaj raznih kvalitativnih efekata kao što su:

1) Vremenski efekat - promene veza ponašanja od jednog do drugog vremenskog perioda
2) Prostorni efekat - promene u ekonomskim funkcijama u različitim regionima
3) Kvalitativne promenljive- podaci o zanimanju, bračnom stanju, polu, igraju važnu ulogu u određivanju ekonomskog ponašanja
4) Grupisanje promenljivih - koje se naročito koristi za jednostavnije slučajeve, kada mu je primena adekvatna.

Lažne promenljive uzimaju dve vrednosti, nulu ili jedan (te se zato zovu i binarne promenljive). Pretpostavimo da je moguće razdvojiti podatke u više kategorija. Takođe, može se pretpostaviti da podaci unutar svake kategorije imaju istu vrednost parametara, ali zapažanja u različitim kategorijama mogu imati različite skupove parametara. Tada istraživač može dopustiti te razlike, upotrebljavajući lažne promenljive.

Razmotrimo model

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

Dalje pretpostavimo da istraživač ima dve kategorije podataka i veruje da će koeficijent X_2 biti različit u te dve kategorije, a takođe očekuje da će svi ostali parametri ostati isti. Npr. neka promenljive, prisutne u modelu, predstavljaju: Y_i - prinos po hektaru; X_{1i} - ulaganja po hektaru i X_{2i} - količina đubriva po hektaru, gde postoje dve različite vrste: M i N.

Potrebno je oceniti sledeću regresiju:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \beta_3 \left(D X_{2i} \right) + \varepsilon_i$$

gde je:

$$D = \begin{cases} 1, \text{ ako je M} \\ 0, \text{ ako je N} \end{cases}$$

Kada lažna promenljiva ima vrednost nula, đubrivo koje se koristi je N i posmatra se regresija:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

Ako pak, lažna promenljiva ima vrednost jedan, koristi se đubrivo M, pa je regresija koja se posmatra:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \left(\beta_2 + \beta_3 \right) X_{2i} + \varepsilon_i$$

Lažna promenljiva D se može koristiti i kao aditivna nezavisna promenljiva, pa tada postoji promena konstantnog člana u dve kategorije

$$Y_i = \beta_0 + \alpha D + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

gde je

$$D = \begin{cases} 1, \text{ ako je M} \\ 0, \text{ ako je N} \end{cases}$$

Ako se koristi đubrivo M konstantni član regresije je $\beta_0 + \alpha$, dok pri korišćenju đubriva N taj član je β_0.

Klasičan primer iz ove oblasti je i primer funkcije potrošnje u dva vremenska perioda: za vreme rata i mira:

$C = \alpha_1 + \beta Y + \varepsilon$, - ratno vreme
$C = \alpha_2 + \beta Y + \varepsilon$, - mirnodopsko vreme

Pri tome je $\alpha_2 > \alpha_1$. Ako β nije identično tada se ništa ne dobija upotrebom lažnih promenljivih, tj. pristrasnost se može izbeći regresujući odvojeno, a da se ne izgubi efikasnost. Sa druge strane, ako je β zajedničko, korisno je koristiti sve podatke, da bi se dobila što efikasnija ocena tog parametra. To se obično postiže kombinovanjem jednačina u jednu jedinu vezu:

$$C = \alpha_1 + \gamma D + \beta Y + \varepsilon$$

gde je

$$D = \begin{cases} 1, \text{ ako je mirnodopsko vreme} \\ 0, \text{ ako je ratno vreme} \end{cases}$$

Za različite vrednosti lažne promenljive, ponovo se dobijaju dve prvobitne jednačine:

$C = \alpha_1 + \beta Y + \varepsilon$, - ratno vreme
$C = \alpha_1 + \gamma + \beta Y + \varepsilon$, - mirnodopsko vreme

gde je $\alpha_1 + \gamma = \alpha_2$

Kod jedinstvene regresije, C je zavisna promenljiva a matrica podataka izgleda

$$\begin{bmatrix} 0 & 1 & Y_1 \\ 0 & 1 & Y_2 \\ 1 & 0 & Y_3 \\ 1 & 0 & Y_4 \\ 1 & 0 & Y_5 \\ 0 & 1 & Y_6 \\ \vdots & \vdots & \vdots \\ 0 & 1 & Y_n \end{bmatrix}$$

gde se pretpostavlja da se prva dva perioda odnose na mir, sledeća tri na rat, što sledi mirnodopski period do kraja.

Razmotrimo jedan drugi primer sa sledećim modelom:

$$Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} + \varepsilon_i$$

Naime, ovaj model treba da objasni ponašanje recimo, ljudi, žena i dece, ali ako samo β_1 varira za tri promenljive to nije dovoljno. Potrebne su dve lažne promenljive, ili uopšte za m kategorija, neophodno je uvesti (m-1) lažnih promenljivih.

$$Y_i = \beta_0 + \beta_1 X_{1i} + \alpha_1 \left(D_1 X_{1i} \right) + \alpha_2 \left(D_2 X_{2i} \right) + \beta_2 X_{2i} + \varepsilon_i$$

gde je:

$$D_1 = \begin{cases} \text{1, ako je } \sim \text{ovek} \\ \text{0, ostali} \end{cases}$$
$$D_2 = \begin{cases} \text{1, ako je `ena} \\ \text{0, ostali} \end{cases}$$

Lažne promenljive se mogu koristiti i u analizi vremenskih seroka za tretman sezonske komponente. Recimo, ukoliko su podaci vremenske serije dati kvartalno i ukoliko se pretpostavi linearni trend uz prisustvo sezonske komponente tad je

$$Y_t = \beta_0 + \beta_1 t + \beta_2 Q_2 + \beta_3 Q_3 + \beta_4 Q_4$$

gde je:

$$Q_2 = \begin{cases} 1, udrugomk\,\mathrm{var}\,talu \\ 0, inačn \end{cases}$$

$$Q_3 = \begin{cases} 1, utrećre\ kvatalu \\ 0, inačn \end{cases}$$

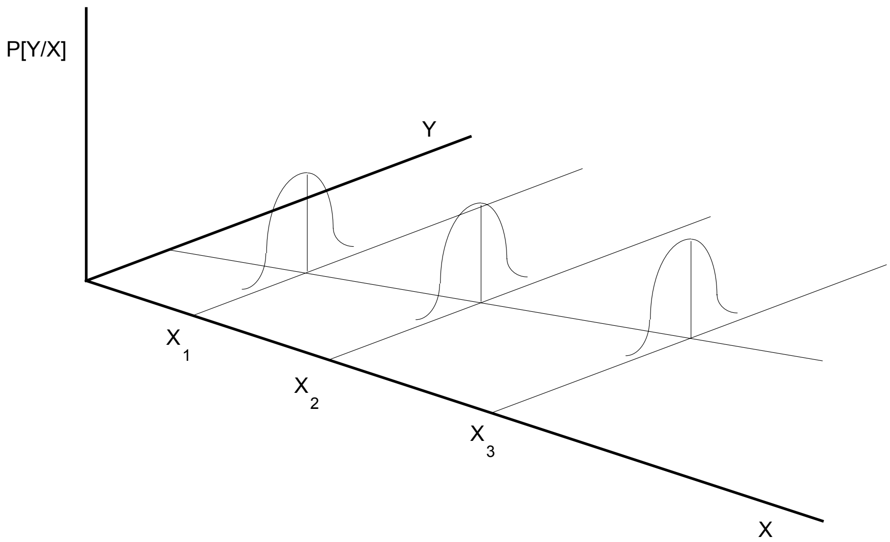
$$Q_4 = \begin{cases} 1, učetvrtomk\,\mathrm{var}\,talu \\ 0, inačn \end{cases}$$

Prvi kvartal je uzet kao bazni. Koeficijenti uz lažne promenljive Q_2 , Q_3 i Q_4 određuju srednji prirast (pozitivan ili negativan) u odgovarajućem kvartalu selektivno u odnosu na bazni kvartal.

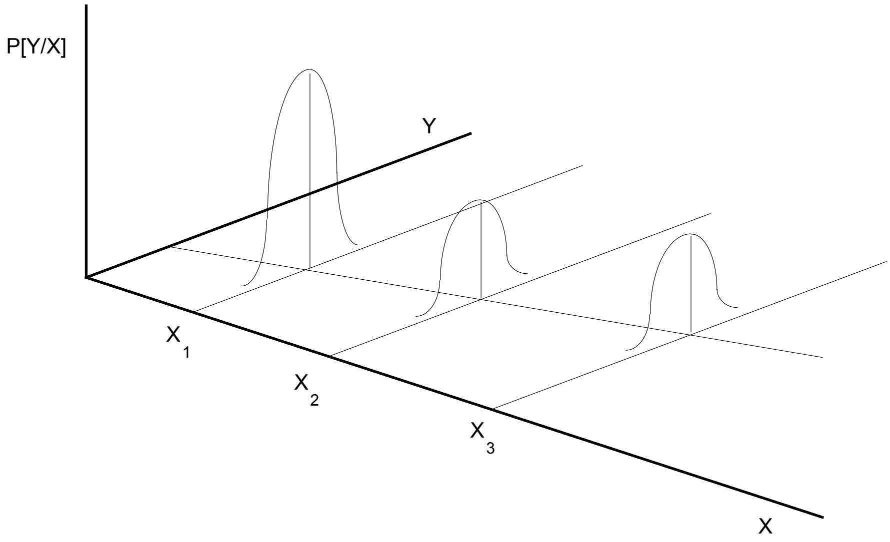
3.4. Heteroskedastičnost i uopštena metoda najmanjih kvadrata

U praksi ekonometrijskih studija, pretpostavka o konstantnoj varijansi za članove poremećaja ponekad ne odgovara stvarnosti.

Pretpostavka homoskedastičnosti (od grčkog: ravnomerno rasprostranjeni), varijansu V(ε) predpostavlja konstantnom duž cele linije regresije (Sl.3.1.).



Kako je to nerealna pretpostavka, za stvarne poremećaje se može pretpostaviti da su heteroskedastični što je ilustrovano na Sl.3.2.



Sl.3.2.

Pretpostavimo da je u linearnom modelu:

$$Y = \alpha + \beta X + \varepsilon$$

varijansa oblika

$$V(\varepsilon) = KX^{m} = \frac{K}{W}$$

gde m uzima vrednost $0 \leq m \leq 2$, i K = const.

Kada je m = 0 tada je $V = K = \sigma^2$ što je bio slučaj koji je do sada razmatran. Neka je

$$W_i = \frac{1}{X_i^m}$$

i definišimo novi član poremećaja:

$$\varepsilon_i^* = \sqrt{W_i}\varepsilon_i$$

Ako je očekivana vrednosti $E(\varepsilon_i) = 0$ tada je i $E(\varepsilon_i^*) = 0$

Odnosno

$$V(\varepsilon_i^*) = W_i \frac{K}{W_i} = K$$

Na predpostavljeni linearni model i poremećaj ε_j^* treba primeniti metodu najmanjih kvadrata:

$$\min Q = \sum_{i=1}^{n} e_i^* = \sum W_i e_i^2 = \sum W_i (Y_i - a - bX_i)^2 \text{ odnosno} \tag{3.1}$$

$$\sum W_i Y_i = a\sum W_i + b\sum W_i Y_i$$

$$\frac{\partial Q}{\partial b} = 2\sum W_i (Y_i - a - bX_i)(-X_i) = 0$$

$$\sum W_i X_i Y_i = a\sum W_i X_i + b\sum W_i X_i^2 \tag{3.2}$$

Opšti slučaj primenjene transformacije:

$V(\varepsilon_i) = xf(x)$ i $Y = \alpha + \beta X + \varepsilon$

i treba transformacijom da se dobije:

$$\varepsilon^* = \frac{\varepsilon}{\sqrt{f(x)}} \text{ jer } \quad V(\frac{\varepsilon}{\sqrt{f(x)}}) = k \Rightarrow \frac{Y}{\sqrt{f(x)}} = \frac{\alpha}{\sqrt{f(x)}} + \frac{\beta}{\sqrt{f(x)}} X + \frac{\varepsilon}{\sqrt{f(x)}}$$

Iz (3.1) i (3.2) sledi:

$$a = \frac{W_i Y_i - b\sum W_i X_i}{\sum W_i}$$

$$b = \frac{W_i \sum W_i X_i Y_i - \sum W_i Y_i \sum W_i X_i}{\sum W_i \sum W_i \ X_i^2 - \sum W_i X_i \sum W_i X_i}$$

odnosno:

$$b = \frac{\sum W_i X_i Y_i - \dfrac{\sum W_i Y_i}{\sum W_i} \sum W_i X_i}{\sum W_i X_i^2 - \dfrac{\sum W_i X_i}{\sum W_i} \sum W_i X_i}$$

Ekvivalentno rešenje u slučaju višestruke regresije se može dobiti sledećom transformacijom. Najpre se definiše matrica D:

$$D = \begin{bmatrix} 1/\sigma_1 & 0 & \cdots & 0 \\ 0 & 1/\sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1/\sigma_N \end{bmatrix} \qquad \text{pri čemu je } \sigma_i = \sqrt{V(\varepsilon_i)}$$

Zatim se sistem:

$y_i = x_i\beta + \varepsilon_i \quad$ i=1, 2, . . . , N

odnosno $Y = X\beta + \varepsilon$

množi sa D kada se dobija:

$$DY = (DX)\beta + D\varepsilon \tag{3.3}$$

Tada je:

$$E\left[D\varepsilon(D\varepsilon)^T\right] = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

$$D\varepsilon = \begin{bmatrix} \varepsilon_1/\sigma_1 \\ \varepsilon_2/\sigma_2 \\ \vdots \\ \varepsilon_N/\sigma_N \end{bmatrix}$$

$$DY = \begin{bmatrix} Y_1/\sigma_1 \\ Y_2/\sigma_2 \\ \vdots \\ Y_N/\sigma_N \end{bmatrix}$$

$$DX = \begin{bmatrix} 1/\sigma_1 & X_{21}/\sigma_1 & \cdots & X_{k1}/\sigma_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1/\sigma_N & X_{2N}/\sigma_N & \cdots & X_{kN}/\sigma_N \end{bmatrix}_{Nxk}$$

Na ove transformisane podatke direktno se može primeniti metoda običnih najmanjih kvadrata kada se dobija ocena

$$b = \left\{(DX)^T DX\right\}^{-1} (DX)^T (DY) \qquad (3.4)$$

$$E\left[D\varepsilon(D\varepsilon)^T\right] = E\begin{bmatrix} \frac{\varepsilon_1^2}{\sigma_1^2} \cdots \frac{\varepsilon_1\varepsilon_N}{\sigma_1\sigma_N} \\ \cdots \\ \frac{\varepsilon_N\varepsilon_1}{\sigma_N\sigma_1} \cdots\cdot \frac{\varepsilon_N^2}{\sigma_N^2} \end{bmatrix}_{NxN} = \begin{bmatrix} \frac{E(\varepsilon_1^2)}{\sigma_1^2} \cdots \frac{E(\varepsilon_1\varepsilon_N)}{\sigma_1\sigma_N} \\ \cdots \\ \frac{E(\varepsilon_N\varepsilon_1)}{\sigma_N\sigma_1} \cdots\cdot \frac{E(\varepsilon_N^2)}{\sigma_N^2} \end{bmatrix}_{NxN} = I_N$$

dakle $E(\frac{\varepsilon_i}{\sigma_i})^2 = 1$ - jedinična varijansa, i $E(\frac{\varepsilon_i\varepsilon_j}{\sigma_i\sigma_j}) = 0$ - nema korelacije izmedju $\frac{\varepsilon_i}{\sigma_i}$ i $\frac{\varepsilon_j}{\sigma_j}$

Ovo rešenje je isto kao i uopštene metode najmanjih kvadrata, ali je dobijeno na jednostavniji način. Ukratko, postupak za njegovo rešavanje se svodi na:

1) Transformisanje shodno jednačini (3.3)
2) Primenu metode običnih najmanjih kvadrata.

3.5 Uopštena metoda najmanjih kvadrata

Jedna od osnovnih pretpostavki u rešavanju osnovne jednačine standardnog linearnog modela:

$$Y = X\beta + \varepsilon$$

metodom najmanjih kvadrata je bila:

$$V(\varepsilon) = E\left(\varepsilon\varepsilon^T\right) = \sigma_e^2 I$$

Međutim, ova pretpostavka se može zameniti sledećom koja je opštija i time manje ograničena:

$$V(\varepsilon) = E\left(\varepsilon\varepsilon^T\right) = \sigma^2 W \qquad (3.3')$$

gde je σ^2 nepoznati faktor. a W poznata, pozitivno definitna matrica reda N. Ovakvom pretpostavkom se obuhvata heteroskedastičnost, tj. slučaj kad varijansa nije konstantna, kao i mogućnost korelacije grešaka.

Pozitivna definitnost matrice W implicira da ne postoji savršena korelacija između bilo kog para grešaka $\left(\varepsilon_i\ ,\varepsilon_j\right)$, a takođe da ne postoji savršena korelacija u celom skupu ε.

Problemu ocene parametara β uz uslov (3.3')se može prići na više načina. Kako je W simetrična i pozitivna definitna matrica, tada mora da postoji nesingularna matrica P reda N takva da je:

$$W^{-1} = PP^T$$

Ako osnovnu jednačinu linearnog modela pomnožimo matricom P sa leve strane dobijamo:

$$PY = PX\beta + P\varepsilon$$

ili:

$$Y_1 = X_1\beta + \varepsilon_1 \tag{3.5}$$

gde su:

$$Y_1 = PY; \quad X_1 = PX; \quad i \quad \varepsilon_1 = P\varepsilon$$

Na osnovu ovih definicija sledi da su greške e homoskedastične i sa nultom kovarijansom, tj.:

$$E\left(\varepsilon_1\varepsilon_1^T\right) = E\left(P\varepsilon\left(P\varepsilon\right)^T\right) = PE(\varepsilon\varepsilon^T)P^T = P(\sigma^2 W)P^T = \sigma^2 P\left(PP^T\right)^{-1}P^T = \sigma^2 I$$

te se prema tome na transformisanu jednačinu (3.5) može primeniti direktno metod najmanjih kvadrata i dobiti ocena:

$$b = \left(X_1^T X\right)^{-1} X_1^T Y_1$$

ili izraženo preko originalnih promenljivih:

$$b = \left(X^T W^{-1}\right) X^T W^{-1} Y \tag{3.6}$$

Ocena (3.6) za vektor parametara β se često naziva Aitken-ova ocena.

Lako se pokazuje da je varijansa ocena (3.6):

$$V(b) = \sigma^2 \left(X^T W^{-1} X \right)^{-1} \tag{3.7}$$

Isto tako, ocena varijanse slučajne greške je:

$$\sigma_e^2 = \frac{e^T W^{-1} e}{(N-k)} \tag{3.8}$$

Na osnovu izraza (3.6), (3.7) i (3.8) može se zaključiti da se na odgovarajuće veličine u uopštenoj metodi najmanjih kvadrata, mogu formalno izvesti, kad se u odgovarajućim izrazima dobijenim standardnom metodom najmanjih kvadrata, matrica X zameni matricom $W^{-1}X$, vektor Y sa vektorom $W^{-1}Y$ i vektor e sa vektorom $W^{-1}e$.

3.6. Autokorelacija

U prethodnim razmatranjima, kod proučavanja linearnih modela, jedna od pretpostavki je bila da nulta - kovarijansa za članove greške implicira: $E\left(\varepsilon, \varepsilon^T\right) = \sigma^2 I$ u čemu članovi sporedne dijagonale daju

$$E\left(\varepsilon_t \varepsilon_{t+s}\right) = 0 \text{ za sve } t \text{ i sve } s \neq 0$$

Pretpostavimo slučaj kada jedan linearni model opisuje zavisnost Y u odnosu na X u bilo kom vremenskom periodu t prilikom analize podataka jedne vremenske serije:

$$Y_t = \alpha + \beta X_t + \varepsilon_t$$

gde su α i β parametri koji se ocenjuju a ε_t greška koja je često serijski korelisana ili autokorelisana. To u stvari znači da je ta greska ε_t u bilo kom vremenskom trenutku t, u korelaciji sa jednom ili više ranijih (prethodnih) vrednosti, ε_{t-1} , ε_{t-2} , itd.

Kada su članovi slučajne greške e autokorelisani, obični najmanji kvadrati ne obezbeđuju najbolje nepristrasno linearno ocenjivanje, parametara α i β.

Za model:

$Y_t = \beta X_t + \varepsilon_t$, t=1, 2, . . . , N (3.9)

pretpostavimo da greške ε_t nisu više nekorelisane, odnosno da se generišu Markovljevim lancem ili najjednostavnijim tipom autoregresione šeme:

$\varepsilon_t = \rho\varepsilon_{t-1} + V_t \quad , \quad t = 2, 3, \ldots N \quad , \quad |\rho| < 1$ (3.10)

gde su

$E(v) = 0$, i $\text{cov}(v) = \sigma_v^2 I$

Kvadriranjem izraza (3.10) i traženjem očekivane vrednosti tog izraza, dobija se:

$\sigma_\varepsilon^2 = \rho^2\sigma^2 + \sigma_v^2$ odnosno (3.11)

$\sigma_\varepsilon^2 = \dfrac{\sigma_v^2}{1-\rho^2}$ što je homoskedastično

Iz (3.9) sledi:

$$\hat{\beta} = \frac{\sum X_t Y_t}{\sum X_t^2} = \frac{\sum X_t(\beta X_t + \varepsilon_t)}{\sum X_t^2} = \beta + \frac{\sum X_t \varepsilon}{\sum X_t^2}$$

odnosno:

$$E(\hat{\beta}) = \beta + E\left(\frac{\sum X_t \varepsilon_t}{\sum X_t^2}\right) = \beta + \frac{1}{\sum X_t^2} E\left(\sum X_t \varepsilon_t\right) = \beta + \frac{\sum E(X_t \varepsilon_t)}{\sum X_t^2} = \beta$$

gde je sumacija izvršena u granicama t = 1 do t = N.

Kako je E(ε_t) jednako nuli i kako je predpostavljeno da je promenljiva X_t deterministička, $\hat{\beta}$ je striktno nepristrasno, tj. serijska korelacija u članu greške ne uvodi bilo kakvu pristrasnost u regresionom ocenjivanju.

Nadalje predpostavimo varijansu od $\hat{\beta}$:

$$V(\beta) = E\left[(\beta-\beta)^2\right] = E\left[\left(\frac{\sum X_t\varepsilon_t}{\sum X_t^2}\right)\right] =$$

$$= \left(\frac{1}{\sum X_t^2}\right)^2 \left[X_1^2\sigma_\varepsilon^2 + X_2^2\sigma_\varepsilon^2 + \ldots + X_N^2\sigma_\varepsilon^2 + 2X_1X_2\varepsilon_1\varepsilon_2 + 2X_2X_3\varepsilon_2\varepsilon_3\ldots\right]$$

Kako su promenljive X fiksirane, dobiće se:

$$V(\hat{\beta}) = \left(\frac{1}{\sum X_t^2}\right)^2 = \left[X_1^2\sigma_\varepsilon^2 + X_2^2\sigma_\varepsilon^2 + \ldots + X_N\sigma_\varepsilon^2 + 2X_1X_2\rho\sigma_\varepsilon^2 + \right.$$

$$\left. + 2X_2X_3\rho\sigma_\varepsilon^2 + + \ldots + 2X_1X_3\rho^2\sigma_\varepsilon^2 + 2X_2X_4\rho^2\sigma_\varepsilon^2 + \ldots\right] =$$

$$= \left(\frac{1}{\sum X_t^2}\right)^2 \sigma_\varepsilon^2\left[\sum X_t^2 + 2\sum_{s=1}^{N-1}\rho^s\sum_{1}^{N-s}X_tX_{t+s}\right]$$

Kada su članovi greške serijski nezavisni ($\rho = 0$), tada $V(\hat{\beta})$ zavisi samo od $\sum X^2$ i σ_ε^2, ali kada je greška u serijskoj korelaciji, varijansa od $\hat{\beta}$ zavisi i od članova $\sum X_tX_{t+s}$ Ako istraživač zna vrednost parametara ρ, tada se odgovarajućom linearnom transformacijom promenljivih, može reducirati jednačina na formu u kojoj ocenjivanje običnim najmanjim kvadratima obezbeđuje minimalnu varijansu ocenjivanja.

Primer: Kada je ρ poznato, transformacija,

$$\varepsilon_t^* = \varepsilon_t - \rho\varepsilon_{t-1}$$

važi i ε_t^* je serijski nekorelisano. Koristeći jednačinu (3.10) dobija se:

$$\varepsilon_t^* = \rho\varepsilon_{t-1} + v_t - \rho\varepsilon_{t-1} = v_t$$

i v_t je serijsko nezavisno i homoskedastično. Na taj način v_t zadovoljava sve pretpostavke za ocenjivanje običnim najmanjim kvadratima u cilju dobijanja najbolje nepristrasne ocene.

Pretpostavimo osnovni model koji odgovara vremenskim periodima t i t-1:

$$Y_t = \beta X_t + \varepsilon_t \qquad (3.12)$$

$$Y_{t-1} = \beta X_{t-1} + \varepsilon_{t-1} \qquad (3.13)$$

Jednačinu (3.13) treba pomnožiti sa ρ, tj.:

$$\rho Y_{t-1} = \rho\beta X_{t-1} + \rho\varepsilon_{t-1}$$

i oduzeti od jednačine (3.12) kada se dobija:

$$\left(Y_t - \rho Y_{t-1}\right) = \beta\left(X_t - \rho X_{t-1}\right) + \left(\varepsilon_t - \rho\varepsilon_{t-1}\right) \qquad (3.14)$$

Definišući promenljive y^* i x^* kao $\left(y_t - \rho y_{t-1}\right)$ i $\left(X_t + \rho x_{t-1}\right)$, respektivno, konačno se dobija

$$Y_t^* = \beta X_t^* + \varepsilon_t^* \qquad (3.15)$$

U jednačini (3.15) član greške je serijski nekorelisan i homoskedastičan.
Od ranije se zna da je:

$$E\left(\varepsilon_1\right) = 0 \quad \text{i} \quad V\left(\varepsilon_1\right) = \frac{\sigma_v^2}{1-\rho^2}$$

pa odogovarajućom transformacijom ε_1 , može se dobiti ε_1^*koje će imati iste statističke osobine kao i ostali ε_s^* .

Kako je $\left(1-\rho\right)^2$konstantno može se pisati:

$$\varepsilon_1^* = \sqrt{\left(1-\rho^2\right)}\varepsilon_1$$

odnosno

$$V\left(\varepsilon_1^*\right) = \left(1-\rho^2\right)\sigma_\varepsilon^2 = \sigma_v^2$$

pa se analogno tome dobija:

$$y_1^* = \sqrt{\left(1-\rho^2\right)}y_1$$
$$x_1^* = \sqrt{\left(1-\rho^2\right)}x_1$$

Sada je problem sveden na T zapažanja i ocenjivanja običnim najmanjim kvadratima vredi za nepristrasnu minimalnu varijansu ocenjivanje β. Napomena: To je tačno metod uopštenih najmanjih kvadrata. Pretpostavimo najopštiji slučaj:

$$Y_t = \beta_0 + \beta_1 X_{1t} + \beta_2 X_{2t} + \ldots + \beta_p X_{pt} + \varepsilon_t$$

gde je član greške generisan sa Markovljevim lancem prvog reda sa parametrom ρ.

Sada se želi dobiti uopštena najmanja ocena primenjujući obične najmanje kvadrate na:

$$Y_t^* = \beta_0^* + \beta_1^* X_{1t}^* + \beta_2^* X_{2t}^* + \ldots + \beta_p^* X_{pt}^* + \varepsilon_t^*$$

gde su:

$$Y_t^* = Y_t - \rho Y_{t-1}$$
$$X_{1t}^* = X_{1t} - \rho Y_{i,t-1}$$
$$Y_1^* = \sqrt{\left(1-\rho^2\right)}Y$$
$$X_1^* = \sqrt{\left(1-\rho^2\right)}X_{i1}$$

Problem se sastoji u tome, što se ne zna prava vrednost za ρ. Kada istraživač sumnja da su članovi greške u serijskoj korelaciji i veruje da su vrednosti, reciom ρ, tada se problem ocenjuje i tretira kao da su u pitanju pravi (istiniti) parametri ρ.

Ako je $\rho \neq \rho^*$, ocenjeni parametri β nisu više ocenjene minimalne varijanse. Ponekad, kada su članovi greške generisani Markovljevim lancem prvog reda, upotreba ρ različitog od ρ, može biti vrlo efikasna (tada se može primeniti metoda običnih najmanjih kvadrata iz netransformisanih podataka).

Efikasnost metode uopštenih kvadrata za različite vrednosti se definiše kao odnos:

$$\frac{V\left(\hat{\beta}^*\right)\rho}{V\left(\hat{\beta}^*\right)\rho^*}$$

gde su:

ρ* - ocena vrednosti i

ρ - tačna vrednost.

Analogno izrazu (3.10) može se pisati:

$$X_t = \lambda X_{t-1} + w_t$$

kada je u pitanju regresiona šema sa trendom ka zajedništvu u konkretnim ekonomskim promenljivim tj. kada se vrši proces generisanja nezavisne promenljive.

Primer:

Neka je $\rho = \lambda = 0$

U tom slučaju $\hat{\beta}^*$ je nepristrasna ocena od b pa se tada dobija:

$$V\left(\hat{\beta}^*\right) = \frac{\sigma_\varepsilon^2}{\sum X_t^2}\left\{1 + \frac{\rho^{*2}}{\left(1-\rho^{*2}\right)^2}\right\}$$

Najmanja varijansa se dobija kada je skup vrednosti ρ^* jednak nuli. Postavlja se pitanje, kako dobiti ρ^* ? Uobičajeno je da se odgovor traži od ostataka običnih najmanjih kvadrata:

$$\rho^* = \frac{\sum e_t \quad {}_{t-1}}{\sum e_{t-1}^2}$$

Ocenjeno ρ^* nije tačna vrednost za ρ , a i ima statističku raspodelu. Ono nije konzistentno, ali pristrasno u malim uzorcima. Pristrasnost je negativna i reda je ρ / T, kada T predstavlja veličinu uzorka.

Kada su nezavisne promenljive takođe u serijskoj korelaciji, tada pristrasnost zavisi takođe od parametara koji generišu njihovu serijsku korelaciju gde je:

$$X_t = \lambda X_{t-1} + w_t$$

a pristrasnost je jednaka:

$$\Pi \approx \left[\left(\rho + \lambda\right) / T\right]$$

Varijansa od ρ^* je reda 1/T.

3.6.1. Durbin-Watson-ov test prisustva autokorelacije

Uvodi se statistika

$$d = \frac{\sum_{i=2}^{N}\left(e_i - e_{i-1}\right)^2}{\sum_{i=1}^{N} e_i^2}$$

e_i - ocenjena vrednost člana slučajne greške.

Razvijanje gornjeg izraza sledi

$$d = \frac{\sum_{i=2}^{N} e_i^2}{\sum_{i=1}^{N} e_i^2} + \frac{\sum_{i=2}^{N} e_{i-1}^2}{\sum_{i=1}^{N} e_i^2} - 2\frac{\sum_{i=2}^{N} e_i^2 e_{i-1}}{\sum_{i=1}^{N} e_i^2}$$

$$\downarrow \qquad \downarrow$$

$$\approx 1 \qquad \approx 1$$

1° Ako ne postoji autokorelacija:

$$\sum_{i=2}^{N} e_i e_{i-1} \approx 0 \quad \Rightarrow \quad \underline{d \approx 2}$$

2° Ako postoji pozitivna autokorelacija

$$e_i \approx e_{i-1} \quad \Rightarrow \quad d = 2 - 2\frac{\sum_{i=2}^{N} e_i^2}{\sum_{i=1}^{N} e_i^2} \quad \Rightarrow \quad \underline{d \approx 0}$$

3° Ako postoji negativna autokorelacija:

$$e_i \approx -e_{i-1} \quad \Rightarrow \quad d = 2 + 2\frac{\sum_{i=2}^{N} e_i^2}{\sum_{i=1}^{N} e_i^2} \quad \Rightarrow \quad \underline{d \approx 4}$$

1°	$d \approx 2$	nema autokorelacije
2°	$0 \le d < 2$	pozitivna autokorelacija
3°	$2 < d \le 4$	negativna autokorelacija