Κεφάλαιο **1**

Συσχέτιση

1.1. ΕΙΣΑΓΩΓΗ

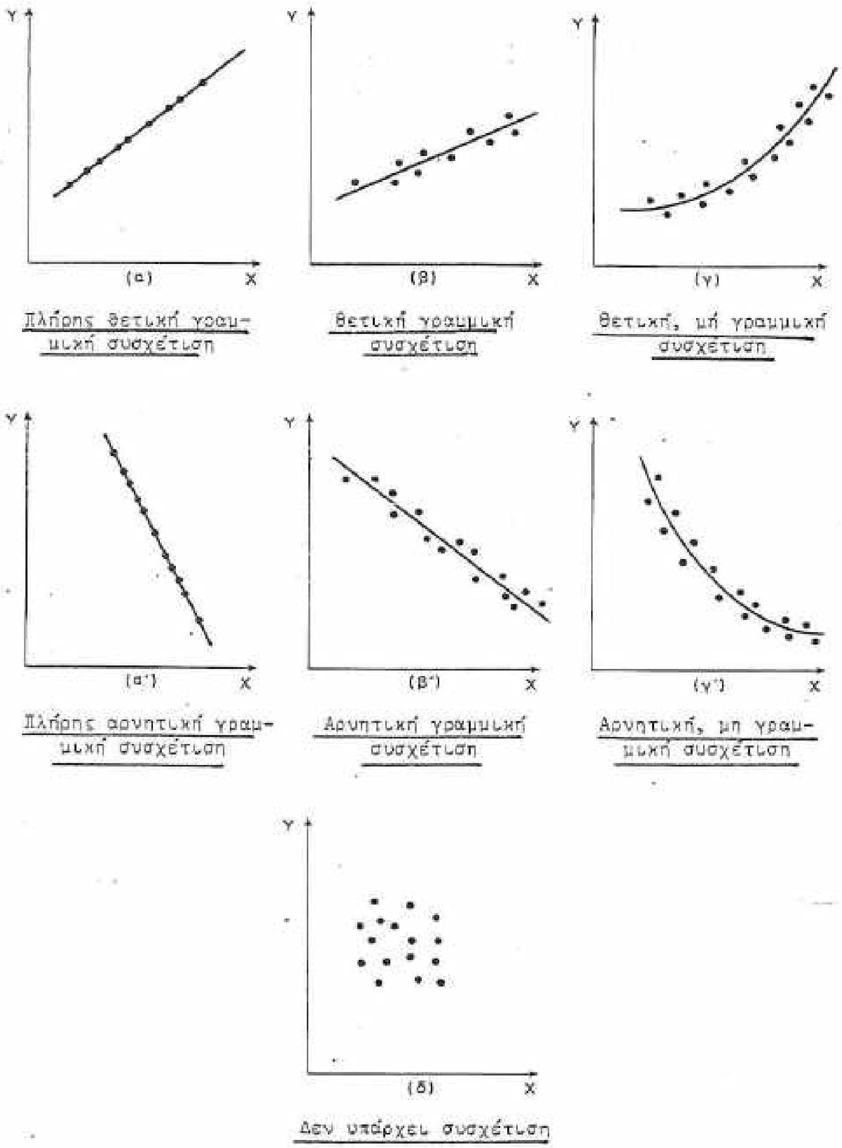
"Συσχέτιση" είναι η μέθοδος που χρησιμοποιείται για τη μέτρηση του βαθμού αλληλοεξάρτησης των οικονομικών μεταβλητών. Αν οι μεταβλητές είναι δύο τότε έχουμε την 1 "απλή συσχέτιση" ενώ αν είναι περισσότερες έχουμε την 2 "πολλαπλή συσχέτιση". Στο κεφάλαιο αυτό θα ασχοληθούμε με την απλή συσχέτιση ενώ η πολλαπλή συσχέτιση θα εξεταστεί μετά την πολλαπλή παλινδρόμηση. ΟΡΙΣΜΟΣ ΣΥΣΧΕΤΙΣΗΣ.

Αν μας δοθεί ένα δείγμα από ζεύγη αντίστοιχων τιμών (παρατηρήσεων) (X_i, Y_i), $i=1,2,\ldots,n$, για δύο μεταβλητές X και Y, μπορούμε να θεωρήσουμε τα σημεία στα οποία απεικονίζονται τα διαταγμένα ζεύγη (X_i, Y_i), σε ένα σύστημα ορθογωνίων αξόνων. Από τα διάφορα υποθετικά διαγράμματα του σχήματος 1.1 προκύπτουν τα εξής:

(i) Η συσχέτιση των μεταβλητών X και Y είναι 1 "γραμμική" [διαγράμματα (α), (α'), (β), (β')] αν τα σημεία (X_i, Y_i) συγκεντρώνονται πάνω ή κοντά σε μια ευθεία, ή 2 "μη γραμμική" [διαγράμματα (γ) και (γ')] αν τα σημεία (X_i, Y_i) συγκεντρώνονται πάνω ή κοντά σε μια καμπύλη ανωτέρου βαθμού. ΓΡΑΜΜΙΚΗ ΜΗ ΓΡΑΜΜΙΚΗ

(ii) Η συσχέτιση είναι 1 "θετική" [διαγράμματα (α), (β), (γ)] αν η κλίση της ευθείας ή της καμπύλης είναι θετική και 2 "αρνητική" [διαγράμματα (α'), (β'), (γ')] αν η κλίση της ευθείας ή της καμπύλης είναι αρνητική. ΘΕΤΙΚΗ ΑΡΝΗΤΙΚΗ

(iii) Η συσχέτιση είναι 1 "πλήρης" όταν εκφράζεται από ακριβή μαθηματική σχέση μεταξύ των δύο μεταβλητών, δηλαδή ό- ΠΛΗΡΗΣ



Σχήμα 1.1: Διάφορες μορφές απλής συσχέτισης.

ταν όλα τα σημεία (X_i, Y_i) του διαγράμματος βρίσκονται ακριβώς πάνω σε κάποια ευθεία ή καμπύλη [διαγράμματα (α), (α')]

iv) Αν η διασπορά των σημείων είναι όπως στο διάγραμμα (δ), τότε δεν υπάρχει συσχέτιση μεταξύ των τιμών των μεταβλητών X και Y. Στο κεφάλαιο αυτό θα ασχοληθούμε αποκλειστικά με τη μέτρηση της γραμμικής συσχέτισης, θετικής ή αρνητικής.

1.2. ΜΕΤΡΗΣΗ ΤΗΣ ΓΡΑΜΜΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ. Ο ΣΥΝΤΕΛΕΣΤΗΣ ΣΥΣΧΕΤΙΣΗΣ

Το διάγραμμα της διασποράς των σημείων (X_i, Y_i) μας δείχνει, σε γενικές γραμμές, το είδος και το βαθμό της γραμμικής συσχέτισης μεταξύ των μεταβλητών X και Y. Συχνά όμως απαιτείται ένα ακριβές αλγεβρικό μέτρο του βαθμού της γραμμικής συσχέτισης. Το μέτρο αυτό δίνεται από το "συντελεστή συσχέτισης". Θα διακρίνουμε το 1 συντελεστή συσχέτισης r_{XY} που εκτιμάται από ένα συγκεκριμένο δείγμα τιμών των X και Y και το 2 συντελεστή συσχέτισης ρ_{XY} των X και Y στους πληθυσμούς.

α. Ο συντελεστής συσχέτισης r_{XY} στο δείγμα

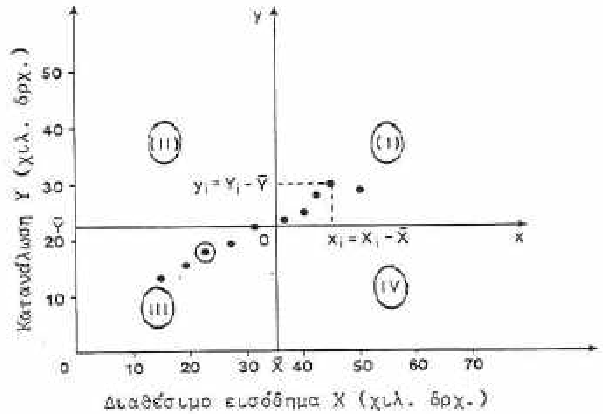
Ας υποθέσουμε ότι διαθέτουμε παρατηρήσεις για το συνολικό διαθέσιμο εισόδημα X και τη συνολική κατανάλωση Y δέκα οικογενειών. Τα συγκεκριμένα στοιχεία δίνονται στον πίνακα 1.1. Το ερώτημα είναι αν υπάρχει γραμμική συσχέτιση μεταξύ του διαθέσιμου εισοδήματος και της κατανάλωσης των οικογενειών αυτών και ποιο είναι το αλγεβρικό μέτρο της συσχέτισης αυτής. ΕΡΩΤΗΜΑ.

Επειδή το μέγεθος του δείγματος είναι μικρό, μπορούμε να κατασκευάσουμε το διάγραμμα των παρατηρήσεων (X_i, Y_i) $i = 1, 2, \ldots, 10$ και να έχουμε μια πρώτη εικόνα για το είδος και το βαθμό της συσχέτισης μεταξύ των μεταβλητών X και Y. Έτσι, από το διάγραμμα του σχήματος 1.2 γίνεται φανερό ότι τα σημεία (X_i, Y_i) συγκεντρώνονται γύρω από μια ευθεία με θετική κλίση και αυτό σημαίνει ότι υπάρχει θετική γραμμική συσχέτιση μεταξύ του διαθέσιμου εισοδήματος και της κατανάλωσης των 10 οικογενειών.

Y'_i: economet1 $\quad$ economet1 = 5,~~99~~15 + 0,484848 economet.
X'_i: economet.

ΠΙΝΑΚΑΣ 1.1: Στοιχεία για τον υπολογισμό του συντελεστή απλής γραμμικής συσχέτισης των μεταβλητών X και Y.

n (1)	X_i (2)	Y_i (3)	x_i (4)	y_i (5)	x_i^2 (6)	y_i^2 (7)	x_iy_i (8)	X_i^2 (9)	Y_i^2 (10)	X_iY_i (11)
1	16	14	-18	-8	324	64	144	256	196	224
2	20	13	-14	-9	196	81	126	400	169	260
3	24	18	-10	-4	100	16	40	576	324	432
4	28	19	- 6	-3	36	9	18	784	361	532
5	32	22	- 2	0	4	0	0	1024	484	704
6	36	23	2	1	4	1	2	1296	529	828
7	40	24	6	2	36	4	12	1600	576	960
8	44	28	10	6	100	36	60	1936	784	1232
9	48	30	14	8	196	64	112	2304	900	1440
10	52	29	18	7	324	49	126	2704	841	1508
n=10	ΣX_i = 340	ΣY_i = 220			Σx_i^2 = 1320	Σy_i^2 = 324	Σx_iy_i = 640	ΣX_i^2 = 12880	ΣY_i^2 = 5164	ΣX_iY_i = 8120
	$\bar{X}$=34	$\bar{Y}$=22								



Σχήμα 1.2: Διάγραμμα των παρατηρήσεων του διαθέσιμου εισοδήματος X και της κατανάλωσης Y των δέκα οικογενειών σε αρχικές τιμές (X_i,Y_i) και σε αποκλίσεις από τους μέσους (x_i,y_i).

Ο αλγεβρικός υπολογισμός του μέτρου της θετικής συσχέτισης των μεταβλητών X και Y γίνεται ως εξής:

Υπολογίζουμε τους μέσους $\bar{X}$ και $\bar{Y}$ των παρατηρήσεων του δείγματος και εκφράζουμε τις τιμές X_i και Y_i των μεταβλητών X και Y σε αποκλίσεις από τους αντίστοιχους μέσους τους:

$$\boxed{x_i = X_i - \bar{X} \quad \text{και} \quad y_i = Y_i - \bar{Y}.}$$

αποκλίσεις από τους αντίστοιχους μέσους τους.

Οι τιμές x_i και y_i δίνονται στις στήλες (4) και (5) του πίνακα 1.1. Η έκφραση των τιμών των μεταβλητών X και Y σε αποκλίσεις από τους μέσους ισοδυναμεί με παράλληλη μετατόπιση των αξόνων X και Y του σχήματος 1.2, έτσι ώστε να διέρχονται από τα σημεία $\bar{Y}$ και $\bar{X}$ αντίστοιχα. Το νέο σύστημα αξόνων έχει ως αρχή το σημείο $(\bar{X},\bar{Y})$ και διαιρεί το επίπεδο στα τεταρτημόρια I, II, III και IV. Αν ένα σημείο βρίσκεται στο πρώτο ή στο τρίτο τεταρτημόριο τότε το γινόμενο των συντεταγμένων του x_iy_i, στο νέο σύστημα αξόνων, είναι θετικό ενώ, αν βρίσκεται στο δεύτερο ή στο τέταρτο τεταρτημόριο, το γινόμενο x_iy_i είναι αρνητικό.

θεωρούμε το άθροισμα των γινομένων $x_i y_i$ για το σύνολο των 10 παρατηρήσεων:

$$\boxed{\sum_{i=1}^{10} x_i y_i .}$$

Αν οι μεταβλητές X και Y συσχετίζονται θετικά τότε, σε θετικές ή αρνητικές τιμές των x_i θα αντιστοιχούν επίσης θετικές ή αρνητικές τιμές των y_i και τα σημεία (x_i, y_i) θα βρίσκονται στα τεταρτημόρια (I) και (III). Στην περίπτωση αυτή τα γινόμενα $x_i y_i$ καθώς και το άθροισμα $\Sigma x_i y_i$ θα είναι θετικά. Αντίθετα, αν οι μεταβλητές X και Y συσχετίζονται αρνητικά, τότε, τα σημεία (x_i, y_i) θα βρίσκονται στα τεταρτημόρια (II) και (IV), οι συντεταγμένες x_i και y_i θα είναι ετερόσημοι αριθμοί και τα γινόμενα $x_i y_i$, καθώς και το άθροισμα $\Sigma x_i y_i$, θα είναι αρνητικά. Αν δεν υπάρχει συσχέτιση μεταξύ των μεταβλητών, τότε, τα σημεία $(x_i y_i)$ θα βρίσκονται και στα τέσσερα τεταρτημόρια και το άθροισμα $\Sigma x_i y_i$ θα τείνει στο μηδέν.

Από τα παραπάνω γίνεται φανερό ότι η αλγεβρική τιμή του αθροίσματος $\Sigma x_i y_i$ μπορεί να χρησιμεύσει ως μέτρο της γραμμικής συσχέτισης των μεταβλητών X και Y.

Όμως η χρησιμοποίηση του αθροίσματος $\Sigma x_i y_i$ ως αλγεβρικού μέτρου της γραμμικής συσχέτισης των X και Y παρουσιάζει τα εξής μειονεκτήματα: (i) εξαρτάται από το μέγεθος του δείγματος διότι, όσο αυξάνεται ο αριθμός των σημείων (x_i, y_i), αυξάνεται και η απόλυτη τιμή του αθροίσματος $\Sigma x_i y_i$, ανεξάρτητα από το αν η συσχέτιση γίνεται υψηλότερη ή όχι, και (ii) εξαρτάται από τις μονάδες μέτρησης των μεταβλητών X και Y αφού, αν υποδιπλασιάσουμε π.χ. τις μονάδες μέτρησης των X και Y, οι τιμές των x_i και y_i διπλασιάζονται και τα γινόμενα $x_i y_i$ άρα και το άθροισμα $\Sigma x_i y_i$, τετραπλασιάζεται.

Το πρώτο μειονέκτημα διορθώνεται αν, διαιρέσουμε το άθροισμα $\Sigma x_i y_i$ με το μέγεθος n του δείγματος:

Cov. r_{XY} $= \boxed{\dfrac{\Sigma x_i y_i}{n}.}$ (1.2.1)

$* \sqrt{\Sigma x_i^2} \cdot \sqrt{\Sigma y_i^2} = \sqrt{\Sigma x_i^2 \cdot \Sigma y_i^2}$

Το δεύτερο μειονέκτημα διορθώνεται αν "τυποποιήσουμε" τις τιμές των x_i και y_i, δηλαδή αν τις διαιρέσουμε με την αντίστοιχη μέση απόκλιση τετραγώνου:

$$\boxed{x_i^* = \frac{x_i}{S_X}} \quad \text{και} \quad \boxed{y_i^* = \frac{y_i}{S_Y}} \qquad (1.2.2)$$

όπου

$$\boxed{S_X = \sqrt{\frac{\Sigma x_i^2}{n}}} \quad \text{και} \quad \boxed{S_Y = \sqrt{\frac{\Sigma y_i^2}{n}}} \qquad (1.2.3)$$

είναι η μέση απόκλιση τετραγώνου των X και Y αντίστοιχα.

Μετά τις διορθώσεις αυτές, το αλγεβρικό μέτρο της απλής γραμμικής συσχέτισης των μεταβλητών X και Y παίρνει τη μορφή:

$$\boxed{r_{XY} = \frac{\Sigma x_i^* y_i^*}{n} = \frac{(\Sigma x_i y_i)/n}{\sqrt{(\Sigma x_i^2)/n}\ \sqrt{(\Sigma y_i^2)/n}}} \qquad (1.2.4)$$

ή

$$\boxed{r_{XY} = \frac{\Sigma x_i y_i}{\sqrt{\Sigma x_i^2}\ \sqrt{\Sigma y_i^2}} \ *} \qquad (1.2.5)$$

και ονομάζεται "συντελεστής συσχέτισης" των μεταβλητών X και Y.

Από τη σχέση (1.2.4) προκύπτει ότι

Cov : συσχέτιση (εξίσ των X κ Y)
V : απόκλιση τετραγώνου.

$$r_{XY} = \frac{Cov(X,Y)}{\sqrt{V(X)V(Y)}} \qquad (1.2.6)$$

$$= \frac{E(XY) - E(X)E(Y)}{\sqrt{V(X)V(Y)}} \qquad (1.2.7)$$

Αν οι τυχαίες μεταβλητές X και Y είναι ανεξάρτητες τότε

$$\boxed{E(XY) = E(X)E(Y)}$$

έλεγχος ανεξαρτησίας των μεταβλητών.

και από τη σχέση (1.2.7) προκύπτει ότι $r_{XY} = 0$. Το αντίστροφο βέβαια δε συμβαίνει: αν $r_{XY} = 0$ τότε οι X και Y δεν είναι οπωσδήποτε ανεξάρτητες. Αν $r_{XY} = 0$ οι μεταβλητές X και Y ονομάζονται "ασυσχέτιστες". Άρα οι έννοιες "ασυσχέτιστες" και "ανεξάρτητες" δεν είναι ισοδύναμες.

Όπως προκύπτει από τα στοιχεία του πίνακα 1.1, η τιμή του συντελεστή συσχέτισης του διαθέσιμου εισοδήματος και της κατανάλωσης, για το συγκεκριμένο δείγμα των 10 οικογενειών, είναι

$$r_{XY}=\frac{640}{\sqrt{324\cdot 1320}}=0.979\ .$$ correlation.

Στο σημείο αυτό πρέπει να τονίσουμε ότι ο συντελεστής συσχέτισης είναι ένας απλός αριθμός χωρίς διαστάσεις και η τιμή του δεν εξαρτάται ούτε από τις μονάδες μέτρησης των X και Y ούτε από την αρχή με βάση την οποία γίνεται η μέτρηση. Πράγματι, μπορούμε εύκολα να αποδείξουμε (να δειχτεί ως άσκηση) ότι αν r_{XY} είναι ο συντελεστής συσχέτισης των X και Y και αν $V=\alpha X+\beta$ και $W=\gamma Y+\delta$ ($\alpha,\beta,\gamma,\delta$ σταθερές, $\alpha\gamma\neq 0$) τότε

$$\boxed{r_{VW}=\frac{\alpha\gamma}{|\alpha\gamma|}\,r_{XY}.}$$

Οι τιμές των μεταβλητών V και W προκύπτουν από τις τιμές των μεταβλητών X και Y αν αλλάξουμε τις μονάδες μέτρησης (πολλαπλασιάζοντας τις τιμές των X και Y επί α και γ αντίστοιχα) και μεταφέροντας την αρχή των αξόνων X και Y στα σημεία -β και -δ αντίστοιχα.

Η σχέση (1.2.5) εκφράζει το συντελεστή συσχέτισης συναρτήσει των τιμών των μεταβλητών X και Y σε αποκλίσεις από τους αντίστοιχους μέσους τους. Αν όμως στη σχέση αυτή θέσουμε $x_i=X_i-X$ και $y_i=Y_i-Y$, εύκολα αποδεικνύεται (να δειχτεί ως άσκηση) ότι ο συντελεστής συσχέτισης εκφράζεται και συναρτήσει των αρχικών τιμών X_i και Y_i των X και Y από τη σχέση:

$$r_{XY}=\frac{n\Sigma X_iY_i-(\Sigma X_i)(\Sigma Y_i)}{\sqrt{n\Sigma X_i^2-(\Sigma X_i)^2}\ \sqrt{n\Sigma Y_i^2-(\Sigma Y_i)^2}}\ . \qquad (1.2.8)$$

Σύμφωνα με τη γνωστή ανισότητα του Schwarz ισχύει:

$$(\Sigma x_iy_i)^2\leq(\Sigma x_i^2)(\Sigma y_i^2)$$

ή

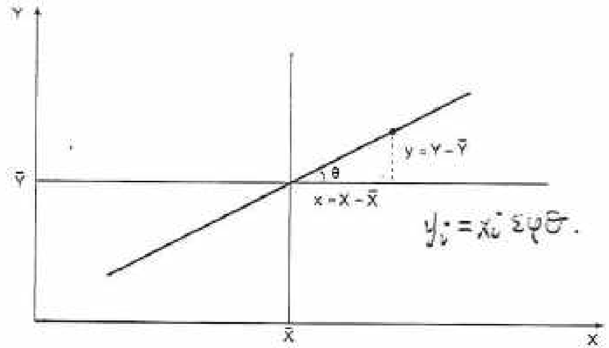
$$\frac{(\Sigma x_iy_i)^2}{(\Sigma x_i^2)(\Sigma y_i^2)}\leq 1.$$

Άρα

$$-1\leq\frac{\Sigma x_iy_i}{\sqrt{\Sigma x_i^2}\ \sqrt{\Sigma y_i^2}}\leq +1, \qquad (1.2.9)$$

από την οποία γίνεται φανερό ότι $-1\leq r_{XY}\leq +1$, δηλαδή, ο συντελεστής συσχέτισης παίρνει τιμές από -1 ως +1.

Όταν η τιμή του r_{XY} είναι θετική αυτό σημαίνει ότι οι μεταβλητές X και Y μεταβάλλονται προς την ίδια κατεύθυνση και όσο η τιμή του r_{XY} πλησιάζει προς το +1 τα σημεία του διαγράμματος συγκεντρώνονται κοντά σε κάποια ευθεία με θετική κλίση. Αν η τιμή του r_{XY} γίνει ίση με +1 τότε όλα τα σημεία βρίσκονται πάνω στην ίδια ευθεία η οποία έχει θετική κλίση και αυτό σημαίνει ότι υπάρχει ακριβής θετική γραμμική



Σχήμα 1.3: Πλήρης θετική γραμμική συσχέτιση.

σχέση μεταξύ των X και Y. Πράγματι, αν όλα τα σημεία του διαγράμματος βρίσκονται πάνω στην ίδια ευθεία τότε, όπως προκύπτει από το σχήμα (1.3), $y_i=x_i\,\varepsilon\varphi\theta$ και

$$r_{XY}=\frac{\Sigma x_iy_i}{\sqrt{\Sigma x_i^2}\ \sqrt{\Sigma y_i^2}}$$

$$=\frac{\Sigma x_i^2\,\varepsilon\varphi\theta}{\sqrt{\Sigma x_i^2}\ \sqrt{\Sigma x_i^2\,\varepsilon\varphi^2\theta}}$$

$$= \frac{\varepsilon\varphi\theta\Sigma x_i^2}{\varepsilon\varphi\theta\Sigma x_i^2}$$

$$=1.$$

Αντίθετα, όταν η τιμή του r_{XY} είναι αρνητική αυτό σημαίνει ότι οι μεταβλητές Χ και Υ μεταβάλλονται προς αντίθετες κατευθύνσεις και όσο η τιμή του r_{XY} πλησιάζει προς το -1 τα σημεία του διαγράμματος συγκεντρώνονται κοντά σε κάποια ευθεία με αρνητική κλίση. Αν η τιμή του r_{XY} γίνει ίση με -1 τότε όλα τα σημεία βρίσκονται πάνω στην ίδια ευθεία η οποία έχει αρνητική κλίση και υπάρχει ακριβής αρνητική γραμμική σχέση μεταξύ των Χ και Υ.

ΑΞΙΑ ΤΟΥ ΣΥΝΤΕΛΕΣΤΗ ΣΥΣΧΕΤΙΣΗΣ.

Σχετικά με την αξία του συντελεστή συσχέτισης ως μέτρου του βαθμού αλληλοεξάρτησης δύο μεταβλητών, πρέπει να παρατηρήσουμε τα εξής:

i) Ο συντελεστής συσχέτισης r_{XY} προσδιορίζει "αποκλειστικά το μέτρο της γραμμικής συσχέτισης" των μεταβλητών Χ και Υ. Αν η τιμή του είναι 0 αυτό σημαίνει ότι δεν υπάρχει γραμμική συσχέτιση μεταξύ των δύο μεταβλητών. Αυτό όμως δεν αποκλείει τη δυνατότητα να υπάρχει μη γραμμική συσχέτιση μεταξύ τους.

ii) Ο υπολογισμός του r_{XY} δεν προσδιορίζει την ευθεία γύρω από την οποία συγκεντρώνονται τα σημεία του διαγράμματος, δηλαδή δεν προσδιορίζει την κλίση και το σταθερό όρο της ευθείας ακόμα και στην περίπτωση που η τιμή του r_{XY} είναι +1 ή -1 που σημαίνει ότι όλα τα σημεία του διαγράμματος συγκεντρώνονται πάνω στην ίδια ευθεία. Η ευθεία αυτή μπορεί να είναι οποιαδήποτε ευθεία του επιπέδου με θετική ή αρνητική κλίση αντίστοιχα.

iii) Ο συντελεστής συσχέτισης δίνει το μέτρο της γραμμικής συσχέτισης των Χ και Υ αλλά δεν προσδιορίζει την αιτιώδη σχέση που τις συνδέει, δηλαδή δεν προσδιορίζει ποια είναι η εξαρτημένη και ποια είναι η ανεξάρτητη μεταβλητή. Έτσι, είναι δυνατόν, η Χ να επηρεάζει την Υ, ή αντίστροφα, ή και οι δύο να συμμεταβάλλονται διότι εξαρτώνται από μια τρί-

τη μεταβλητή ή, τέλος, η συσχέτιση που βρέθηκε στο δείγμα να είναι τυχαία. Αυτό πρέπει να το έχουμε πάντα υπόψη όταν ερμηνεύουμε την τιμή του συντελεστή συσχέτισης διότι αλλιώς υπάρχει κίνδυνος να οδηγηθούμε σε εσφαλμένα συμπεράσματα. Π.χ. αν ο συντελεστής συσχέτισης μεταξύ του ύψους των μισθών των αστυνομικών και της κατανάλωσης οινοπνευματωδών ποτών είναι 0.95, αυτό δε σημαίνει ούτε ότι οι αστυνομικοί είναι μεγάλοι καταναλωτές οινοπνευματωδών ποτών ούτε ότι η αύξηση της κατανάλωσης οινοπνευματωδών ποτών προκαλεί αύξηση του μισθού των αστυνομικών. Απλώς η συσχέτιση των δύο μεγεθών οφείλεται στο ότι και τα δύο εξαρτώνται από τις μεταβολές κάποιας τρίτης μεταβλητής όπως π.χ. το ύψος του εθνικού εισοδήματος. Π.Χ.

β. Ο συντελεστής συσχέτισης ρ_{XY} στους πληθυσμούς

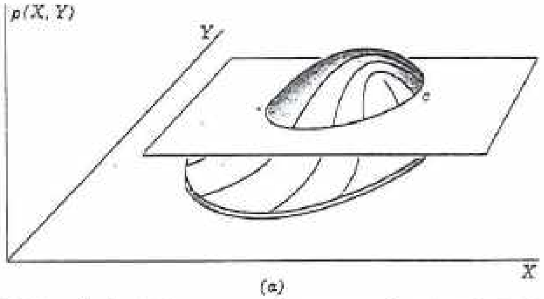
Ο συντελεστής συσχέτισης των Χ και Υ στους πληθυσμούς ορίζεται όπως και ο συντελεστής συσχέτισης στο δείγμα:

$$\boxed{\rho_{XY} = \frac{1}{N}\Sigma x_i^* y_i^*} \qquad (1.2.10)$$

όπου τα x_i^* και y_i^* ορίζονται από τις (1.2.2) και (1.2.3), και Ν είναι το μέγεθος των πληθυσμών.

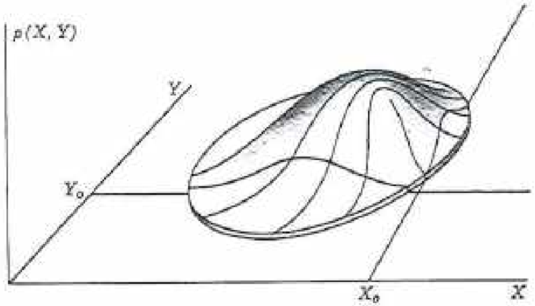
Ενώ ο συντελεστής συσχέτισης είναι ένας συγκεκριμένος αριθμός -άγνωστος στον ερευνητή- ο συντελεστής συσχέτισης του δείγματος είναι μια τυχαία μεταβλητή που η τιμή της αλλάζει από δείγμα σε δείγμα. Έτσι, συχνά, χρειάζεται να προβούμε σε στατιστική επαγωγή και να προσδιορίσουμε κάποιο διάστημα εμπιστοσύνης για την τιμή του ρ_{XY} στους πληθυσμούς από την τιμή του r_{XY} σε ένα συγκεκριμένο δείγμα.

Για να προχωρήσουμε στη στατιστική επαγωγή απαιτούνται ορισμένες υποθέσεις για τη συμπεριφορά των πληθυσμών. Η συνήθης υπόθεση είναι ότι η από κοινού συνάρτηση κατανομής πιθανότητας των Χ και Υ είναι η δικανονική κατανομή (σχήμα 1.4). Αν θεωρήσουμε μια τομή της δικανονικής κατανομής με ένα επίπεδο παράλληλο προς το επίπεδο ΧΥ, η τομή θα είναι μια έλελλείψεις ίσης πιθανότητας

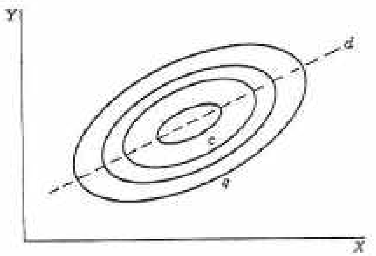


Σχήμα 1.4: Η δικανονική κατανομή πιθανότητας.

λειψη α η οποία θα περιέχει όλα τα σημεία (X_i,Y_i) που έχουν την ίδια πιθανότητα $p(X_i,Y_i)$ να εμφανιστούν (σχήμα 1.5). Τέτοιες ελλείψεις ίσης πιθανότητας (ισοϋψείς) που προκύπτουν από την τομή της πολυκανονικής κατανομής του σχήματος 1.5 σε διάφορα ύψη, δίνονται στο σχήμα 1.6 και είναι η συνήθης



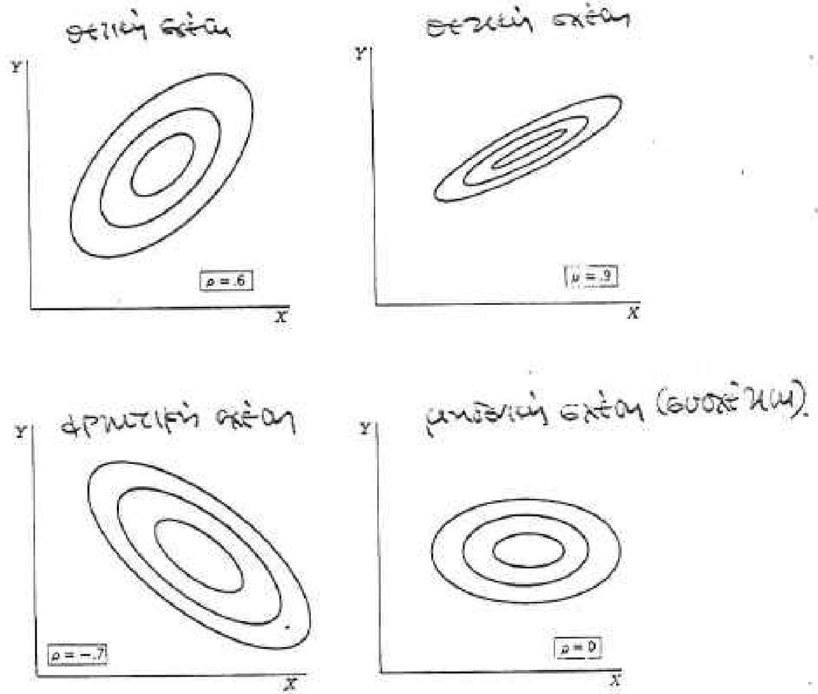
Σχήμα 1.5: Έλλειψη ίσης πιθανότητας από μια δικανονική κατανομή.



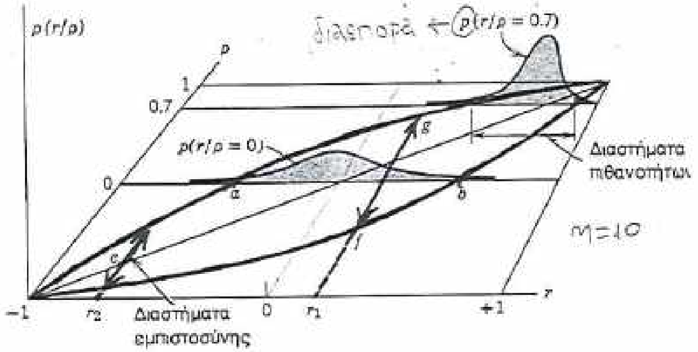
Σχήμα 1.6: Απεικόνιση της δικανονικής κατανομής στο επίπεδο (X,Y) με τη βοήθεια ελλείψεων ίσης πιθανότητας.

πρακτική της απεικόνισης του χώρου των τριών διαστάσεων στο επίπεδο με τη βοήθεια των ισοϋψών καμπύλων. Όλες οι ισοϋψείς ελλείψεις έχουν κοινό το μεγάλο άξονα και όσο η δικανονική κατανομή επιμηκύνεται και συγκεντρώνεται γύρω από τον μεγάλο άξονα των ισοϋψών τόσο η τιμή του ρ_{XY} αυξάνεται. Στο σχήμα 1.7 δίνονται μερικά παραδείγματα πληθυσμών με τις τιμές του συντελεστή συσχέτισης που αντιστοιχεί σαυτούς.

Η κατανομή του r_{XY} δεν είναι η ίδια για όλες τις τιμές του ρ_{XY} και για κάθε μέγεθος δείγματος n. Στο σχήμα 1.8 δίνεται η κατανομή πιθανότητας του r_{XY} για διάφορες τιμές του ρ_{XY} και για μέγεθος δείγματος n=10. Από το σχήμα 1.8 προκύπτει ότι η κατανομή του r_{XY} παρουσιάζει μεγαλύτερη διασπορά για τιμές του ρ_{XY} κοντά στο μηδέν και μικρότερη διασπορά όσο η τιμή του ρ_{XY} πλησιάζει το +1 ή το -1. Αυτό σημαίνει ότι ο συντελεστής συσχέτισης r_{XY} του δείγματος είναι πιο αποτελεσματική εκτίμηση του ρ_{XY} για τιμές του ρ_{XY} κοντά στο +1 ή το -1. Όταν οι μεταβλητές X και Y παρουσιάζουν πλήρη γραμμική συσχέτιση (ρ=1 ή -1), όλα τα σημεία (X_i,Y_i) όλων των ισοϋψών συγκεντρώνονται πάνω στον μεγάλο άξονα και το δείγμα θα δίνει την τιμή $r_{XY}=1$ η οποία θα είναι ακριβής εκτίμηση του ρ_{XY}. Ακόμα, πρέπει να παρατηρήσουμε ότι, η κατανομή του r_{XY} για τιμές του ρ_{XY} κοντά στο μηδέν είναι κανονική, ενώ για τιμές του ρ κοντά στο +1 ή το -1 η κατανομή του r δεν είναι συμμετρική. Γίνεται λοιπόν φανερό ότι, για κάθε τιμή του ρ_{XY} ορίζεται η υπό συνθήκη πιθανότητα του r_{XY}. Αν για κάθε τέτοια κατανομή ορίσουμε τα σημεία a και b ανάμεσα στα οποία βρίσκεται το 95% του εμβαδού της και θεωρήσουμε τις δύο καμπύλες που προκύπτουν αν ενώσουμε όλα τα σημεία a και όλα τα σημεία b, αυτές ορίζουν το 95% διάστημα εμπιστοσύνης για την αληθή τιμή του ρ_{XY} που αντιστοιχεί στις διάφορες τιμές του ρ_{XY} και για μέγεθος δείγματος n=10. Τέτοια διαστήματα εμπιστοσύνης μπορούμε να κατασκευάσουμε για κάθε μέγεθος δείγματος. Στο σχήμα 1.9 δίνονται αρκετά τέτοια διαστήματα εμπιστοσύνης για διάφορα μεγέθη δειγμάτων. Έτσι αν από ένα δείγμα μεγέθους n=10 έχουμε υπολογίσει ότι $r_{XY}=0.9$ τότε το 95%

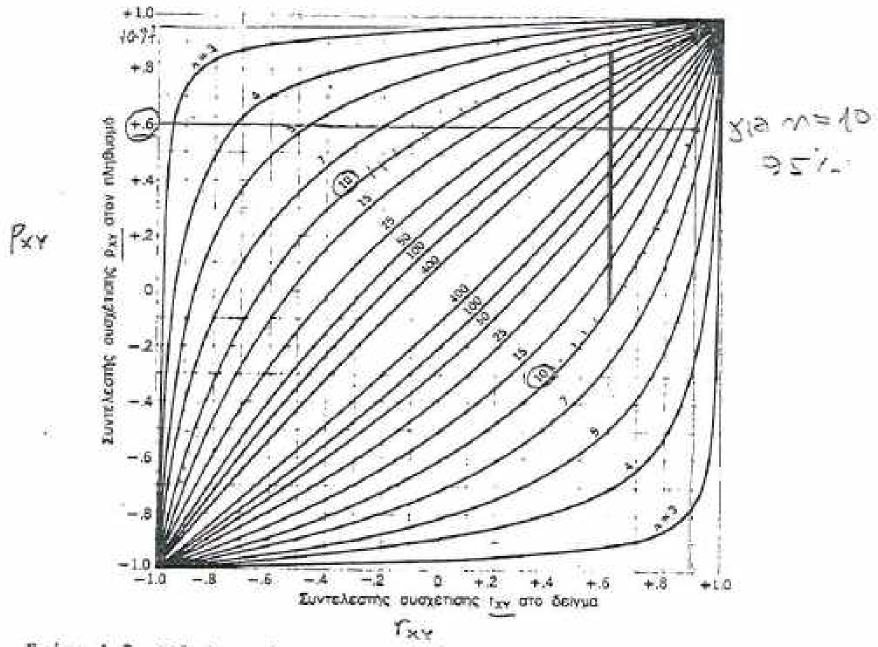


Σχ. 1.7: Παραδείγματα συντελεστών συσχέτισης στους πληθυσμούς.



Σχ. 1.8: Οι υπό συνθήκη κατανομές πιθανότητας $P(r/\rho)$.

το r είναι μια τιμή του ρ.



Σχήμα 1.9: 95% διαστήματα εμπιστοσύνης για την τιμή του ρ_{XY} στον πληθυσμό για διάφορα μεγέθη δείγματος, όταν η από κοινού συνάρτηση κατανομής πιθανότητας των X και Y είναι η δικανονική κατανομή.

διάστημα εμπιστοσύνης για την αληθή τιμή του ρ_{XY} στους πληθυσμούς είναι:

$$\boxed{0.60 < \rho < 0.97}$$ √

και επειδή το μηδέν δεν περιλαμβάνεται στο διάστημα αυτό η υπόθεση

$$H_0: \rho_{XY}=0$$

απορρίπτεται. δηλαδή υπάρχει στατιστικά σημαντική γραμμική συσχέτιση μεταξύ των X και Y στους πληθυσμούς.

Παίρνουμε την κάθετο του $r_{XY}=0.9$ ή βρίσκουμε τις τομές με την άνω και κάτω καμπύλη του n=10. Έτσι προσδιορίζουμε το διάστημα εμπιστοσύνης για την αληθή τιμή ρ_{XY}.

1.3. ΣΥΝΤΕΛΕΣΤΗΣ ΜΕΡΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ

Ας θεωρήσουμε ξανά το παράδειγμα της κατανάλωσης και του διαθέσιμου εισοδήματος. Όπως είδαμε οι δύο μεταβλητές συσχετίζονται θετικά και μάλιστα υψηλά (r_{XY}=0,979). Όμως, είναι φανερό ότι, και οι δύο μεταβλητές εξαρτώνται από τον αριθμό Ζ των μελών σε κάθε οικογένεια. Έτσι, θα μπορούσαμε να έχουμε καλύτερη εικόνα για τη συσχέτιση των μεταβλητών Χ και Υ αν οι δέκα οικογένειες του δείγματος είχαν τον ίδιο αριθμό μελών.

Ο "συντελεστής μερικής συσχέτισης πρώτης τάξης" $r_{XY\cdot Z}$ μετρά το βαθμό της συσχέτισης των Χ και Υ όταν η Ζ παραμένει σταθερή. Αποδεικνύεται ότι

$$\boxed{r_{XY\cdot Z}=\frac{r_{XY}-r_{XZ}r_{YZ}}{\sqrt{(1-r_{XZ}^2)(1-r_{YZ}^2)}}} \qquad (1.3.1)$$

όπου

r_{XY} είναι ο συντελεστής συσχέτισης των Χ και Υ

r_{XZ} είναι ο συντελεστής συσχέτισης των Χ και Ζ.

r_{YZ} είναι ο συντελεστής συσχέτισης των Υ και Ζ.

Ανάλογα ορίζονται και οι συντελεστές μερικής συσχέτισης δεύτερης, τρίτης, ..., k-τάξης, που εκφράζουν το βαθμό της συσχέτισης δύο μεταβλητών όταν η τιμή δύο, τριών, ..., k άλλων μεταβλητών που τις επηρεάζουν παραμένουν σταθερές. Στο θέμα αυτό θα επανέλθουμε στο κεφάλαιο 3 της πολλαπλής παλινδρόμησης.

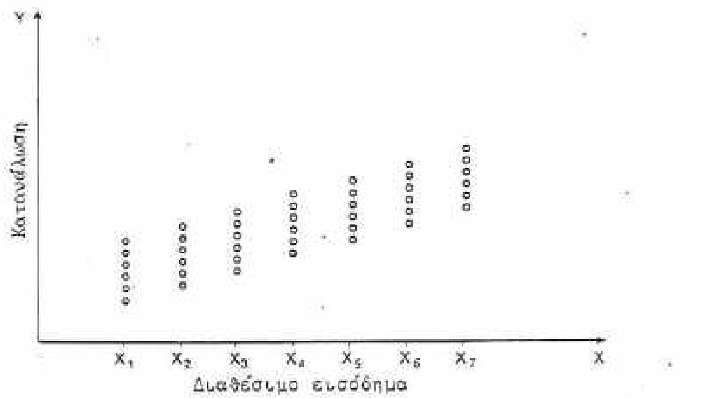
Κεφάλαιο 2

Απλή Παλινδρόμηση

2.1. ΕΙΣΑΓΩΓΗ

Στο παράδειγμα του πρώτου κεφαλαίου είδαμε ότι, η τιμή r=0,979 του συντελεστή συσχέτισης βεβαιώνει μεν ότι υπάρχει υψηλή θετική γραμμική συσχέτιση μεταξύ του διαθέσιμου εισοδήματος και της κατανάλωσης των οικογενειών του δείγματος, αλλά, δεν προσδιορίζει ούτε αν το διαθέσιμο εισόδημα επηρεάζει την κατανάλωση ή αντίστροφα, ούτε την κλίση και το σταθερό όρο της ευθείας γύρω από την οποία συγκεντρώνονται τα σημεία. Όμως στην ποσοτική οικονομική ανάλυση δεν ενδιαφερόμαστε, συνήθως, για τον υπολογισμό του βαθμού της συσχέτισης δύο οικονομικών μεγεθών, αλλά κυρίως, για τον προσδιορισμό της αιτιώδους συναρτησιακής σχέσης που συνδέει τα μεγέθη αυτά.

Έστω ότι επιθυμούμε π.χ. να εξετάσουμε πια είναι η συναρτησιακή σχέση που συνδέει την κατανάλωση με το διαθέσιμο εισόδημα της οικογένειας. Αν η κατανάλωση Υ πολλών οικογενειών που έχουν το ίδιο διαθέσιμο εισόδημα Χ δοθεί για διάφορες περιπτώσεις σε διάγραμμα, τότε, το διάγραμμα αυτό θα έχει πιθανότατα τη μορφή του σχήματος 2.1. Από το διάγραμμα γίνεται φανερό ότι το διαθέσιμο εισόδημα επηρεάζει την κατανάλωση. Στόχος μας είναι να αποδόσουμε την ποσοτική σχέση με μια μαθηματική συνάρτηση η οποία να συνδέει τις αντίστοιχες τιμές των μεταβλητών Χ και Υ. Γεωμετρικά, ο προσδιορισμός της συνάρτησης αυτής ισοδυναμεί με την προσαρμογή μιας καμπύλης στα σημεία του διαγράμματος. Η καμπύλη αυτή καλεί-



Σχήμα 2.1: Η κατανάλωση πολλών οικογενειών που έχουν το ίδιο διαθέσιμο εισόδημα για διάφορες τιμές του διαθέσιμου εισοδήματος.

ται "παλίνδρόμηση της Y πάνω στη X" και ως μαθηματική συνάρτηση είναι χρήσιμη για τη σύντομη και ακριβή περιγραφή της σχέσης των Y και X και ακόμη μπορεί να προσδιορίσει το ύψος της κατανάλωσης μιας οικογένειας αν δοθεί το διαθέσιμο εισόδημά της.

Στο κεφάλαιο αυτό θα ασχοληθούμε με την "απλή γραμμική παλινδρόμηση" δηλαδή με τις μεθόδους προσαρμογής αποκλειστικά μιας ευθείας γραμμής στα σημεία του δείγματος. Είναι δυνατόν βέβαια η μεταβλητή Y να συνδέεται με τη X με μια πιο σύνθετη γραμμική ή μη γραμμική σχέση. Το θέμα όμως αυτό θα μας απασχολήσει στα κεφάλαια 3 και 4. Προς το παρόν υποθέτουμε ότι η κατάλληλη καμπύλη είναι μια ευθεία γραμμή της οποίας επιθυμούμε να προσδιορίσουμε αλγεβρικά την κλίση και το σταθερό όρο.

2.2. Η ΑΠΛΗ ΓΡΑΜΜΙΚΗ ΠΑΛΙΝΔΡΟΜΗΣΗ. Η ΜΕΘΟΔΟΣ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

Ας θεωρήσουμε πάλι το δείγμα των παρατηρήσεων για την κατανάλωση Y και το διαθέσιμο εισόδημα X των 10 οικογενειών του πρώτου κεφαλαίου και ας υποθέσουμε ότι η κατανάλωση είναι γραμμική συνάρτηση του διαθέσιμου εισοδήματος. Η υπόθεση αυτή εκφράζει απλώς την επιθυμία μας να προσαρμόσουμε στα σημεία του δείγματος μια ευθεία χωρίς αυτό να σημαίνει ότι είμαστε βέβαιοι ότι η πραγματική σχέση που συνδέει την Y και την X είναι γραμμική. Επειδή η μεταβλητή Y (κατανάλωση) εξαρτάται από το διαθέσιμο εισόδημα, θα την ονομάσουμε "εξαρτημένη" μεταβλητή. Το διαθέσιμο εισόδημα X θεωρούμε ότι δεν επηρεάζεται από την κατανάλωση και επειδή ο ερευνητής μπορεί να επιλέξει κατά την κρίση του το ύψος του εισοδήματος των οικογενειών, των οποίων θα μελετήσει την καταναλωτική συμπεριφορά, θα το ονομάσουμε "ανεξάρτητη" ή "ερμηνευτική" μεταβλητή. Στόχος μας είναι να προσαρμόσουμε στα σημεία του δείγματος την καλλίτερη ευθεία.

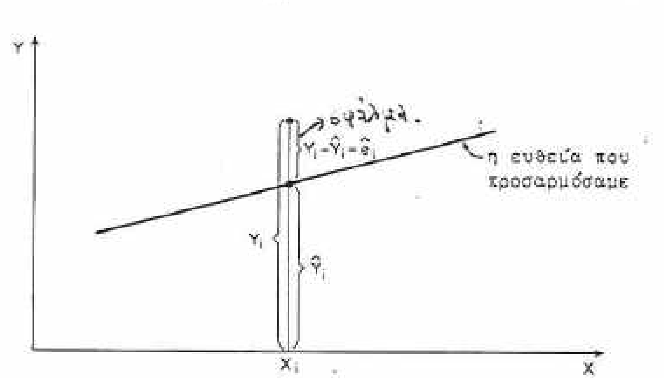
ΑΝΕΞΑΡΤΗΤΗ ΕΞΑΡΤΗΜΕΝΗ ΜΕΤΑΒΛΗΤΗ

Στο ερώτημα ποια είναι η καλλίτερη ευθεία, η απάντηση είναι απλή: αυτή που δίνει τα μικρότερα σφάλματα. Ένα τέτοιο σφάλμα δίνεται στο σχήμα 2.2 και ορίζεται ως η κάθετη απόσταση της παρατηρούμενης τιμής Y_i από την προσαρμοσμένη ευθεία, δηλαδή η απόσταση:

ΟΡΙΣΜΟΣ ΣΦΑΛΜΑΤΟΣ

$$\boxed{\hat{e}_i = Y_i - \hat{Y}_i} \qquad (2.2.1)$$

όπου $\hat{Y}_i$ είναι η τιμή της Y_i που προκύπτει από την προσαρμοσμένη ευθεία. Το σφάλμα $\hat{e}_i$ είναι θετικό αν η παρατηρούμενη τιμή Y_i βρίσκεται πάνω από την ευθεία που προσαρμόσαμε και αρνητικό αν η Y_i βρίσκεται κάτω από την ευθεία.



Σχήμα 2.2: Το σφάλμα από την προσαρμογή μιας ευθείας.

Υπάρχουν πολλά κριτήρια για την ελαχιστοποίηση των σφαλμάτων. Έτσι, μπορούμε να ελαχιστοποιήσουμε το άθροισμα όλων των σφαλμάτων

1) $$\sum_{i=1}^{10} \hat{e}_i = \sum_{i=1}^{10} (Y_i - \hat{Y}_i) \qquad (2.2.2)$$

ή το άθροισμα των απολύτων τιμών των σφαλμάτων

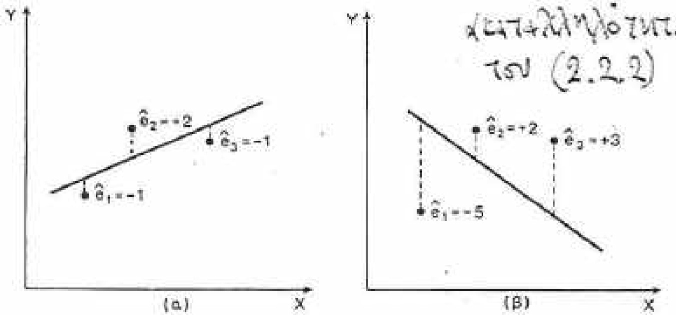
2) $$\sum_{i=1}^{10} |\hat{e}_i| = \sum_{i=1}^{10} |Y_i - \hat{Y}_i| \qquad (2.2.3)$$

ή το άθροισμα των τετραγώνων των σφαλμάτων

3) $$\sum_{i=1}^{10} \hat{e}_i^2 = \sum_{i=1}^{10} (Y_i - \hat{Y}_i)^2 \; \checkmark \qquad (2.2.4)$$

Τα δύο πρώτα κριτήρια δεν είναι ικανά πάντα να οδηγήσουν στην καλλίτερη προσαρμογή. Όπως προκύπτει από το σχήμα 2.3 το

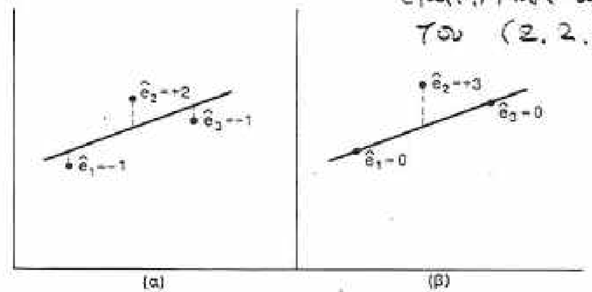
ακαταλληλότητα του (2.2.2)



Σχήμα 2.3: Η ακαταλληλότητα του κριτηρίου $\Sigma(Y_i-\hat{Y}_i)$ για την προσαρμογή της καλλίτερης ευθείας.

κριτήριο (2.2.2) ικανοποιείται και από τις δύο προσαρμογές (α) και (β), αφού και στις δύο περιπτώσεις το άθροισμα των σφαλμάτων είναι μηδέν, ενώ είναι φανερό ότι η προσαρμογή (α) είναι καλλίτερη από την (β). Επίσης, όπως προκύπτει από το σχήμα 2.4, το κριτήριο (2.2.3) ικανοποιείται καλλίτερα από την προσαρμογή (β) ενώ είναι φανερό ότι η προσαρμογή (α) είναι καλλίτερη. Αντίθετα, το τρίτο κριτήριο, γνωστό ως "κριτήριο ελαχίστων τετραγώνων", οδηγεί, πάντοτε όπως θα δείξουμε αργότερα, στην καλλίτερη προσαρμογή. Θα χρησιμοποιήσουμε λοιπόν το κριτήριο ελαχίστων τετραγώνων (Ε.Τ.) για να προσαρμόσουμε στις παρατηρήσεις του δείγματος την καλλίτερη ευθεία:

ακαταλληλότητα του (2.2.3).



Σχήμα 2.4: Η ακαταλληλότητα του κριτηρίου $\Sigma|Y_i-\hat{Y}_i|$ για την προσαρμογή της καλλίτερης ευθείας.

$$\boxed{\hat{Y}_i = \hat{\alpha}_0 + \hat{b}X_i, \quad i=1,2,\ldots,n} \qquad (2.2.5)$$

όπου

$\hat{Y}_i$: είναι η τιμή της Y_i από την παλινδρόμηση που εκτιμήσαμε,

$\hat{\alpha}_0$: η εκτίμηση του σταθερού όρου της προσαρμογής,

$\hat{b}$: η εκτίμηση της κλίσης της προσαρμογής,

n : το μέγεθος του δείγματος.

Ε.Τ.: ελάχιστα τετράγωνα.

Αν

$$\boxed{\hat{e}_i = Y_i - \hat{Y}_i} \qquad (2.2.6)$$

είναι το σφάλμα από την προσαρμογή της ευθείας Ε.Τ., τότε, το κριτήριο επιλέγει τις τιμές $\hat{\alpha}_0$ και $\hat{b}$ του σταθερού όρου και της κλίσης που ελαχιστοποιούν το άθροισμα των τετραγώνων:

$$\begin{aligned}\sum_{i=1}^{n} \hat{e}_i^2 &= \sum_{i=1}^{n} (Y_i - \hat{Y}_i)^2 \\ &= \sum_{i=1}^{n} (Y_i - \hat{\alpha}_0 - \hat{b}X_i)^2 \\ &= f(\hat{\alpha}_0, \hat{b}) \end{aligned} \qquad (2.2.7)$$

Αν $f_{\hat{\alpha}_0}$, $f_{\hat{b}}$, $f_{\hat{\alpha}_0\hat{\alpha}_0}$, $f_{\hat{b}\hat{b}}$, $f_{\hat{\alpha}_0\hat{b}}$ είναι οι παράγωγοι πρώτης και δεύτερης τάξης της (2.2.7) ως προς $\hat{\alpha}_0$ και $\hat{b}$ τότε για να υπάρχει το ελάχιστο της $f(\hat{\alpha}_0,\hat{b})$ πρέπει να ικανοποιούνται οι συνθήκες πρώτης τάξης:

Συνθήκες 1ης τάξης 1) $f_{\hat{\alpha}_0}=0$ και $f_{\hat{b}}=0$ (2.2.8)

καθώς και οι συνθήκες δεύτερης τάξης:

Συνθήκη δεύτερης τάξης 2) $f_{\hat{\alpha}_0\hat{\alpha}_0}>0$, $f_{\hat{b}\hat{b}}>0$ και $f_{\hat{\alpha}_0\hat{\alpha}_0}f_{\hat{b}\hat{b}}-(f_{\hat{\alpha}_0\hat{b}})^2>0$. (2.2.9)

Οι συνθήκες πρώτης τάξης είναι:

$$f_{\hat{\alpha}_0}=-2\Sigma(Y_i-\hat{\alpha}_0-\hat{b}X_i)=0 \qquad (2.2.10)$$

$$f_{\hat{b}}=-2\Sigma X_i(Y_i-\hat{\alpha}_0-\hat{b}X_i)=0 \qquad (2.2.11)$$

από τις οποίες προκύπτουν οι λεγόμενες "κανονικές εξισώσεις"

$$n\hat{\alpha}_0+(\Sigma X_i)\hat{b}=\Sigma Y_i \qquad (2.2.12)$$

$$(\Sigma X_i)\hat{\alpha}_0+(\Sigma X_i)\hat{b}=\Sigma X_iY_i. \qquad (2.2.13)$$

Οι κανονικές εξισώσεις αποτελούν ένα γραμμικό σύστημα δύο εξισώσεων με δύο άγνωστους και η επίλυσή του προσδιορίζει το σταθερό όρο και την κλίση της ευθείας Ε.Τ.:

$$\boxed{\hat{\alpha}_0=\frac{\Sigma X_i^2\Sigma Y_i-\Sigma X_i\Sigma X_iY_i}{n\Sigma X_i^2-(\Sigma X_i)^2}} \qquad (2.2.14)$$

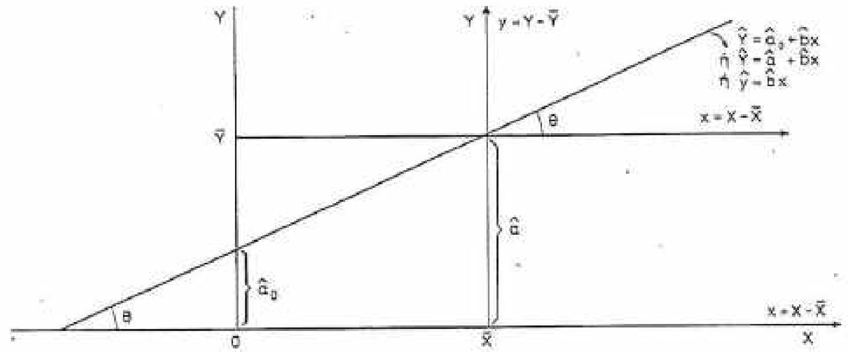
$$\boxed{\hat{b}=\frac{n\Sigma X_iY_i-\Sigma X_i\Sigma Y_i}{n\Sigma X_i^2-(\Sigma X_i)^2}}\,. \qquad (2.2.15)$$

Ο αναγνώστης εύκολα μπορεί να διαπιστώσει ότι οι συνθήκες δεύτερης τάξης (2.2.9) ικανοποιούνται στο σημείο $(\hat{\alpha}_0,\hat{b})$ που προσδιορίζεται από τις συνθήκες πρώτης τάξης.

Όπως προκύπτει από την (2.2.12) ο σταθερός όρος $\hat{\alpha}_0$ εκφράζεται συναρτήσει της κλίσης $\hat{b}$ από την

$$\boxed{\begin{aligned}\hat{\alpha}_0&=\frac{\Sigma Y_i}{n}-\hat{b}\,\frac{\Sigma X_i}{n}\\&=\bar{Y}-\hat{b}\bar{X}.\end{aligned}} \qquad (2.2.16)$$

Μπορούμε να απλουστεύσουμε σημαντικά τους υπολογισμούς που απαιτούνται στις (2.2.14) και (2.2.15) αν εκφράσουμε τις τιμές της μεταβλητής X σε αποκλίσεις από τους μέσους. Όπως προκύπτει από το σχήμα 2.5, η έκφραση των τιμών της X σε αποκλίσεις από το μέσο της $\bar{X}$ ισοδυναμεί με μεταφορά του άξο-



Σχήμα 2.5: Η γεωμετρία της προσαρμογής Ε.Τ. στα τρία συστήματα αξόνων (X,Y), (x,Y) και (x,y).

να των Y στο σημείο $\bar{X}$, δηλαδή με μεταφορά της αρχής των αξόνων από το σημείο (0,0) στο σημείο $(\bar{X},0)$. Στο νέο σύστημα αξόνων (x,Y) ο σταθερός όρος δεν είναι πλέον $\hat{\alpha}_0$, αλλά $\hat{\alpha}$ ενώ η κλίση $\hat{b}$=εφθ δε μεταβάλλεται. Στην προκείμενη περίπτωση ελαχιστοποιούμε το

$$\boxed{\Sigma\hat{e}_i^2=\Sigma(Y_i-\hat{\alpha}-\hat{b}x_i)}$$
$$=f(\hat{\alpha},\hat{b}).$$

Οι κανονικές εξισώσεις είναι τώρα:

$$n\hat{\alpha}+(\Sigma x_i)\hat{b}=\Sigma Y_i$$

$$(\Sigma x_i)\hat{\alpha}+(\Sigma x_i^2)\hat{b}=\Sigma x_iY_i$$

και επειδή

$$\Sigma x_i = \Sigma(X_i - \bar{X})$$
$$= \Sigma X_i - n\bar{X}$$
$$= n\bar{X} - n\bar{X}$$
$$= 0,$$

οι κανονικές εξισώσεις απλουστεύονται στις:

$$n\hat{\alpha} = \Sigma Y_i$$
$$(\Sigma x_i^2)\hat{b} = \Sigma x_i Y_i$$

από τις οποίες προκύπτουν οι εκτιμήσεις Ε.Τ.:

τύπος για την εύρεση του $\hat{\alpha}$.

$$\boxed{\hat{\alpha} = \frac{\Sigma Y_i}{n} = \bar{Y}} \qquad (2.2.17)$$

$$\hat{b} = \frac{\Sigma x_i Y}{\Sigma x_i^2}\,. \qquad (2.2.18)$$

Από τις (2.2.16) και (2.2.17) προκύπτει ότι ο σταθερός όρος $\hat{\alpha}_0$ στο σύστημα αξόνων (X,Y) και ο σταθερός όρος $\hat{\alpha}$ στο σύστημα (x,Y) συνδέονται με τη σχέση:

Σύνδεση α_0 μ' α.

$$\boxed{\hat{\alpha}_0 = \hat{\alpha} - \hat{b}\bar{X}.} \qquad (2.2.19)$$

Αν ενδιαφερόμαστε μόνο για την κλίση της ευθείας Ε.Τ. τότε μπορούμε να εργαστούμε στο σύστημα αξόνων (x,y) εκφράζοντας και τις δύο μεταβλητές X και Y σε αποκλίσεις από τους αντίστοιχους μέσους. Όπως προκύπτει από το σχήμα 2.5 ο σταθερός όρος της ευθείας Ε.Τ. στο σύστημα (x,y) είναι 0 αφού η ευθεία διέρχεται από την αρχή των αξόνων $(\bar{X},\bar{Y})$. Συνεπώς, απομένει για εκτίμηση μόνο η κλίση $\hat{b}$. Εδώ έχουμε

$$\Sigma \hat{e}_i^2 = \Sigma(y_i - \hat{y}_i)^2$$
$$= \Sigma(y_i - \hat{b}x_i)^2$$
$$= f(\hat{b})$$

και η συνθήκη πρώτης τάξης για την ύπαρξη του ελαχίστου

$$f_{\hat{b}} = -2\Sigma x_i(y_i - \hat{b}x_i) = 0$$

δίνει την εκτίμηση της κλίσης:

$$\hat{b} = \frac{\Sigma x_i y_i}{\Sigma x_i^2} \qquad (2.2.20)$$

$\alpha = \bar{Y} - b\bar{X}$

Όπως προκύπτει από τις (2.2.15), (2.2.18) και (2.2.20)

τύπος για την εύρεση του $\hat{b}$.

$$\boxed{\hat{b} = \frac{n\Sigma X_i Y_i - \Sigma X_i \Sigma Y_i}{n\Sigma X_i^2 - (\Sigma X_i)^2} = \frac{\Sigma x_i Y_i}{\Sigma x_i^2} = \frac{\Sigma x_i y_i}{\Sigma x_i^2}} \qquad (2.2.21)$$

αφού και οι τρεις σχέσεις εκφράζουν την κλίση της ευθείας Ε.Τ. (1) η πρώτη συναρτήσει των αρχικών τιμών των μεταβλητών, (2) η δεύτερη συναρτήσει των αρχικών τιμών της Y και των αποκλίσεων της X από το μέσο της και (3) η τρίτη συναρτήσει των τιμών και των δύο μεταβλητών σε αποκλίσεις από τους μέσους τους.

Για το συγκεκριμένο παράδειγμα της κατανάλωσης και του διαθέσιμου εισοδήματος των 10 οικογενειών εύκολα υπολογίζουμε από τα στοιχεία του πίνακα 2.1, ότι:

$$\hat{\alpha} = 22 \quad \text{και} \quad \hat{b} = 0,485.$$

Άρα η ευθεία Ε.Τ. για το συγκεκριμένο δείγμα είναι η

$$\hat{Y} = 22 + 0,485x. \qquad (2.2.22)$$

Αν επιθυμούμε, μπορούμε να εκφράσουμε την κατανάλωση Y συναρτήσει των αρχικών τιμών του διαθέσιμου εισοδήματος X

22: κατανάλωση επιβίωσης για $X=\bar{X}=34$. (αυτόνομη κατανάλωση).

$$\hat{Y} = 22 + 0,485(X - \bar{X})$$
$$= 22 + 0,485(X - 34) \qquad , (\bar{X}=34)$$
$$= (22 - 0,485 \cdot 34) + 0,485X$$

5,51: κατανάλωση επιβίωσης για $X=0$

$$= 5,51 + 0,485X. \qquad (2.2.23)$$

Η κλίση των ευθειών (2.2.22) και (2.2.23) είναι η ίδια και εκφράζει την εκτίμηση της οριακής ροπής προς κατανάλωση:

ΠΙΝΑΚΑΣ 2.1: Στοιχεία για τον υπολογισμό των εκτιμητριών ελαχίστων τετραγώνων του απλού γραμμικού υποδείγματος.

n (1)	X_i (2)	Y_i (3)	x_i (4)	y_i (5)	x_i^2 (6)	y_i^2 (7)	x_iy_i (8)	X_i^2 (9)	Y_i^2 (10)	X_iY_i (11)	x_iY_i (12)
1	16	14	-18	-8	324	64	144	256	196	224	-252
2	20	13	-14	-9	196	81	126	400	169	260	-182
3	24	18	-10	-4	100	16	40	576	324	432	-180
4	28	19	- 6	-3	36	9	18	784	361	532	-114
5	32	22	- 2	0	4	0	0	1024	484	704	- 44
6	36	23	2	1	4	1	2	1296	529	828	46
7	40	24	6	2	36	4	12	1600	576	960	144
8	44	28	10	6	100	36	60	1936	784	1232	280
9	48	30	14	8	196	64	112	2304	900	1440	420
10	52	29	18	7	324	49	126	2704	841	1508	522
n=10	$\Sigma X_i=$ 340	$\Sigma Y_i=$ 220			$\Sigma x_i^2=$ 1320	$\Sigma y_i^2=$ 324	$\Sigma x_iy_i=$ 640	$\Sigma X_i^2=$ 12880	$\Sigma Y_i^2=$ 64	$\Sigma X_iY_i=$ 8120	$\Sigma x_iY_i=$ 640
	$\bar{X}=34$	$\bar{Y}=22$									

$\hat{\alpha}=\bar{Y}=22,$

$$\hat{b}=\frac{n\Sigma X_iY_i-\Sigma X_i\Sigma Y_i}{n\Sigma X_i^2-(\Sigma X_i)^2}=\frac{\Sigma x_iY_i}{\Sigma x_i^2}=\frac{\Sigma x_iy_i}{\Sigma x_i^2}=0,485$$

$\hat{\alpha}_0=\hat{\alpha}-\hat{b}\bar{X}=5,51$

$$\boxed{\frac{\partial\hat{Y}}{\partial X}=0,485.}$$

Αντίθετα ο μεν σταθερός όρος της (2.2.22) εκφράζει την κατανάλωση επιβίωσης του καταναλωτή με διαθέσιμο εισόδημα $\bar{X}$ (του μέσου καταναλωτή) ενώ ο σταθερός όρος της (2.2.23) εκφράζει την κατανάλωση επιβίωσης του καταναλωτή με διαθέσιμο εισόδημα X=0.

Χρησιμοποιώντας τώρα την ευθεία (2.2.23) μπορούμε να υπολογίσουμε την "εκτίμηση" της κατανάλωσης μιας οικογένειας αν γνωρίζουμε το διαθέσιμο εισόδημά της. Έτσι, αν μια οικογένεια έχει διαθέσιμο εισόδημα 30 χιλ. δρχ., η εκτίμηση της κατανάλωσής της είναι:

$$\hat{Y}=5,51+0,485\cdot 30=20,06 \text{ χιλ. δρχ.}$$ Π.Χ.

Η ευθεία (2.2.22) δίνει το ίδιο ακριβώς αποτέλεσμα:

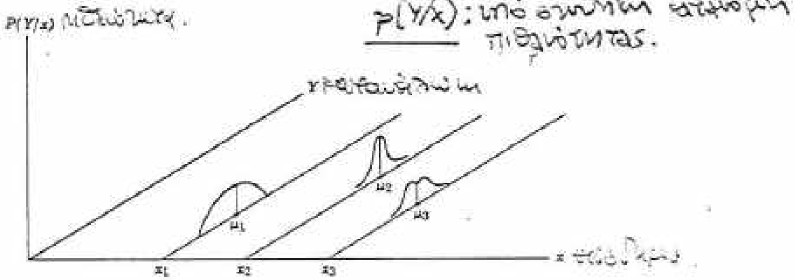
$$\hat{Y}=22+0,485(30-34)=20,06 \text{ χιλ. δρχ.}$$

Στο σημείο αυτό πρέπει να επισημάνουμε ότι, η εφαρμογή της μεθόδου ελαχίστων τετραγώνων για την προσαρμογή της καλλίτερης ευθείας στα σημεία ενός συγκεκριμένου δείγματος, είναι μια μηχανική υπολογιστική διαδικασία. Σε κάθε δείγμα που προέρχεται από τους πληθυσμούς των X και Y μπορούμε να προσαρμόσουμε την ευθεία Ε.Τ. και είναι προφανές ότι, κάθε μια από τις ευθείες αυτές θα έχει διαφορετικό σταθερό όρο και διαφορετική κλίση.

Είναι γνωστό ότι τα οικονομικά φαινόμενα δεν είναι κατευθυνόμενα πειράματα ώστε να μπορούμε να τα επαναλάβουμε όσες φορές θέλουμε διατηρώντας σταθερές τις τιμές της ανεξάρτητης μεταβλητής ή προσδιορίζοντάς τις κατά βούληση. Η συνήθης κατάσταση που αντιμετωπίζουμε είναι η εξής: διαθέτουμε ένα μόνο δείγμα παρατηρήσεων για την εξαρτημένη μεταβλητή Y και την ανεξάρτητη μεταβλητή X και από το μοναδικό αυτό δείγμα πρέπει να εκτιμήσουμε, με τη μέθοδο Ε.Τ., τη γραμμική σχέση που συνδέει την Y με τη X. Έτσι, στο παράδειγμα της κατανάλωσης και του διαθέσιμου εισοδήματος, διαθέτουμε ένα δείγμα παρατηρήσεων για 10 οικογένειες, για τις οποίες έγινε η σχετική έρευνα, και από το δείγμα αυτό, εκτιμήσαμε τη γραμμική σχέση που συνδέει την κατανάλωση με το διαθέσιμο εισόδημα, υπολογίσαμε δηλαδή την κατανάλωση επιβίωσης και την οριακή ροπή προς κατανάλωση των 10 οικογενειών. Σκοπός όμως μιας τέτοιας έρευνας δεν είναι συνήθως η εκτίμηση π.χ. της οριακής ροπής προς κατανάλωση των 10 οικογενειών του συγκεκριμένου δείγματος, αλλά ο προσδιορισμός ενός διαστήματος εμπιστοσύνης μέσα στο οποίο θα βρίσκεται, με κάποια πιθανότητα, η αληθής τιμή της οριακής ροπής προς κατανάλωση για ολόκληρο τον πληθυσμό των οικογενειών από τις οποίες προέρχεται το δείγμα. Για να γίνει όμως αυτό απαιτούνται ορισμένες πληροφορίες για τη στατιστική συμπεριφορά των πληθυσμών. Αν οι πληροφορίες αυτές δεν υπάρχουν τότε κάνουμε εμείς ορισμένες υποθέσεις για τη στατιστική συμπεριφορά των πληθυσμών. Οι υποθέσεις αυτές μετατρέπουν το "ακριβές μαθηματικό υπόδειγμα" που συνδέει τις δύο μεταβλητές σε ένα "στατιστικό υπόδειγμα" με τη βοήθεια του οποίου μπορούμε να προχωρήσουμε στον προσδιορισμό διαστημάτων εμπιστοσύνης και στον έλεγχο υποθέσεων για τις αληθείς τιμές των παραμέτρων α και b στους πληθυσμούς. Προχωρούμε λοιπόν στη διατύπωση των ελάχιστων υποθέσεων που απαιτούνται για τη στατιστική επαγωγή σχετικά με τις αληθείς τιμές των παραμέτρων στους πληθυσμούς. Αργότερα θα είμαστε σε θέση να ελέγξουμε αν οι υποθέσεις αυτές ισχύουν πράγματι ή όχι για τους πληθυσμούς από τους οποίους προέρχεται το συγκεκριμένο δείγμα.

2.3. ΤΟ ΣΤΑΤΙΣΤΙΚΟ ΥΠΟΔΕΙΓΜΑ ΤΗΣ ΑΠΛΗΣ ΓΡΑΜΜΙΚΗΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

Ας θεωρήσουμε ξανά το παράδειγμα της κατανάλωσης και του διαθέσιμου εισοδήματος και ας υποθέσουμε ότι διαθέτουμε παρατηρήσεις για την κατανάλωση Y πολλών οικογενειών που έχουν το ίδιο διαθέσιμο εισόδημα x (εκφρασμένο σε απόκλιση από το μέσο για την απλούστευση των υπολογισμών). Είναι φανερό ότι όλες αυτές οι οικογένειες δε θα έχουν την ίδια κατανάλωση παρά το ότι έχουν το ίδιο διαθέσιμο εισόδημα. Αντίθετα η κατανάλωση Y των οικογενειών αυτών θα ακολουθεί κάποια κατανομή με κάποιο μέσο και κάποια διακύμανση. Η κατανομή πιθανότητας της κατανάλωσης Y για το δεδομένο ύψος x του διαθέσιμου εισοδήματος θα είναι η "υπό συνθήκη κατανομή πιθανότητας" p(Y/x) και θα έχουμε μία διαφορετική τέτοια κατανομή για κάθε ύψος του διαθέσιμου εισοδήματος x. Ένα πιθανό σύνολο από τέτοιες κατανομές δίνεται στο σχήμα 2.6. Ε-



Σχήμα 2.6: Πιθανές υπό συνθήκη κατανομές P(Y/x).

να συγκεκριμένο δείγμα παρατηρήσεων για την κατανάλωση και το διαθέσιμο εισόδημα n οικογενειών θα μας εφοδιάσει με μία παρατήρηση από καθένα από τους πληθυσμούς p(Y/x) και πρέπει από τη μία αυτή παρατήρηση να κάνουμε στατιστική επαγωγή για τις παραμέτρους ολόκληρης της κατανομής. Είναι φανερό πως αυτό είναι αδύνατο. Αν επιθυμούμε συνεπώς να προχωρήσουμε σε στατιστική επαγωγή για τους πληθυσμούς αυτούς τότε θα χρειαστούμε ορισμένες γενικές υποθέσεις για τη στατιστική συμπεριφορά όλων των κατανομών p(Y/x). Οι υποθέσεις αυτές είναι οι εξής:

$\boxed{(\upsilon.1)}$: Οι μέσοι $E(Y_i)$ των κατανομών $p(Y_i/x_i)$ βρίσκονται πάνω στην ίδια ευθεία η οποία είναι η "αληθής γραμμική σχέση στους πληθυσμούς" (σχήμα 2.7): Υποθέσεις

$$\boxed{E(Y_i)=\alpha+bx_i,\ i=1,2,\ldots,n}$$

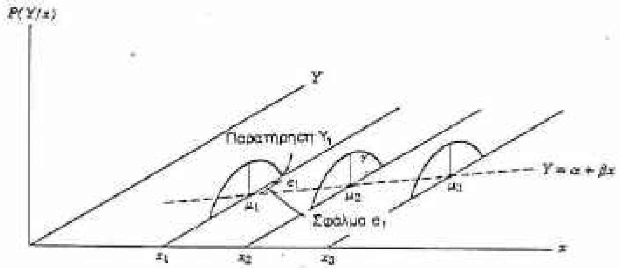
όπου οι παράμετροι α και b θα εκτιμηθούν από το συγκεκριμένο δείγμα. Η υπόθεση αυτή ερμηνεύει τα πιστεύω μας ότι η πραγματική σχέση που συνδέει τις μεταβλητές Y και X είναι γραμμική.

(υ.2): Οι κατανομές $p(Y_i/x_i)$ έχουν την ίδια διακύμανση σ^2 για κάθε x_i, $i=1,2,\ldots,n$.

(υ.3): Οι κατανομές $p(Y_i/x_i)$ είναι στατιστικά ανεξάρτητες, δηλαδή η τιμή Y_i δεν επηρεάζει την τιμή Y_j, $i,j=1,2,\ldots,n$, $(i\neq j)$ και αντίστροφα. Πιο πάνω υποθέσαμε ακόμη ότι:

(υ.4): Οι τιμές $x_1,x_2,\ldots,x_n$ της ανεξάρτητης μεταβλητής παραμένουν σταθερές σε επανειλημμένα δείγματα και αυτό σημαίνει ότι το πείραμα θεωρείται κατευθυνόμενο.

(υ.5): Για λόγους που θα εξηγήσουμε στη γενική περίπτωση της πολλαπλής παλινδρόμησης υποθέτουμε, ότι,οι τιμές των μεταβλητών X και Y έχουν μετρηθεί χωρίς σφάλματα.



Σχήμα 2.7: Οι υπό συνθήκη κατανομές P(Y/x) κάτω από την υπόθεση (υ.1)

Αποτελεί συνήθη πρακτική να γράφουμε το υπόδειγμα της απλής γραμμικής παλινδρόμησης με τη μορφή:

$$\boxed{Y_i=\alpha+bx_i+e_i,\ i=1,2,\ldots,n} \qquad (2.3.1)$$

όπου η τυχαία μεταβλητή e_i εκφράζει την απόκλιση της Y_i από το μέσο της

$$\boxed{E(Y_i)=\alpha+bx_i} \qquad (2.3.2)$$

όπου ξέρουμε το x_i μπορεί να μετρηθεί το X_i

και ονομάζεται "σφάλμα". Οι τυχαίες μεταβλητές Y_i και e_i έχουν την ίδια κατανομή αλλά με διαφορετικό μέσο και οι υποθέσεις (υ.1)-(υ.5) παίρνουν, για το υπόδειγμα (2.3.1), τη μορφή:

(υ.1)	$E(e_i)=0$, $i=1,2,\ldots,n$.	ΣΥΝΟΛΟ ΑΣΘΕΝΩΝ ΥΠΟΘΕΣΕΩΝ
(υ.2)'	$E(e_i^2)=\sigma^2$, $i=1,2,\ldots,n$.	
(υ.3)	Οι κατανομές των e_i και e_j είναι στατιστικά ανεξάρτητες για κάθε $i,j=1,2,\ldots,n$ $(i\neq j)$.	
(υ.4)	Όπως πριν.	
(υ.5)	Όπως πριν.	

Σημειώνουμε ότι μέχρι τώρα δεν έχουμε κάνει καμιά υπόθεση για το είδος των κατανομών $p(Y_i/x_i)$. Υποθέσαμε μόνο ότι οι κατανομές αυτές είναι ανεξάρτητες, ότι έχουν τους μέσους τους πάνω στην ίδια ευθεία, που είναι η αληθής σχέση που συνδέει την Y με τη X, και ότι, έχουν την ίδια διακύμανση. Το σύνολο των υποθέσεων (υ.1)-(υ.5), που το ονομάζουμε "σύνολο των ασθενών υποθέσεων", θα μας επιτρέψει, όπως θα δείξουμε παρακάτω, (1) να εκτιμήσουμε τις παραμέτρους α και b από το συγκεκριμένο δείγμα και (2) να υπολογίσουμε το μέσο και τη διακύμανση των εκτιμητριών αυτών για επανειλημμένες δοκιμές με τις ίδιες τιμές $x_1,x_2,\ldots,x_n$ της X. Αν επιθυμούμε να προχωρήσουμε σε στατιστική επαγωγή για τις αληθείς τιμές των παραμέτρων στους πληθυσμούς θα απαιτηθεί, όπως θα δούμε, και η πρόσθετη υπόθεση της κανονικότητας των κατανομών $p(Y_i/x_i)$:

(υ.6) $\boxed{Y_i \sim N(\alpha+bx_i,\sigma^2)}$ ή $\boxed{e_i \sim N(0,\sigma^2)}$.

Γενικά, η ύπαρξη των σφαλμάτων e_i, δηλαδή των αποκλίσεων των τιμών Y_i από το μέσο τους $E(Y_i)$, μπορεί να οφείλεται σε δύο βασικούς λόγους: ΣΦΑΛΜΑ

(1). "Σε σφάλματα μέτρησης". Στο παράδειγμα της κατανάλωσης και του διαθέσιμου εισοδήματος υπάρχουν πολλοί λόγοι για τη μη ακριβή μέτρηση της κατανάλωσης όπως η απόκρυψη του πραγματικού ύψους της κατανάλωσης, ο ανακριβής υπολογισμός της από τις επιμέρους οικογένειες, κλπ. Αν τα σφάλματα μέτρησης της Y είναι συστηματικά τότε υπάρχουν προβλήματα στη διεξαγωγή της στατιστικής επαγωγής διότι τότε δε θα ισχύει η βασική υπόθεση (υ.1) αφού $E(e_i)\neq 0$. Τα προβλήματα για τη στατιστική επαγωγή είναι πολύ σοβαρότερα όταν υπάρχουν σφάλματα μέτρησης και στην ανεξάρτητη μεταβλητή X. Το πρόβλημα αυτό θα μας απασχολήσει στο γενικότερο πλαίσιο της πολλαπλής παλινδρόμησης. Προς το παρόν υποθέτουμε ότι η μέτρηση των τιμών Y_i και X_i των μεταβλητών Y και X έχει γίνει χωρίς σφάλματα.

2. "Σε παράλειψη ουσιωδών ερμηνευτικών μεταβλητών". Είναι γνωστό ότι στη διαμόρφωση του ύψους της καταναλωτικής δαπάνης μιας οικογένειας συντελούν και άλλοι παράγοντες εκτός από το διαθέσιμο εισόδημα. Οι λόγοι που οι μεταβλητές αυτές δεν εισάγονται αναλυτικά στο υπόδειγμα είναι πολλοί:

i) Η θεωρία που προσδιορίζει τη συμπεριφορά του καταναλωτή, όπως και κάθε θεωρία, δεν είναι πλήρης. Είμαστε βέβαιοι ότι το διαθέσιμο εισόδημα προσδιορίζει σε μεγάλο βαθμό την κατανάλωση μιας οικογένειας αλλά ίσως να αγνοούμε ή να είμαστε αβέβαιοι για την επίδραση άλλων μεταβλητών λιγότερο σημαντικών. Το σφάλμα e_i μπορεί να χρησιμοποιηθεί για να εκφράσει την επίδραση όλων αυτών των μεταβλητών που έχουν παραληφθεί από άγνοια ή αβεβαιότητα.

ii) Έστω και αν γνωρίζουμε μερικές από τις ερμηνευτικές μεταβλητές που έχουμε παραλείψει, είναι δυνατόν να μη διαθέτουμε επαρκή στατιστικά στοιχεία γι' αυτές. Η περιουσία π.χ. επηρεάζει, κατά κανόνα, το ύψος της κατανάλωσης μιας οικογένειας, αλλά, επίσης κατά κανόνα, δεν υπάρχουν στατιστικά στοιχεία για τα περιουσιακά στοιχεία των οικογενειών. Έτσι είμαστε υποχρεωμένοι να παραλείψουμε τη μεταβλητή αυτή από την παλινδρόμηση παρά το ότι πιστεύουμε ότι ασκεί σημαντική επίδραση στη διαμόρφωση της Y.

iii) Στη διαμόρφωση του ύψους της κατανάλωσης επιδρά, εκτός από το διαθέσιμο εισόδημα, ο αριθμός των παιδιών, το φύλο τους, η ηλικία των μελών της οικογένειας, η μόρφωσή τους, η θρησκεία τους, η γεωγραφική περιοχή στην οποία ζει η οικογένεια, κλπ. Είναι όμως δυνατόν, η συνολική επίδραση όλων αυτών των παραγόντων ή μερικών από αυτούς, να είναι πρακτικά μικρή ή μη συστηματική και έτσι, για λόγους κόστους συλλογής των στοιχείων και κόστους διεξαγωγής των σχετικών υπολογισμών, να μην κρίνεται σκόπιμη η αναλυτική εισαγωγή τους στο υπόδειγμα. Άλλωστε, πολλοί από τους παράγοντες που αναφέραμε πιο πάνω είναι ποιοτικοί και είναι δύσκολο και δαπανηρό να μετρηθούν. Η συνολική επίδραση όλων ή μερικών από τους παράγοντες αυτούς μπορεί να εκφραστεί με την παρουσία τους σφάλματος e_i.

iv) Ακόμα και αν εισαχθούν στο υπόδειγμα όλες οι σχετικές ερμηνευτικές μεταβλητές είναι βέβαιο, ότι, θα εξακολουθούν να υπάρχουν μικρές έστω αποκλίσεις μεταξύ των τιμών Y_i που δε θα μπορέσουμε να τις ερμηνεύσουμε όσο και αν προσπαθήσουμε. Αυτό οφείλεται στη φύση των κοινωνικών φαινομένων, στη διαμόρφωση των οποίων υπεισέρχονται διάφοροι ψυχολογικοί παράγοντες που δεν ερμηνεύονται ούτε εκφράζονται ποσοτικά.

v) Τέλος, έστω και αν μπορούμε να ερμηνεύσουμε ικανοποιητικά τη συμπεριφορά της Y με δύο ή τρεις μεταβλητές και η θεωρία δε μας επιβάλλει την εισαγωγή και άλλων μεταβλητών, λόγοι απλότητας του υποδείγματος ή και κόστους των υπολογισμών μας οδηγούν συχνά στην παράλειψή τους.

Για όλους αυτούς τους λόγους που αναφέραμε η φύση των σφαλμάτων e_i παίζει, όπως θα διαπιστώσουμε στο κεφάλαιο 4, πολύ σημαντικό ρόλο στην ανάλυση της παλινδρόμησης και την αξιοπιστία της στατιστικής επαγωγής.

2.4. Η ΜΕΣΗ ΤΙΜΗ ΚΑΙ Η ΔΙΑΚΥΜΑΝΣΗ ΤΩΝ ΕΚΤΙΜΗΤΡΙΩΝ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ. ΤΟ ΘΕΩΡΗΜΑ ΤΩΝ GAUSS-MARKOV

Ας θεωρήσουμε το σχήμα 2.8 και ας υποθέσουμε ότι η αληθής παλινδρόμηση $E(Y)=\alpha+bx$ είναι η διακεκομμένη ευθεία του σχήματος. "Η αληθής παλινδρόμηση είναι άγνωστη και θα παραμείνει άγνωστη" και ο ερευνητής καλείται να χρησιμοποιήσει το συγκεκριμένο δείγμα παρατηρήσεων (X_i, Y_i), $i=1,2,\dots,n$ και την κατάλληλη μέθοδο για να εκτιμήσει όσο το δυνατόν καλλίτερα τις τιμές των παραμέτρων α και b της αληθούς παλινδρόμησης. Αν $\hat{a}$ και $\hat{b}$ είναι οι "εκτιμήσεις" ελαχίστων τετραγώνων των α και b, τότε αυτές δίνονται, όπως είδαμε στην παράγραφο 2.4, από τους τύπους

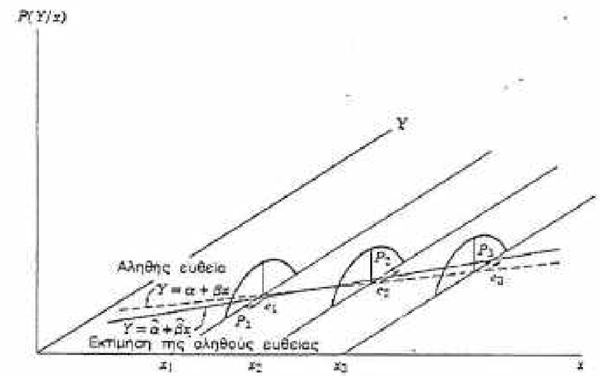
εκτιμήσεις α ή b.

$$\boxed{\hat{a}=\bar{Y}=\frac{\sum_{i=1}^{n} Y_i}{n}, \quad \hat{b}=\frac{\sum_{i=1}^{n} x_i Y_i}{\sum_{i=1}^{n} x_i^2}} \tag{2.4.1}$$

ενώ η εκτίμηση της αληθούς παλινδρόμησης είναι η

$$\boxed{\hat{Y}=\hat{a}+\hat{b}x} \tag{2.4.2}$$

και είναι η συνεχής γραμμή του σχήματος 2.8.



Σχήμα 2.8: Η αληθής παλινδρόμηση και η παλινδρόμηση που εκτιμήθηκε από ένα συγκεκριμένο δείγμα.

Θεωρούμε τώρα ότι επαναλαμβάνουμε το πείραμα άπειρες φορές διατηρώντας τις τιμές $x_1, x_2, \dots, x_n$ της μεταβλητής X σταθερές. Σε κάθε τέτοια δοκιμή θα έχουμε διαφορετικές τιμές $Y_1, Y_2, \dots, Y_n$ της μεταβλητής Y που προέρχονται από τις κατανομές $p(Y_i/x_i)$, $i=1,2,\dots,n$ και συνεπώς διαφορετικές εκτιμήσεις για τις παραμέτρους α και b. Το σύνολο των εκτιμήσεων $\hat{a}$ (ή $\hat{b}$) που θα προκύψει από τις επανειλημμένες αυτές δοκιμές

του πειράματος ονομάζουμε "εκτιμήτρια $\hat{a}$ (ή $\hat{b}$)" και είναι μία τυχαία μεταβλητή (αφού είναι συναρτήσεις των $Y_1, Y_2, \dots, Y_n$ που είναι τυχαίες μεταβλητές) που θα ακολουθεί κάποια κατανομή πιθανότητας με κάποιο μέσο και κάποια διακύμανση. Για να τονίσουμε το γεγονός ότι οι εκτιμήτριες $\hat{a}$ και $\hat{b}$ εξαρτώνται από το μέγεθος n του δείγματος που χρησιμοποιούμε κατά τις επανειλημμένες δοκιμές, θα έπρεπε να χρησιμοποιήσουμε το συμβολισμό $\hat{a}_n$ και $\hat{b}_n$ αλλά δε θα διατηρήσουμε το δείκτη n πιστεύοντας ότι ο όρος "εκτιμήτρια" θα υπενθυμίζει στον αναγνώστη το γεγονός αυτό σε όλη την ανάλυση που θα ακολουθήσει τόσο για την απλή όσο και για την πολλαπλή παλινδρόμηση.

εκτιμήτρια $\hat{a}$ (ή $\hat{b}$).
↓
$\hat{a}_n$ ή $\hat{b}_n$

Τα βασικά ερωτήματα που εύλογα προκύπτουν τώρα είναι:

ΒΑΣΙΚΑ ΕΡΩΤΗΜΑΤΑ

(i) Πως οι κατανομές των εκτιμητριών $\hat{a}$ και $\hat{b}$ συγκεντρώνονται γύρω από τους "στόχους" τους που είναι οι τιμές α και b των παραμέτρων στην αληθή παλινδρόμηση, και

(ii) Πιο είναι το είδος της κατανομής των εκτιμητριών $\hat{a}$ και $\hat{b}$; (γραμμική ή μη γραμμική).

Το πρώτο ερώτημα αναφέρεται στον υπολογισμό του (1) "μέσου" και της (2) "διακύμανσης" των εκτιμητριών $\hat{a}$ και $\hat{b}$ με βάση τις υποθέσεις που έχουμε κάνει για τις κατανομές $p(Y_i/x_i)$ και, όπως θα διαπιστώσουμε στη συνέχεια, για τον υπολογισμό του μέσου και της διακύμανσης των εκτιμητριών ελαχίστων τετραγώνων αρκεί το σύνολο των ασθενών υποθέσεων (υ.1) έως (υ.5). Το σύνολο των ασθενών υποθέσεων αρκεί επίσης για την απόδειξη του θεωρήματος των Gauss-Markov που διατυπώνεται παρακάτω. Η απάντηση στο δεύτερο ερώτημα θα δοθεί στην παράγραφο 2.6.

Αποδεικνύεται ότι για τις κατανομές των εκτιμητριών $\hat{a}$ και $\hat{b}$ ισχύουν τα εξής:

$$\boxed{E(\hat{a})=a \qquad V(\hat{a})=\frac{\sigma^2}{n}} \tag{2.4.3}$$

$$\boxed{E(\hat{b})=b \qquad V(\hat{b})=\frac{\sigma^2}{\Sigma x_i^2}.} \tag{2.4.4}$$

Λόγω της μεγαλύτερης σπουδαιότητας που έχει για την εμπειρική ανάλυση η παράμετρος b (που εκφράζει την οριακή ροπή ή την ελαστικότητα της Y ως προς X) θα αποδείξουμε τα πιο πάνω αποτελέσματα μόνο για την εκτιμήτρια $\hat{b}$. Η απόδειξη των σχετικών αποτελεσμάτων για την εκτιμήτρια $\hat{\alpha}$ είναι ευκολότερη και αφήνεται στον αναγνώστη ως άσκηση. Ο αναγνώστης μπορεί ακόμη να αποδείξει ως άσκηση, ότι, για την εκτιμήτρια $\hat{\alpha}_0=\bar{Y}-\hat{b}\bar{X}$ του σταθερού όρου στην αρχική παλινδρόμηση $Y_i=\alpha_0+bX_i$ ισχύουν τα εξής:

$$E(\hat{\alpha}_0)=\alpha_0 \quad \text{και} \quad V(\hat{\alpha}_0)=\sigma^2\left(\frac{1}{n}+\frac{\bar{X}^2}{\Sigma x_i^2}\right) = \sigma^2\frac{\Sigma X_i^2}{n\Sigma x_i^2}. \qquad (2.4.5)$$

Η εκτιμήτρια $\hat{b}$ γράφεται:

$$\boxed{\hat{b}=\frac{\Sigma x_i Y_i}{\Sigma x_i^2}=\Sigma\left(\frac{x_i}{\Sigma x_i^2}\right)Y_i=\Sigma w_i Y_i} \qquad (2.4.6)$$

όπου

$$w_i=\frac{x_i}{\Sigma x_i^2}. \qquad (2.4.7)$$

Λόγω της υπόθεσης (υ.4) οι ποσότητες w_i είναι σταθερές και ο αναγνώστης εύκολα αποδεικνύει ότι

$$\Sigma w_i=0, \quad \Sigma w_i^2=1/\Sigma x_i^2 \quad \text{και} \quad \Sigma w_i x_i=1. \qquad (2.4.8)$$

Από τη (2.4.6) γίνεται φανερό ότι η εκτιμήτρια $\hat{b}$ είναι "γραμμικός συνδυασμός των ανεξάρτητων τυχαίων μεταβλητών Y_i" με σταθμιστές τις σταθερές ποσότητες w_i και, σύμφωνα με τη στατιστική θεωρία, θα ισχύουν τα εξής:

$$\begin{aligned} E(\hat{b}) &= \Sigma w_i E(Y_i) \\ &= \Sigma w_i(\alpha+bx_i) && [\text{λόγω της (υ.1)}] \\ &= \alpha\Sigma w_i + b\Sigma w_i x_i \\ &= b, && [\text{λόγω της (2.4.8)}] \end{aligned}$$

και

$$\begin{aligned} V(\hat{b}) &= \Sigma w_i^2 V(Y_i) && [\text{λόγω της (υ.3)}] \\ &= \sigma^2\Sigma w_i^2 && [\text{λόγω της (υ.2)}] \\ &= \frac{\sigma^2}{\Sigma x_i^2} && [\text{λόγω της (2.4.8)}]. \end{aligned}$$

Από τα αποτελέσματα αυτά γίνεται φανερό ότι οι εκτιμήτριες ελαχίστων τετραγώνων $\hat{\alpha}$ και $\hat{b}$ είναι "αμερόληπτες" εκτιμήτριες των αληθών τιμών των παραμέτρων α και b στους πληθυσμούς.

Στη συνέχεια θα αποδείξουμε το Θεώρημα των Gauss-Markov το συμπέρασμα του οποίου αποτελεί το σημαντικότερο λόγο για τη χρησιμοποίηση της μεθόδου ελαχίστων τετραγώνων για την εκτίμηση των παραμέτρων της γραμμικής παλινδρόμησης.

Θεώρημα Gauss-Markov: Στο σύνολο των γραμμικών και αμερόληπτων εκτιμητριών των παραμέτρων α και b της γραμμικής παλινδρόμησης, οι εκτιμήτριες ελαχίστων τετραγώνων $\hat{\alpha}$ και $\hat{b}$ έχουν τη μικρότερη διακύμανση δηλαδή είναι οι πιο "αποτελεσματικές". ΘΕΩΡΗΜΑ

Θα αποδείξουμε το θεώρημα για την εκτιμήτρια $\hat{b}$. Ο αναγνώστης ας αποδείξει το θεώρημα για την εκτιμήτρια $\hat{\alpha}$ ως άσκηση.

Ας θεωρήσουμε μια τυχούσα εκτιμήτρια $\bar{b}$ της παραμέτρου b, η οποία είναι γραμμική συνάρτηση των τυχαίων μεταβλητών $Y_1, Y_2, \ldots, Y_n$, δηλαδή

$$\bar{b}=\Sigma c_i Y_i. \qquad (2.4.9)$$

Οι σταθμιστές c_i μπορούν να γραφτούν ως εξής:

$$c_i=w_i+d_i$$

όπου w_i είναι οι σταθμιστές (2.4.7) των Ε.Τ. και d_i είναι οι αποκλίσεις των c_i από τους w_i.

Θα δείξουμε ότι αν η τυχούσα γραμμική εκτιμήτρια $\bar{b}$ είναι και αμερόληπτη εκτιμήτρια της b τότε θα είναι λιγότερο αποτελεσματική από την εκτιμήτρια $\hat{b}$ των ελαχίστων τετραγώνων.

BLUE: best linear unbiased estimators.

Για να είναι η $\bar{b}$ αμερόληπτη εκτιμήτρια της b πρέπει να ισχύει

$$E(\bar{b})=b. \qquad (2.4.10)$$

Αλλά

$$\begin{aligned} E(\bar{b}) &= \Sigma c_i E(Y_i) \\ &= \Sigma c_i(\alpha+bx_i) \\ &= \alpha\Sigma c_i + b\Sigma c_i x_i \end{aligned}$$

και για να ισχύει η (2.4.10) πρέπει

$$\Sigma c_i=0 \quad \text{ή} \quad \Sigma(w_i+d_i)=0 \qquad (2.4.11)$$

και

$$\Sigma c_i x_i=1 \quad \text{ή} \quad \Sigma(w_i+d_i)x_i=1. \qquad (2.4.12)$$

Επειδή $\Sigma w_i=0$, από τη (2.4.11) προκύπτει ότι:

$$\Sigma d_i=0 \qquad (2.4.13)$$

και επειδή $\Sigma w_i x_i=1$, από τη (2.4.12) προκύπτει ότι:

$$\Sigma d_i x_i=0. \qquad (2.4.14)$$

Η διακύμανση της εκτιμήτριας $\bar{b}$ είναι:

$$\begin{aligned} V(\bar{b}) &= \Sigma c_i^2 V(Y_i) \\ &= \sigma^2\Sigma c_i^2 \\ &= \sigma^2\Sigma(w_i+d_i)^2 \\ &= \sigma^2\Sigma(w_i^2+d_i^2+2w_i d_i) \\ &= \sigma^2\Sigma w_i^2+\sigma^2\Sigma d_i^2 \\ &= V(\hat{b})+\sigma^2\Sigma d_i^2, \end{aligned} \qquad (2.4.15)$$

διότι

$$\Sigma w_i d_i=\frac{\Sigma x_i d_i}{\Sigma x_i^2}=0, \quad \text{λόγω της } (2.4.14).$$

Επειδή $\Sigma d_i^2 \geq 0$ (το άθροισμα Σd_i^2 γίνεται μηδέν τότε και μόνο τότε αν $d_i=0$ για κάθε $i=1,2,\ldots,n$, δηλαδή τότε και μόνο τότε αν $\bar{b}=\hat{b}$ αφού θα ισχύει $c_i=w_i$) εύκολα συμπεραίνουμε από τη (2.4.15), ότι:

$$V(\bar{b}) > V(\hat{b})$$

και αυτό αποδεικνύει το θεώρημα.

Πρέπει να επισημάνουμε ότι η εκτιμήτρια $\hat{b}$ των ελαχίστων τετραγώνων είναι η πιο αποτελεσματική (άριστη) μόνο μέσα στο σύνολο των αμερόληπτων και γραμμικών (ως προς Y_i) εκτιμητριών της b και δεν πρέπει να αποκλίσουμε την ύπαρξη άλλων μη αμερόληπτων ή και μη γραμμικών εκτιμητριών που να έχουν μικρότερη διακύμανση από τη $\hat{b}$.

2.5. ΕΚΤΙΜΗΣΗ ΤΗΣ ΔΙΑΚΥΜΑΝΣΗΣ σ^2 ΤΩΝ ΣΦΑΛΜΑΤΩΝ

Στις σχέσεις

$$V(\hat{\alpha})=\frac{\sigma^2}{n} \quad \text{και} \quad V(\hat{b})=\frac{\sigma^2}{\Sigma x_i^2} \qquad (2.5.1)$$

η ποσότητα σ^2, που εκφράζει την κοινή διακύμανση των κατανομών $P(Y_i/x_i)$ ή των $P(e_i/x_i)$ είναι, κατά κανόνα, άγνωστη στον ερευνητή ο οποίος διαθέτει μόνο τα n ζεύγη των παρατηρήσεων (X_i,Y_i). Για να υπολογίσουμε λοιπόν την αριθμητική τιμή των $V(\hat{\alpha})$ και $V(\hat{b})$ απαιτείται μια εκτίμηση της σ^2 και ως τέτοια θα λάβουμε τη διακύμανση των καταλοίπων $\hat{e}_i$ της παλινδρόμησης ελαχίστων τετραγώνων.

Από τη (2.3.1) γνωρίζουμε ότι

$$Y_i=\alpha+bx_i+e_i, \quad i=1,2,\ldots,n \qquad (2.5.2)$$

και αν λάβουμε τους μέσους των δύο μελών της, τότε

$$\bar{Y}_i=\alpha+\bar{e}_i \qquad (2.5.3)$$

διότι $\bar{x}_i=0$. Από τις (2.5.1) και (2.5.3), αν αφαιρέσουμε κατά μέλη, προκύπτει ότι

$$y_i=bx_i+(e_i-\bar{e}_i), \quad i=1,2,\ldots,n. \qquad (2.5.4)$$

Από την παλινδρόμηση ελαχίστων τετραγώνων έχουμε ότι

$$Y_i=\hat{Y}_i+\hat{e}_i, \quad i=1,2,\ldots,n. \qquad (2.5.5)$$

ή

$$(Y_i-\bar{Y}_i)=(\hat{Y}_i-\bar{\hat{Y}}_i)+\hat{e}_i \quad (\text{διότι } \bar{Y}_i=\bar{\hat{Y}}_i)^{1} \qquad (2.5.6)$$

ή

$$y_i=\hat{y}_i+\hat{e}_i \qquad (2.5.7)$$

ή

$$\begin{aligned}\hat{e}_i&=y_i-\hat{y}_i\\&=bx_i+(e_i-\bar{e}_i)-\hat{b}x_i\\&=-(\hat{b}-b)x_i+(e_i-\bar{e}_i), \quad i=1,2,\ldots,n.\end{aligned}$$

Άρα

$$\Sigma\hat{e}_i^2=(\hat{b}-b)^2\Sigma x_i^2+\Sigma(e_i-\bar{e}_i)^2-2(\hat{b}-b)\Sigma x_i(e_i-\bar{e}_i)$$

και

$$E(\Sigma\hat{e}_i)=(n-2)\sigma^2 \qquad (2.5.8)$$

διότι

$$\begin{aligned}E[(\hat{b}-b)^2\Sigma x_i^2]&=\Sigma x_i^2E(\hat{b}-b)^2\\&=\Sigma x_i^2V(\hat{b})\\&=\Sigma x_i^2\frac{\sigma^2}{\Sigma x_i^2}\\&=\sigma^2,\end{aligned}$$

$$\begin{aligned}E[\Sigma(e_i-\bar{e}_i)^2]&=E[\Sigma e_i^2-\frac{1}{n}(\Sigma e_i)^2]\\&=(n-1)\sigma^2\end{aligned}$$

1. $\hat{Y}_i=\hat{\alpha}+\hat{b}x_i$, άρα $\bar{\hat{Y}}_i=\hat{\alpha}+\hat{b}\bar{x}_i=\hat{\alpha}=\bar{Y}$.

και[1]

$$\begin{aligned}E[(\hat{b}-b)\Sigma x_i(e_i-\bar{e}_i)]&=E\left[\frac{\Sigma x_ie_i}{\Sigma x_i^2}(\Sigma x_ie_i-\bar{e}_i\Sigma x_i)\right]\\&=E\left[\frac{(\Sigma x_ie_i)^2}{\Sigma x_i^2}\right]\\&=\sigma^2.\end{aligned}$$

Συνεπώς η

$$\hat{\sigma}^2=\frac{\Sigma\hat{e}_i^2}{n-2} \qquad (2.5.9)$$

είναι "αμερόληπτη εκτιμήτρια" της αληθούς διακύμανσης σ^2 διότι,

$$\begin{aligned}E(\hat{\sigma}^2)&=\frac{E(\Sigma\hat{e}_i^2)}{(n-2)}\\&=\frac{(n-2)\sigma^2}{(n-2)}\\&=\sigma^2. \qquad (2.5.10)\end{aligned}$$

Ο υπολογισμός της $\hat{\sigma}^2$ μπορεί να γίνει εύκολα ως εξής: από την

$$y_i=\hat{y}_i+\hat{e}_i$$

προκύπτει, ότι:

$$\begin{aligned}\Sigma y_i^2&=\Sigma\hat{y}_i^2+\Sigma\hat{e}_i^2+2\Sigma\hat{y}_i\hat{e}_i\\&=\Sigma\hat{y}_i^2+\Sigma\hat{e}_i^2 \qquad (2.5.11)\end{aligned}$$

διότι

$$\Sigma\hat{y}_i\hat{e}_i=\hat{b}\Sigma x_i\hat{e}_i$$

1. Να ληφθεί υπόψη ότι $\hat{b}=\frac{\Sigma x_iY_i}{\Sigma x_i^2}=\frac{\Sigma x_i(\alpha+bx_i+e_i)}{\Sigma x_i^2}=b+\frac{\Sigma x_ie_i}{\Sigma x_i^2}$, άρα $\hat{b}-b=\frac{\Sigma x_ie_i}{\Sigma x_i^2}$.

$$=\hat{b}\Sigma x_i(y_i-\hat{b}x_i) =$$
$$=\hat{b}\Sigma x_i y_i-\hat{b}^2\Sigma x_i^2 =$$
$$=\hat{b}\Sigma x_i y_i-\hat{b}\Sigma x_i y_i = \quad [\text{λόγω της } (2.2.21)]$$
$$=0 \tag{2.5.12}$$

Συνεπώς

$$\Sigma\hat{e}_i^2=\Sigma y_i^2-\Sigma\hat{y}_i^2$$
$$=\Sigma y_i^2-\Sigma(\hat{b}x_i)^2$$
$$=\Sigma y_i^2-\hat{b}^2\Sigma x_i^2$$
$$=\Sigma y_i^2-\hat{b}\Sigma x_i y_i \tag{2.5.13}$$

και

$$\hat{\sigma}^2=\frac{\Sigma y_i^2-\hat{b}\Sigma x_i y_i}{n-2}. \tag{2.5.14}$$

Ο αναγνώστης μπορεί ακόμη να δείξει ότι

$$\hat{\sigma}^2=\frac{\Sigma Y_i^2-n\hat{\alpha}^2-\hat{b}\Sigma x_i Y_i}{n-2}. \tag{2.5.15}$$

Άρα η εκτίμηση των διακυμάνσεων $V(\hat{\alpha})$ και $V(\hat{b})$ είναι:

$$\widehat{V(\hat{\alpha})}=\frac{\hat{\sigma}^2}{n} \quad \text{και} \quad \widehat{V(\hat{b})}=\frac{\hat{\sigma}^2}{\Sigma x_i^2}. \tag{2.5.16}$$

Επιστρέφοντας στο συγκεκριμένο παράδειγμα της κατανάλωσης και του διαθέσιμου εισοδήματος, εύκολα υπολογίζουμε από τα στοιχεία του πίνακα 2.1 ότι

$$\hat{\sigma}^2=1.7$$
$$\widehat{V(\hat{\alpha})}=0.17$$
$$\widehat{V(\hat{b})}=0.0013.$$

2.6. Η ΚΑΤΑΝΟΜΗ ΠΙΘΑΝΟΤΗΤΑΣ ΤΩΝ ΕΚΤΙΜΗΤΡΙΩΝ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

Μέχρι τώρα είδαμε πως υπολογίζονται οι εκτιμήτριες ελαχίστων τετραγώνων $\hat{\alpha}$ και $\hat{b}$ των παραμέτρων α και b της αληθούς παλινδρόμησης, δείξαμε ότι είναι αμερόληπτες εκτιμήτριες, ότι είναι οι πιο αποτελεσματικές από όλες τις γραμμικές και αμερόληπτες εκτιμήτριες των α και b και εκτιμήσαμε τη διακύμανσή τους.

Στόχος μας όμως, όπως έχουμε ήδη τονίσει, δεν είναι μόνο ο υπολογισμός των $\hat{\alpha}$ και $\hat{b}$ από το συγκεκριμένο δείγμα αλλά, κυρίως, ο προσδιορισμός διαστημάτων εμπιστοσύνης μέσα στα οποία θα βρίσκονται, με κάποια πιθανότητα, οι αληθείς τιμές α και b των παραμέτρων, για να μπορέσουμε να προβούμε σε ελέγχους υποθέσεων σχετικά με τις τιμές των α και b στους πληθυσμούς και να διατυπώσουμε προβλέψεις.

Για να προχωρήσουμε στην εργασία αυτή είναι απαραίτητο να γνωρίζουμε, εκτός από το μέσο και τη διακύμανση, το είδος της κατανομής των εκτιμητριών $\hat{\alpha}$ και $\hat{b}$ και, επειδή αυτές είναι γραμμικές συναρτήσεις των Y_i, αρκεί να γνωρίζουμε τις κατανομές των τυχαίων μεταβλητών Y_i.

Στο σημείο αυτό εισάγεται η "ισχυρή υπόθεση" (υ.6) της κανονικότητας των κατανομών $P(Y_i/x_i)$ ή $P(e_i/x_i)$:

$$Y_i \sim N(\alpha+bx,\sigma^2)$$

ή

$$e_i \sim N(0,\sigma^2)$$

και επειδή οι εκτιμήτριες $\hat{\alpha}$ και $\hat{b}$ είναι γραμμικές συναρτήσεις των Y_i θα ακολουθούν και αυτές την κανονική κατανομή και συγκεκριμένα

$$\hat{\alpha} \sim N\left(\alpha, \frac{\sigma^2}{n}\right) \tag{2.6.1}$$

και

$$\hat{b} \sim N\left(b, \frac{\sigma^2}{\Sigma x_i^2}\right). \tag{2.6.2}$$

Η υπόθεση (υ.6) της κανονικότητας των κατανομών $P(Y_i/x_i)$

ή $P(e_i/x_i)$ δεν είναι απαραίτητη όταν το μέγεθος του δείγματος είναι μεγάλο διότι στην περίπτωση αυτή η κατανομή των $\hat{a}$ και $\hat{b}$ θα τείνει προς την κανονική, σύμφωνα με το "Κεντρικό Οριακό θεώρημα". Βέβαια, το Κεντρικό Οριακό θεώρημα, στη συνήθη διατύπωσή του, αποδεικνύει την κανονικότητα του δειγματικού μέσου των Y_i αλλά στη γενικευμένη διατύπωσή[1] του εξασφαλίζει επίσης τη σύγκλιση προς την κανονική κατανομή του αθροίσματος καθώς και των γραμμικών συνδυασμών των Y_i όπως είναι η εκτιμήτρια $\hat{b}$.

Ακόμη, πρέπει να παρατηρήσουμε ότι, κάτω από την υπόθεση (υ.6) αποδεικνύεται ότι οι εκτιμήτριες ελαχίστων τετραγώνων έχουν τη μικρότερη διακύμανση από όλες τις αμερόληπτες εκτιμήτριες των a και b γραμμικές ή όχι. Το θεώρημα αυτό, γνωστό ως θεώρημα του Rao[2], είναι βέβαια ισχυρότερο από το θεώρημα των Gauss-Markov που περιορίζεται μόνο στην κλάση των γραμμικών εκτιμητριών.

2.7. ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ ΚΑΙ ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ ΓΙΑ ΤΙΣ ΤΙΜΕΣ ΤΩΝ ΠΑΡΑΜΕΤΡΩΝ ΣΤΟΥΣ ΠΛΗΘΥΣΜΟΥΣ

Αφού προσδιορίσαμε την κατανομή των εκτιμητριών $\hat{a}$ και $\hat{b}$ μπορούμε τώρα να προχωρήσουμε στον προσδιορισμό διαστημάτων εμπιστοσύνης για τις αληθείς τιμές a και b των παραμέτρων στον πληθυσμό.

Από την τυποποίηση της κατανομής (2.6.2) προκύπτει, ότι, η τυχαία μεταβλητή

$$z = \frac{\hat{b}-b}{\sqrt{V(\hat{b})}} = \frac{\hat{b}-b}{\sigma/\sqrt{\Sigma x_i^2}} = \frac{(\hat{b}-b)\sqrt{\Sigma x_i^2}}{\sigma} \tag{2.7.1}$$

ακολουθεί την τυποποιημένη κανονική κατανομή N(0,1). Αν ήταν γνωστή η ποσότητα σ θα μπορούσαμε να προχωρήσουμε στη στατιστική επαγωγή με τη βοήθεια των πινάκων της τυποποιημένης κανονικής κατανομής. Όπως έχουμε τονίσει όμως, η σ είναι άγνωστη και εκτιμάται από τη $\hat{\sigma}$. Πρέπει λοιπόν να αναζητήσουμε το είδος της κατανομής της τυχαίας μεταβλητής

$$\frac{\hat{b}-b}{\sqrt{\widehat{V(\hat{b})}}} = \frac{(\hat{b}-b)\sqrt{\Sigma x_i^2}}{\hat{\sigma}}.$$

1. Για τις διάφορες γενικεύσεις του Κεντρικού Οριακού Θεωρήματος βλέπε H. Cramer, *Mathematical Methods of Statistics*, Princeton University Press, Princeton, N.J., 1946, chap. 17.

2. Για την απόδειξη του θεωρήματος του Rao βλέπε C.R. Rao, *Linear Statistical Inference and its Applications*, Wiley, New York, 1965, σελ. 258.

Από τη στατιστική θεωρία γνωρίζουμε[1] ότι η τυχαία μεταβλητή

$$\upsilon^2 = \frac{(n-2)\hat{\sigma}^2}{\sigma^2} \tag{2.7.2}$$

ακολουθεί την κατανομή χ^2 με (n-2) βαθμούς ελευθερίας και ότι η κατανομή της υ είναι ανεξάρτητη[1] από την κατανομή της z. Άρα η τυχαία μεταβλητή

$$t = \frac{z}{\sqrt{\frac{\upsilon^2}{n-2}}} = \frac{\frac{(\hat{b}-b)\sqrt{\Sigma x_i^2}}{\sigma}}{\frac{\hat{\sigma}}{\sigma}} = \frac{(\hat{b}-b)\sqrt{\Sigma x_i^2}}{\hat{\sigma}} \tag{2.7.3}$$

ακολουθεί την κατανομή t του Student με n-2 βαθμούς ελευθερίας.

Αν $t_{0,025}$ είναι η τιμή της t που αφήνει δεξιά της το 2.5% της κατανομής, δηλαδή

$$P(-t_{0,025} < t < t_{0,025}) = 0,95 \tag{2.7.4}$$

τότε, λόγω της (2.7.3), θα έχουμε:

$$P\left(-t_{0,025} < \frac{(\hat{b}-b)\sqrt{\Sigma x_i^2}}{\hat{\sigma}} < t_{0,025}\right) = 0,95 \tag{2.7.5}$$

ή

1. Οι σχετικές αποδείξεις θα δοθούν στη γενική περίπτωση της πολλαπλής παλινδρόμησης.

$$P\left(\hat{b}-t_{0,025}\frac{\hat{\sigma}}{\sqrt{\Sigma x_i^2}}<b<\hat{b}+t_{0,025}\frac{\hat{\sigma}}{\sqrt{\Sigma x_i^2}}\right)=0,95. \qquad (2.7.6)$$

Άρα, το διάστημα εμπιστοσύνης μέσα στο οποίο θα βρίσκεται, με πιθανότητα 95%, η αληθής τιμή της παραμέτρου b στον πληθυσμό είναι

$$\hat{b}\pm t_{0,025}\frac{\hat{\sigma}}{\sqrt{\Sigma x_i^2}}=\hat{b}\pm t_{0,025}\sqrt{V(\hat{b})} \qquad (2.7.7)$$

Ακολουθώντας την ίδια διαδικασία μπορούμε εύκολα να δείξουμε ότι το 95% διάστημα εμπιστοσύνης για την α είναι

$$\hat{\alpha}\pm t_{0,025}\frac{\hat{\sigma}}{\sqrt{n}}=\hat{\alpha}\pm t_{0,025}\sqrt{V(\hat{\alpha})}. \qquad (2.7.8)$$

Για το συγκεκριμένο παράδειγμα της κατανάλωσης και του διαθέσιμου εισοδήματος, το διάστημα μέσα στο οποίο βρίσκεται, με πιθανότητα 95%, η αληθής τιμή της οριακής ροπής προς κατανάλωση b στον πληθυσμό, είναι

$$0,485\pm 2.306\cdot 0.036$$

ή

$$0,402<b<0,568.$$

Σχετικά με τον έλεγχο υποθέσεων, η συνήθης υπόθεση που ελέγχεται είναι η

$$H_0:\ b=0. \qquad (2.7.9)$$

Το ερώτημα που τίθεται είναι αν πρέπει να προβούμε σε μονόπλευρο ή αμφίπλευρο έλεγχο της υπόθεσης H_0. Σ'αυτό θα μας κατευθύνει η οικονομική θεωρία. Αν η οικονομική θεωρία μας διαβεβαιώνει ότι η οριακή ροπή b είναι θετική, τότε, η H_0 ελέγχεται κατά της

$$H_1:\ b>0$$

και ο έλεγχος είναι μονόπλευρος. Στην περίπτωση αυτή, κάτω από την υπόθεση H_0, ο λόγος t στη (2.7.3) παίρνει την τιμή

$$t=\frac{\hat{b}\sqrt{\Sigma x_i^2}}{\hat{\sigma}} \qquad (2.7.10)$$

και το διάστημα απόρριψης για την H_0 ορίζεται από την τιμή $t_{0,05}$ των πινάκων, με $(n-2)$ βαθμούς ελευθερίας, που αφήνει στα δεξιά της το 5% της κατανομής. Αν η υπολογιζόμενη τιμή t από τη (2.7.10) είναι μεγαλύτερη από την τιμή $t_{0,05}$ των πινάκων, τότε, απορρίπτουμε την H_0 και δεχόμαστε την H_1, δηλαδή δεχόμαστε ότι η τιμή b της παραμέτρου στον πληθυσμό είναι στατιστικά μεγαλύτερη από το μηδέν.

Αν η οικονομική θεωρία δε μπορεί να μας διαβεβαιώσει για το πρόσημο της b τότε η H_0 ελέγχεται κατά της

$$H_1:\ b\neq 0$$

και ο έλεγχος είναι αμφίπλευρος. Στην περίπτωση του αμφίπλευρου ελέγχου η H_0 απορρίπτεται αν η τιμή $b=0$ δεν περιέχεται στο διάστημα εμπιστοσύνης (2.7.7) ή, ισοδύναμα, αν η απόλυτη τιμή του λόγου t από τη (2.7.10) είναι μεγαλύτερη από τη θεωρητική τιμή $t_{0,025}$ των πινάκων.

Για το παράδειγμα της κατανάλωσης και του διαθέσιμου εισοδήματος είναι φανερό ότι πρέπει να ελέγξουμε την υπόθεση

$$H_0:\ b=0$$

κατά της

$$H_1:\ b>0.$$

Η τιμή του λόγου t από τη (2.7.10) είναι

$$t=\frac{0.485\sqrt{1320}}{\sqrt{1.7}}$$

$$=13.51$$

και αν τη συγκρίνουμε με τη θεωρητική τιμή $t_{0,05}=1,86$ των πινάκων, για 8 βαθμούς ελευθερίας, διαπιστώνουμε ότι $t>t_{0,05}$, συνεπώς απορρίπτουμε την H_0 και δεχόμαστε ότι η οριακή ροπή προς κατανάλωση στον πληθυσμό είναι θετική και στατιστικά σημαντική σε επίπεδο σημαντικότητας 5%.

Μπορούμε, τέλος, να ελέγξουμε υποθέσεις της μορφής

$$H_0: b=\gamma$$

κατά της

$$H_1: b\neq\gamma \quad \text{(αμφίπλευρος έλεγχος)}$$

ή της

$$H_1: b\gtrless\gamma \quad \text{(μονόπλευρος έλεγχος)}$$

αρκεί να εξετάσουμε αν η τιμή γ περιέχεται στο διάστημα εμπιστοσύνης (2.7.7) για αμφίπλευρο έλεγχο ή αν $t>t_{0,05}$ για μονόπλευρο έλεγχο, όπου η στατιστική t υπολογίζεται από τη (2.7.3) θέτοντας b=γ, δηλαδή από την

$$t=\frac{(\hat{b}-\gamma)\sqrt{\Sigma x_i^2}}{\hat{\sigma}}. \qquad (2.7.11)$$

2.8. Ο ΣΥΝΤΕΛΕΣΤΗΣ ΑΠΛΟΥ ΠΡΟΣΔΙΟΡΙΣΜΟΥ ΚΑΙ Η ΑΝΑΛΥΣΗ ΤΗΣ ΔΙΑΚΥΜΑΝΣΗΣ ΣΤΗΝ ΑΠΛΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Στο σημείο αυτό, ας εξετάσουμε, τη σχέση που συνδέει την εκτίμηση του γωνιακού συντελεστή $\hat{b}$ της απλής παλινδρόμησης με το συντελεστή συσχέτισης r_{XY} των μεταβλητών X και Y που υπολογίσαμε στο πρώτο κεφάλαιο.

Από τη σχέση (2.2.21) προκύπτει ότι:

$$\hat{b}=\frac{\Sigma x_i y_i}{\Sigma x_i^2}$$

$$=\frac{\Sigma x_i y_i}{\sqrt{\Sigma x_i^2}\sqrt{\Sigma y_i^2}}\frac{\sqrt{\Sigma y_i^2}}{\sqrt{\Sigma x_i^2}}$$

ή

$$\hat{b}=r_{XY}\frac{S_Y}{S_X} \qquad (2.8.1)$$

όπου r_{XY} είναι ο συντελεστής συσχέτισης των X και Y και S_Y, S_X η μέση απόκλιση τετραγώνων των Y και X αντίστοιχα. Άρα, η εκτίμηση $\hat{b}$ ελαχίστων τετραγώνου από το συγκεκριμένο δείγμα είναι ανάλογη προς το συντελεστή συσχέτισης r_{XY} των παρατηρήσεων του δείγματος και ο συντελεστής αναλογίας είναι ο λόγος S_Y/S_X.

Στη συνέχεια μπορούμε να φτάσουμε σε ορισμένα ενδιαφέροντα συμπεράσματα.

Από τη (2.5.11) έχουμε ότι

$$\Sigma y_i^2=\Sigma\hat{y}_i^2+\Sigma e_i^2. \qquad (2.8.2)$$

Η σχέση αυτή ερμηνεύεται ως εξής: Η συνολική διακύμανση των "εκ παρατηρήσεως" τιμών Y_i, i=1,2,..., της Y, γύρω από το μέσο τους $\bar{Y}_i$, αναλύεται σε δύο συνιστώσες, οι οποίες, λόγω της (2.5.12), είναι ορθογώνιες. Η πρώτη συνιστώσα είναι η διακύμανση των "εκ προσαρμογής" τιμών $\hat{Y}_i$ γύρω από το μέσο τους $\bar{\hat{Y}}_i$ και είναι το μέρος της διακύμανσης των Y_i που ερμηνεύεται από τη γραμμική παλινδρόμηση της Y πάνω στη X. Η δεύτερη συνιστώσα είναι η διακύμανση των "καταλοίπων" $\hat{e}_i$ της παλινδρόμησης και εκφράζει το μέρος της διακύμανσης των Y_i που παραμένει ανερμήνευτο από την παλινδρόμηση.

Αν πάρουμε το λόγο της διακύμανσης που ερμηνεύεται από την παλινδρόμηση προς τη συνολική διακύμανση έχουμε:

$$\frac{\Sigma\hat{y}_i^2}{\Sigma y_i^2}=\frac{\Sigma(\hat{b}x_i)^2}{\Sigma y_i^2}$$

$$=\hat{b}^2\frac{\Sigma x_i^2}{\Sigma y_i^2}$$

$$=r_{XY}^2, \qquad \text{[λόγω της (2.8.1)]}. \qquad (2.8.3)$$

Ερμηνεύοντας τη (2.8.3) σε σχέση με το παράδειγμα της κατανάλωσης Y και του διαθέσιμου εισοδήματος X (για το οποίο έχουμε υπολογίσει ότι r_{XY}=0,979) διαπιστώνουμε ότι, το 95,8% ($=r^2\times100$) της συνολικής διακύμανσης της Y στο δείγμα ερμηνεύεται από τη γραμμική επίδραση της X πάνω στην Y ενώ το υπόλοιπο 4,2% παραμένει ανερμήνευτο και οφείλεται είτε στην παράλειψη άλλων ερμηνευτικών μεταβλητών είτε σε τυχαία σφάλματα.

Το τετράγωνο του συντελεστή συσχέτισης r^2_{XY} συμβολίζεται, συνήθως, με R^2 και ονομάζεται "συντελεστής απλού προσδιορισμού" για την παλινδρόμηση.

Από τις σχέσεις (2.8.2) και (2.8.3) εύκολα προκύπτει, ότι,

$$R^2=\frac{\Sigma\hat{y}_i^2}{\Sigma y_i^2}$$

$$=1-\frac{\Sigma\hat{e}_i^2}{\Sigma y_i^2}, \qquad (2.8.4)$$

από την οποία γίνεται φανερό, ότι, η μεγαλύτερη τιμή του R^2 είναι η μονάδα και αυτό συμβαίνει μόνο αν $\Sigma\hat{e}_i^2=0$, δηλαδή όταν όλα τα σημεία (X_i,Y_i) βρίσκονται πάνω στην ίδια ευθεία με κλίση $\hat{b}$ και σταθερό όρο $\hat{a}_0$ στο σύστημα αξόνων (X,Y) η $\hat{a}$ στο σύστημα (x,Y) ή 0 στο σύστημα (x,y). Η μικρότερη τιμή του R^2 είναι το μηδέν και αυτό συμβαίνει μόνο αν $\Sigma\hat{e}_i^2=\Sigma y_i^2$, δηλαδή όταν η γραμμική επίδραση της X πάνω στην Y είναι μηδενική.

Είναι φανερό ότι ο ρόλος του συντελεστή προσδιορισμού R^2 στην αξιολόγηση της παλινδρόμησης που εκτιμήσαμε είναι σημαντικός αφού, όσο η τιμή του πλησιάζει προς τη μονάδα, τόσο μεγαλύτερο είναι το μέρος της διακύμανσης της Y που ερμηνεύεται από τη γραμμική παλινδρόμηση, δηλαδή τόσο καλλίτερη είναι η προσαρμογή της ευθείας ελαχίστων τετραγώνων στις παρατηρήσεις του δείγματος. Περισσότερα για την αξία του R^2 ως μέτρου προσαρμογής της παλινδρόμησης θα αναφέρουμε στο κεφάλαιο της πολλαπλής παλινδρόμησης.

Στη συνέχεια διατυπώνουμε ένα ακόμη κριτήριο για τον έλεγχο της στατιστικής σημαντικότητας της παλινδρόμησης που εκτιμήσαμε, ο οποίος δεν είναι βέβαια ανεξάρτητος από τους ελέγχους υποθέσεων που διατυπώσαμε στα προηγούμενα.

Έχουμε δείξει ότι

$$z=\frac{(\hat{b}-b)\sqrt{\Sigma x_i^2}}{\sigma}\sim N(0,1)$$

άρα,

$$z^2=\frac{(\hat{b}-b)^2\Sigma x_i^2}{\sigma^2}\sim \chi^2_1. \qquad (2.8.5)$$

Είδαμε επίσης ότι

$$\upsilon^2=\frac{(n-2)\hat{\sigma}^2}{\sigma^2}\sim \chi^2_{n-2}. \qquad (2.8.6)$$

και ότι οι κατανομές (2.8.5) και (2.8.6) είναι ανεξάρτητες.

Σύμφωνα με τη στατιστική θεωρία, ο λόγος

$$F=\frac{z^2}{\upsilon^2/(n-2)}$$

$$=\frac{(\hat{b}-b)^2\Sigma x_i^2}{\hat{\sigma}^2}$$

$$=\frac{(\hat{b}-b)^2\Sigma x_i^2}{\Sigma\hat{e}^2/(n-2)}\sim F^1_{n-2} \qquad (2.8.7)$$

δηλαδή ακολουθεί την κατανομή F με 1 και (n-2) βαθμούς ελευθερίας.

Κάτω από την υπόθεση:

$$H_0:\ b=0$$

η τιμή του λόγου F είναι:

$$F=\frac{Q_1}{Q_2/(n-2)} \qquad (2.8.8)$$

όπου

$$Q_1=\hat{b}^2\Sigma x_i^2$$

$$=\Sigma\hat{y}_i^2 \qquad (2.8.9)$$

είναι το μέρος της διακύμανσης της Y που ερμηνεύεται από την παλινδρόμηση, και

$$Q_2=\Sigma\hat{e}_i^2 \qquad (2.8.10)$$

είναι το μέρος της διακύμανσης της Y που παραμένει ανερμήνευτο.

Άρα η υπόθεση ότι δεν υπάρχει στατιστικά σημαντική γραμμική σχέση ($b=0$) μεταξύ της Y και της X μπορεί να ελεγχθεί και ως εξής: αν η τιμή του λόγου F, που υπολογίζεται από τη (2.8.8) κάτω από την υπόθεση H_0, είναι μεγαλύτερη από την τιμή $F_{0,05}$ των πινάκων για (1, n-2) βαθμούς ελευθερίας, τότε απορρίπτουμε την υπόθεση H_0 και δεχόμαστε ότι υπάρχει στατιστικά σημαντική γραμμική σχέση μεταξύ της Y και της X σε επίπεδο σημαντικότητας 5%.

Στον πίνακα 2.2 παρουσιάζεται η πιο πάνω ανάλυση της διακύμανσης:

ΠΙΝΑΚΑΣ 2.2: Η ανάλυση της διακύμανσης στην απλή γραμμική παλινδρόμηση.

Διακύμανση που ερμηνεύεται	Αντίστοιχο άθροισμα τετραγώνων	Βαθμοί ελευθερίας	Μέσο άθροισμα τετραγώνων
Από τη X	$Q_1=\Sigma\hat{y}_i^2$	1	$\Sigma\hat{y}_i^2/1$
Από τα κατάλοιπα	$Q_2=\Sigma\hat{e}_i^2$	n-2	$\Sigma\hat{e}_i^2/(n-2)$
Συνολική διακύμανση	$\Sigma y_i^2=Q_1+Q_2$	n-1	$F=\dfrac{Q_1}{Q_2/(n-2)}$

Οι ποσότητες Q_1 και Q_2 μπορούν να υπολογιστούν από τις αρχικές παρατηρήσεις ως εξής:

$$Q_1=\hat{b}^2\Sigma x_i^2=\hat{b}\Sigma x_i y_i \tag{2.8.11}$$

και

$$Q_2=\Sigma\hat{y}_i^2-Q_1$$

$$=\Sigma y_i^2-\hat{b}\Sigma x_i y_i . \tag{2.8.12}$$

Η ανάλυση της διακύμανσης για το παράδειγμα της κατανάλωσης και του διαθέσιμου εισοδήματος δίνεται στον πίνακα 2.3:

ΠΙΝΑΚΑΣ 2.3: Ανάλυση της διακύμανσης για την κατανάλωση των 10 οικογενειών

Διακύμανση της κατανάλωσης (Y) που ερμηνεύεται:	Αντίστοιχο άθροισμα τετραγώνων	Βαθμοί ελευθερίας	Μέσο άθροισμα τετραγώνων
Από το διαθέσιμο εισόδημα (X)	310.4	1	310.4
Από τα κατάλοιπα	13.6	8	13.6/8=1.7
Συνολική διακύμανση	324.0	9	$F=\dfrac{310.4}{1.7}=182.6$

Από τους πίνακες της κατανομής F_8^1 (με 1 και 8 βαθμούς ελευθερίας) δίνεται ότι $F_{0,05}=5.32$ και επειδή η τιμή του λόγου F=182.6, που υπολογίζεται από την παλινδρόμηση, υπερβαίνει κατά πολύ την τιμή $F_{0,05}$ των πινάκων, απορρίπτουμε την υπόθεση H_0 και δεχόμαστε ότι υπάρχει στατιστικά σημαντική γραμμική σχέση μεταξύ της κατανάλωσης και του διαθέσιμου εισοδήματος των οικογενειών του δείγματος. Η γραμμική αυτή σχέση είναι βέβαια στατιστικά σημαντική όχι μόνο σε επίπεδο σημαντικότητας 5% αλλά ακόμη και σε επίπεδο σημαντικότητας 1‰ όπως εύκολα μπορούμε να διαπιστώσουμε συγκρίνοντας την τιμή του λόγου F=182.6 με την τιμή $F_{0,001}=29.3$ των πινάκων για 1 και 8 βαθμούς ελευθερίας.

2.9. Η ΠΑΡΟΥΣΙΑΣΗ ΤΩΝ ΑΠΟΤΕΛΕΣΜΑΤΩΝ ΣΤΗΝ ΑΠΛΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Υπάρχουν πολλοί τρόποι για να αναφέρουμε τα αποτελέσματα των υπολογισμών από την εκτίμηση μιας απλής παλινδρόμησης. Εδώ θα υιοθετήσουμε την ακόλουθη μορφή παρουσίασης, χρησιμοποιώντας το παράδειγμα της κατανάλωσης και του διαθέσιμου εισοδήματος που έχουμε επεξεργαστεί στο κεφάλαιο αυτό:

$$\hat{Y}_i=5.51+0.485X , \quad R^2=0.958$$
$$(1.28)\ (0.036) \quad B.E.=8 \tag{2.9.1}$$
$$(t=4.305)\ (t=13.472)$$

Οι αριθμοί στις πρώτες παρενθέσεις δίνουν τις εκτιμήσεις των τυπικών σφαλμάτων (της μέσης απόκλισης τετραγώνου) των αντίστοιχων συντελεστών και οι αριθμοί στις δεύτερες παρενθέσεις τις εκτιμήσεις των λόγων t κάτω από την υπόθεση H_0 ότι η τιμή του αντίστοιχου συντελεστή στον πληθυσμό είναι μηδέν (π.χ. 4,305=5,51:1,28).

Ένα πλεονέκτημα από την παρουσίαση των αποτελεσμάτων στη μορφή αυτή είναι ότι μπορούμε να διαπιστώσουμε αμέσως αν κάθε ένας από τους συντελεστές που εκτιμήσαμε είναι στατιστικά σημαντικός. Στην περίπτωση του παραδείγματός μας, επειδή η θεωρητική τιμή $t_{0,025}$ από τους πίνακες (για 8 βαθμούς ελευθερίας) είναι ίση με 2,306, είναι φανερό ότι και ο σταθερός όρος και η κλίση της ευθείας που εκτιμήσαμε είναι στατιστικά σημαντικοί. Φυσικά, μπορούμε να ελέγξουμε και κάθε άλλη υπόθεση σχετικά με τις τιμές των συντελεστών στους πληθυσμούς χρησιμοποιώντας την τιμή του λόγου t από τη σχέση (2.7.11).

Η τιμή του συντελεστή απλού προσδιορισμού R^2 στη (2.9.1) δίνει, αν πολλαπλασιαστεί επί 100, το ποσοστό της διακύμανσης της Y που ερμηνεύεται από την παλινδρόμηση. Έτσι, στο παράδειγμά μας, το 95,6% της διακύμανσης της κατανάλωσης ερμηνεύεται από την παλινδρόμηση ενώ το 4,4% παραμένει ανερμήνευτο. Τέλος, οι βαθμοί ελευθερίας πρέπει να αναφέρονται πάντοτε για να είναι δυνατή η εύρεση της θεωρητικής τιμής της κατανομής t από το σχετικό πίνακα.

2.10. ΧΡΗΣΙΜΟΠΟΙΗΣΗ ΤΗΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ ΓΙΑ ΤΗ ΔΙΑΤΥΠΩΣΗ ΠΡΟΒΛΕΨΕΩΝ

Όπως τονίσαμε στην εισαγωγή του εγχειριδίου αυτού, ένας από τους βασικούς σκοπούς της ποσοτικής ανάλυσης είναι και η διατύπωση προβλέψεων. Στην παράγραφο αυτή θα προσπαθήσουμε να απαντήσουμε στα εξής ερωτήματα σχετικά με τη διατύπωση προβλέψεων:

i) Αν έχουμε "πολλές οικογένειες" που διαθέτουν ένα ορισμένο ύψος διαθέσιμου εισοδήματος, έστω X_0, πόση "προβλέπεται" ότι θα είναι η "μέση κατανάλωση" $E(Y_0/X_0)=\hat{\mu}_0$ των οικογενειών αυτών και ποιο θα είναι π.χ. το 95% διάστημα εμπιστοσύνης μέσα στο οποίο θα βρίσκεται η αληθής τιμή της μέσης κατανάλωσης σ' ολόκληρο τον πληθυσμό των οικογενειών που έχουν διαθέσιμο εισόδημα X_0;

ii) Αν έχουμε "μια μόνο οικογένεια" με διαθέσιμο εισόδημα X_0, πόση "προβλέπεται" ότι θα είναι η κατανάλωση $\hat{Y}_0$ της οικογένειας αυτής και ποιο θα είναι το 95% διάστημα εμπιστοσύνης μέσα στο οποίο θα βρίσκεται η αληθής τιμή Y_0 της κατανάλωσης της οικογένειας αυτής.

Πρέπει να τονίσουμε εδώ, ότι, η απάντηση στα δύο αυτά ερωτήματα θα δοθεί με βάση τις πληροφορίες που έχουμε αποκομίσει μέχρι τώρα από την ανάλυση του δείγματος των 10 οικογενειών, δηλαδή με βάση τη διαπίστωση ότι η σχέση που συνδέει την κατανάλωση με το διαθέσιμο εισόδημα είναι η (2.9.1).

i) Πρόβλεψη για το μέσο μ_0

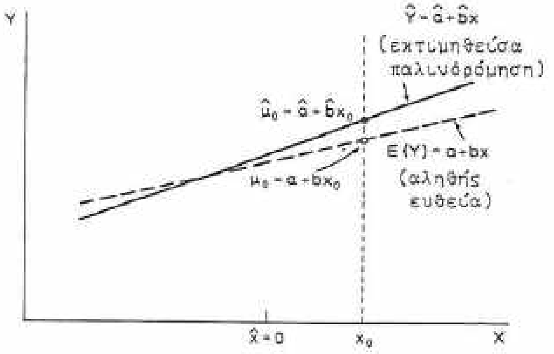
Επειδή η παλινδρόμηση $\hat{Y}_i=\hat{\alpha}+\hat{b}x_i$, που εκτιμήσαμε από το δείγμα των 10 οικογενειών, είναι η ευθεία πάνω στην οποία βρίσκονται ακριβώς οι εκτιμήσεις των $E(Y_i/x_i)$, είναι φανερό ότι η κατάλληλη εκτιμήτρια για το μέσο $E(Y_0/X_0)$ θα είναι η

$$\hat{\mu}_0=\hat{\alpha}+\hat{b}x_0 \qquad (2.10.1)$$

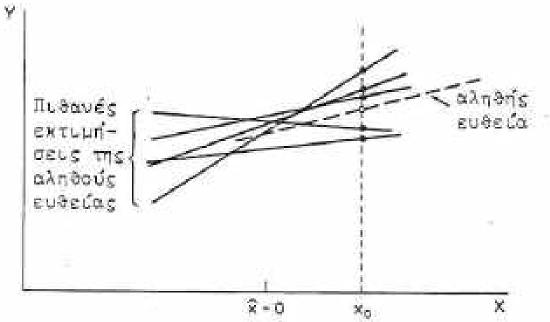
όπου η τιμή x_0 εκφράζεται σε απόκλιση από το μέσο $\bar{X}_i$ του δείγματος των 10 οικογενειών (σχήμα 2.9α). Η εκτιμήτρια $\hat{\mu}_0$ είναι αμερόληπτη εκτιμήτρια της αληθούς τιμής $\mu_0=\alpha+bx_0$ στον πληθυσμό, διότι:

$$\begin{aligned} E(\hat{\mu}_0) &= E(\hat{\alpha})+x_0E(\hat{b}) \\ &= \alpha+bx_0 \\ &= \mu_0 . \end{aligned} \qquad (2.10.2)$$

Επειδή η εκτιμήτρια $\hat{\mu}_0$ είναι γραμμική συνάρτηση των $\hat{\alpha}$



(α)



(β)

Σχήμα 2.9: Η εκτιμήτρια $\hat{\mu}_0$ και ο "στόχος" μ_0: α) μια μεμονωμένη εκτίμηση του μ_0. (β) Πολλές εκτιμήσεις του μ_0 από διάφορες παλινδρομήσεις στις οποίες διατηρούμε σταθερές τις τιμές της X.

και $\hat{b}$, που έχουν συνδιακύμανση ίση με μηδέν[1], θα έχουμε:

---

1. $\hat{\alpha}=\frac{\Sigma Y_i}{n}=\frac{\Sigma(\alpha+bx_i+e_i)}{n}=\alpha+\frac{\Sigma e_i}{n}$, άρα $\hat{\alpha}-\alpha=\frac{\Sigma e_i}{n}$ (1)

$$\hat{b}=\frac{\Sigma x_i(\alpha+bx_i+e_i)}{\Sigma x_i^2}=b+\Sigma w_i e_i,\ \text{όπου}\ w_i=\frac{x_i}{\Sigma x_i^2}.\ \text{Άρα}\ \hat{b}-b=\Sigma w_i e_i.\quad (2)$$

Από τις (1) και (2) έχουμε:

$$Cov(\hat{\alpha},\hat{b})=E(\hat{\alpha}-\alpha)(\hat{b}-b)=E\left(\frac{\Sigma e_i}{n}\cdot\Sigma w_i e_i\right)=\frac{1}{n}E(\Sigma w_i e_i^2)=\frac{1}{n}\Sigma w_i E(e_i^2)$$

$$=\frac{\sigma^2}{n}\Sigma w_i=0,\ \text{λόγω της (2.4.8).}$$

$$V(\hat{\mu}_0)=V(\hat{\alpha})+x_0^2V(\hat{b})$$
$$=\frac{\sigma^2}{n}+x_0^2\frac{\sigma^2}{\Sigma x_i^2}$$
$$=\sigma^2\left(\frac{1}{n}+\frac{x_0^2}{\Sigma x_i^2}\right).\qquad (2.10.3)$$

Ακόμη, κάτω από την ισχυρή υπόθεση (υ.6), οι εκτιμήτριες $\hat{\alpha}$ και $\hat{b}$ ακολουθούν την κανονική κατανομή και συνεπώς η $\hat{\mu}_0$, ως γραμμική συνάρτηση των $\hat{\alpha}$ και $\hat{b}$, θα ακολουθεί επίσης την κανονική κατανομή:

$$\hat{\mu}_0\sim N\left[\alpha+bx_0,\ \sigma^2\left(\frac{1}{n}+\frac{x_0^2}{\Sigma x_i^2}\right)\right].\qquad (2.10.4)$$

Σύμφωνα με τη μεθοδολογία που ακολουθήσαμε για τις εκτιμήτριες $\hat{\alpha}$ και $\hat{b}$, εύκολα προκύπτει ότι, το 95% διάστημα εμπιστοσύνης για την αληθή τιμή μ_0 της μέσης κατανάλωσης του πληθυσμού των οικογενειών που έχουν διαθέσιμο εισόδημα X_0 είναι το

$$\hat{\mu}_0\pm t_{0,025}\ \hat{\sigma}\sqrt{\frac{1}{n}+\frac{x_0^2}{\Sigma x_i^2}}.\qquad (2.10.5)$$

Αν το διαθέσιμο εισόδημα είναι $X_0=50$ χιλ. δραχμές, τότε, η μέση κατανάλωση των οικογενειών με διαθέσιμο εισόδημα 50 χιλ. δραχμών, εκτιμάται σε

$$\hat{\mu}_0=\hat{\alpha}+\hat{b}(X_0-\bar{X})$$
$$=22+0,485(50-34)$$
$$=29,76\ \text{χιλ. δραχμές}$$

και το 95% διάστημα εμπιστοσύνης για την αληθή τιμή της μέσης κατανάλωσης στον πληθυσμό των οικογενειών που έχουν διαθέσιμο εισόδημα 50 χιλ. δραχμές, είναι:

$$28,13<\mu_0<31,39\ \text{χιλ. δραχμές.}$$

ii) Πρόβλεψη για τη μεμονωμένη τιμή Y_0

Είναι και πάλι φανερό ότι η καλλίτερη εκτιμήτρια για την κατανάλωση της μεμονωμένης οικογένειας με διαθέσιμο εισόδημα X_0 είναι η τιμή $\hat{Y}_0$ που προσδιορίζεται από την παλινδρόμηση που εκτιμήσαμε:

$$\hat{Y}_0=\hat{\alpha}+\hat{b}x_0$$
$$=\hat{\mu}_0$$

η οποία είναι αμερόληπτη εκτιμήτρια της αληθούς τιμής Y_0. Στην προσπάθειά μας να προσδιορίσουμε το διάστημα εμπιστοσύνης για την Y_0 αντιμετωπίζουμε όλα τα προβλήματα που αντιμετωπίσαμε και για τον προσδιορισμό του διαστήματος εμπιστοσύνης για το μέσο μ_0 στην προηγούμενη περίπτωση. Έχουμε όμως και ένα πρόσθετο πρόβλημα που οφείλεται στο ότι προσπαθούμε να εκτιμήσουμε μία μόνο τιμή $\hat{Y}_0$ αντί του πιο σταθερού μέσου $E(\hat{Y}_0/X_0)$. Επομένως, στη διακύμανση (2.10.3) του μέσου $\hat{\mu}_0$ πρέπει να προσθέσουμε και τη διακύμανση του σφάλματος $\hat{e}_i$ της μεμονωμένης παρατήρησης. Άρα,

$$V(\hat{Y}_0)=\sigma^2\left(\frac{1}{n}+\frac{x_0^2}{\Sigma x_i^2}\right)+\sigma^2$$

$$=\sigma^2\left(\frac{1}{n}+\frac{x_0^2}{\Sigma x_i^2}+1\right) \qquad (2.10.6)$$

και το 95% διάστημα εμπιστοσύνης για την αληθή τιμή Y_0 είναι

$$\hat{\mu}_0 \pm t_{0,025}\,\hat{\sigma}\sqrt{\frac{1}{n}+\frac{x_0^2}{\Sigma x_i^2}+1}\,. \qquad (2.10.7)$$

Έτσι, αν το διαθέσιμο εισόδημα της οικογένειας είναι $X_0=50$ χιλ. δραχμές, τότε, η εκτίμηση της κατανάλωσής της είναι

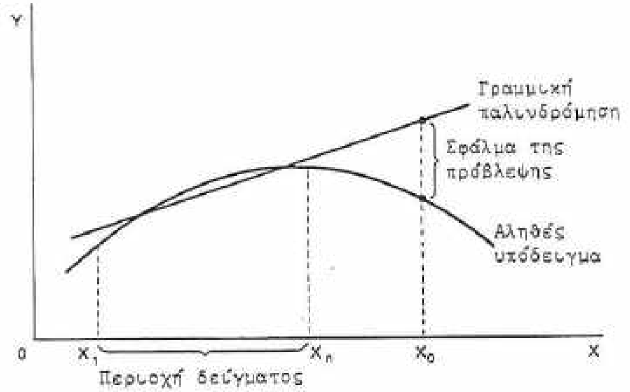
$$\hat{Y}_0=\hat{\mu}_0=29,76 \text{ χιλ. δραχμές}$$

και το 95% διάστημα εμπιστοσύνης για την αληθή τιμή της κατανάλωσης Y_0 είναι

$$26,34 < Y_0 < 33,18 \text{ χιλ. δραχμές.}$$

Από τις σχέσεις (2.10.3), (2.10.5), (2.10.6) και (2.10.7) γίνεται φανερό ότι, όσο αυξάνεται η τιμή του $x_0=X_0-\bar{X}$, δηλαδή όσο απομακρυνόμαστε από το μέσο $\bar{X}$ του δείγματος από το οποίο εκτιμήσαμε την παλινδρόμηση, τόσο η διακύμανση των $\hat{\mu}_0$ και $\hat{Y}_0$ αυξάνεται με αποτέλεσμα τα διαστήματα εμπιστοσύνης για τις αληθείς τιμές μ_0 και Y_0 να διευρύνονται και οι προβλέψεις να μην είναι αποτελεσματικές.

Πρέπει ακόμη να επισημάνουμε και ένα δεύτερο κίνδυνο που αντιμετωπίζουμε όταν η τιμή X_0 απομακρύνεται από το μέσο $\bar{X}$ του δείγματος ή από την περιοχή του δείγματος: αν η πραγματική σχέση που συνδέει τις μεταβλητές Y και X δεν είναι ευθεία (σχήμα 2.10) τότε η παλινδρόμηση $\hat{Y}=\hat{\alpha}+\hat{b}x$ που εκτι-



Σχήμα 2.10: Το σφάλμα της πρόβλεψης όταν το αληθές υπόδειγμα είναι μη γραμμικό.

μήσαμε μπορεί μεν να είναι μια καλή προσέγγιση της πραγματικής καμπύλης στην περιοχή του δείγματος αλλά όχι και πέρα από την περιοχή του δείγματος. Αν προσπαθήσουμε λοιπόν να διατυπώσουμε προβλέψεις για τιμές X_0 απομακρυσμένες από την περιοχή που καλύπτει το δείγμα, τότε, υπεισέρχεται στις προβλέψεις μας σοβαρό σφάλμα που οφείλεται στην κακή εξειδίκευση του υποδείγματος. Το σφάλμα εξειδίκευσης θα μειωθεί στο ελάχιστο αν προσαρμόσουμε στο δείγμα την κατάλληλη καμπύλη αντί της ευθείας. Το θέμα όμως της σωστής εξειδίκευσης του υποδείγματος θα μας απασχολήσει στο κεφάλαιο 4.

2.11. ΟΙ ΕΚΤΙΜΗΤΡΙΕΣ ΜΕΓΙΣΤΗΣ ΠΙΘΑΝΟΦΑΝΕΙΑΣ ΤΩΝ ΠΑΡΑΜΕΤΡΩΝ ΤΗΣ ΑΠΛΗΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

Εκτός από τις εκτιμήτριες ελαχίστων τετραγώνων, μπορούμε, κάτω από το σύνολο των ασθενών υποθέσεων (υ.1) έως (υ.5) και της ισχυρής υπόθεσης (υ.6) της κανονικότητας των υπό συνθήκη κατανομών πιθανότητας $P(Y_i/x_i)$, να υπολογίσουμε και τις εκτιμήτριες Μέγιστης Πιθανοφάνειας των παραμέτρων α, b και σ^2 της παλινδρόμησης

$$Y_i=\alpha+bx_i+e_i .$$

Κάτω από τις υποθέσεις (υ.1) έως (υ.6), η υπό συνθήκη κατανομή πιθανότητας των Y_i, όταν δίνεται η τιμή x_i της ανεξάρτητης μεταβλητής X είναι η

$$P(Y_i/x_i)=\frac{1}{2\pi\sigma^2}\, e^{-\frac{1}{2\sigma^2}[Y_i-(\alpha+bx_i)]^2}, \quad i=1,2,\ldots,n \qquad (2.11.1)$$

η οποία έχει μέσο $\alpha+bx_i$ και διακύμανση σ^2.

Λόγω της ανεξαρτησίας των κατανομών (2.11.1), η από κοινού συνάρτηση πιθανότητας των $Y_1,Y_2,\ldots,Y_n$ είναι το γινόμενο των επιμέρους συναρτήσεων πιθανότητας:

$$P(Y_1,Y_2,\ldots,Y_n/x_1,x_2,\ldots,x_n)=(2\pi\sigma^2)^{-\frac{n}{2}}\, e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}[Y_i-(\alpha+bx_i)]^2} \qquad (2.11.2)$$

και η συνάρτηση πιθανοφάνειας των α, b και σ^2 είναι, συνεπώς, η

$$L(\alpha,b,\sigma^2)=(2\pi\sigma^2)^{-\frac{n}{2}}\, e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}[Y_i-(\alpha+bx_i)]^2} \qquad (2.11.3)$$

ή, σε λογαριθμική μορφή,

$$Q=\ln L=-\frac{n}{2}\ln 2\pi-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}\sum_{i=1}^{n}[Y_i-(\alpha+bx_i)]^2. \qquad (2.11.4)$$

Οι συνθήκες πρώτης τάξης για τη μεγιστοποίηση της (2.11.4) οδηγούν στο σύστημα των εξισώσεων:

$$\left.\begin{aligned}
\frac{\partial Q}{\partial \alpha}&=\frac{1}{\sigma^2}\sum_{i=1}^{n}(Y_i-\tilde{\alpha}-\tilde{b}x_i)=0\\
\frac{\partial Q}{\partial b}&=\frac{1}{\sigma^2}\sum_{i=1}^{n}x_i(Y_i-\tilde{\alpha}-\tilde{b}x_i)=0\\
\frac{\partial Q}{\partial \sigma^2}&=-\frac{n}{2}\frac{1}{\tilde{\sigma}^2}+\frac{1}{2\tilde{\sigma}^4}\sum_{i=1}^{n}(Y_i-\tilde{\alpha}-\tilde{b}x_i)^2=0
\end{aligned}\right\} \qquad (2.11.5)$$

του οποίου η επίλυση προσδιορίζει τις εκτιμήτριες $\tilde{\alpha}$, $\tilde{b}$ και $\tilde{\sigma}^2$ μέγιστης πιθανοφάνειας των αληθών παραμέτρων α, b και σ^2:

$$\tilde{\alpha}=\bar{Y},\quad \tilde{b}=\frac{\Sigma x_iY_i}{\Sigma x_i^2} \quad \text{και} \quad \tilde{\sigma}^2=\frac{1}{n}\Sigma(Y_i-\tilde{\alpha}-\tilde{b}x_i)^2. \qquad (2.11.6)$$

Όπως εύκολα διαπιστώνουμε, οι εκτιμήτριες $\tilde{\alpha}$ και $\tilde{b}$ μέγιστης πιθανοφάνειας ταυτίζονται με τις εκτιμήτριες $\hat{\alpha}$ και $\hat{b}$ των ελαχίστων τετραγώνων και συνεπώς είναι αμερόληπτες εκτιμήτριες των α και b, ακολουθούν την κανονική κατανομή, και σύμφωνα με τη (Β.4.10) του παραρτήματος Β, είναι οι πιο αποτελεσματικές μέσα στην κλάση όλων των εκτιμητριών που ακολουθούν (και ασυμπτωτικά) την κανονική κατανομή, αμερόληπτων ή όχι. Η εκτιμήτρια $\tilde{\sigma}^2$ της διακύμανσης δεν είναι αμερόληπτη εκτιμήτρια της σ^2 διότι, όπως διαπιστώσαμε στην παράγραφο 2.5, η αμερόληπτη εκτιμήτρια είναι η

$$\hat{\sigma}^2=\frac{1}{n-2}\Sigma(Y_i-\hat{\alpha}-\hat{b}x_i)^2 .$$

Για να προχωρήσουμε στους ελέγχους υποθέσεων και στον προσδιορισμό διαστημάτων εμπιστοσύνης χρησιμοποιώντας τη $\tilde{\sigma}^2$, πρέπει να αποδείξουμε ότι

$$\frac{n\tilde{\sigma}^2}{\sigma^2}\sim\chi_{n-2}$$

και ότι η στατιστική αυτή είναι ανεξάρτητη από τις κατανομές $\tilde{\alpha}$ και $\tilde{b}$. Η απόδειξη αυτή θα δοθεί για τη γενικότερη περίπτωση της πολλαπλής παλινδρόμησης.

Κεφάλαιο 3

Πολλαπλή παλινδρόμηση

3.1. ΕΙΣΑΓΩΓΗ

Η πολλαπλή παλινδρόμηση είναι η επέκταση της απλής παλινδρόμησης στην περίπτωση που έχουμε περισσότερες από μια ανεξάρτητες μεταβλητές. Είναι η κατάλληλη μέθοδος για να εκτιμήσουμε την επίδραση που ασκούν συγχρόνως πάνω στη μεταβλητή Y ορισμένες μεταβλητές $X_1, X_2, \ldots, X_k$, χωρίς αυτό να σημαίνει ότι δε μπορούμε να προσδιορίσουμε τη μεμονωμένη επίδραση κάθε μιας από τις $X_1, X_2, \ldots, X_k$ πάνω στην Y. Η εισαγωγή περισσότερων ερμηνευτικών μεταβλητών έχει ως σκοπό αφενός τη μείωση των σφαλμάτων $\hat{e}_i$, άρα και της διακύμανσης $\hat{\sigma}^2$, και αφετέρου την εξάλειψη του σφάλματος εξειδίκευσης που διαπράττουμε όταν παραλείπουμε μια ή περισσότερες ερμηνευτικές μεταβλητές που ασκούν σημαντική επίδραση πάνω στην εξαρτημένη μεταβλητή.

Όπως είναι γνωστό από την οικονομική θεωρία, στη διαμόρφωση του ύψους της κατανάλωσης μιας οικογένειας επιδρούν, εκτός από το διαθέσιμο εισόδημα, και άλλοι παράγοντες όπως η περιουσία, ο αριθμός των παιδιών, η ηλικία των μελών της οικογένειας, ο τόπος διαμονής κλπ. Όσοι από τους παράγοντες αυτούς είναι μετρήσιμοι μπορούν να εισαχθούν ως ερμηνευτικές μεταβλητές στην παλινδρόμηση και να προσδιορισθεί η επίδραση που ασκούν, συνολικά αλλά και μεμονωμένα, στη διαμόρφωση του ύψους της κατανάλωσης.

Στο κεφάλαιο αυτό θα ασχοληθούμε με τη "γραμμική πολλαπλή παλινδρόμηση" η οποία είναι η επέκταση της απλής γραμ-

μικής παλινδρόμησης που εξετάσαμε στο προηγούμενο κεφάλαιο. Στην απλή γραμμική παλινδρόμηση προσαρμόσαμε στο δείγμα των παρατηρήσεων (Y_i, X_i), $i=1,2,\ldots,n$ την καλλίτερη ευθεία. Αν έχουμε δύο ανεξάρτητες μεταβλητές, τότε οι παρατηρήσεις μας θα είναι οι διαταγμένες τριάδες (Y_i, X_{i1}, X_{i2}), $i=1,2,\ldots,n$ και θα αντιπροσωπεύουν σημεία στο χώρο των τριών διαστάσεων. Στα σημεία αυτά θα προσαρμόσουμε το "επίπεδο ελαχίστων τετραγώνων". Γενικά, αν έχουμε κ ανεξάρτητες μεταβλητές τότε, οι παρατηρήσεις $(Y_i, X_{i1}, X_{i2}, \ldots, X_{ik})$, $i=1,2,\ldots,n$, θα είναι σημεία του χώρου των k+1 διαστάσεων και στα σημεία αυτά θα προσαρμόσουμε το "πολυεπίπεδο" ελαχίστων τετραγώνων.

3.2. ΤΟ ΚΛΑΣΙΚΟ ΓΡΑΜΜΙΚΟ ΥΠΟΔΕΙΓΜΑ ΤΗΣ ΠΟΛΛΑΠΛΗΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

Έστω ότι διαθέτουμε ένα δείγμα n παρατηρήσεων για τις μεταβλητές $Y, X_1, X_2, \ldots, X_k$, όπου Y είναι η εξαρτημένη μεταβλητή και $X_1, X_2, \ldots, X_k$ είναι οι ανεξάρτητες μεταβλητές που επηρεάζουν τη διαμόρφωση της Y.

Αναλυτικά οι n παρατηρήσεις του δείγματος είναι οι εξής:

$$\begin{array}{cccc} Y_1 & X_{11} & X_{12} \ldots & X_{1k} \\ Y_2 & X_{21} & X_{22} \ldots & X_{2k} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ Y_n & X_{n1} & X_{n2} & X_{nk} \end{array} \qquad (3.2.1)$$

Αν επιθυμούμε να προσαρμόσουμε στις n παρατηρήσεις του δείγματος (3.2.1) το καλλίτερο πολυεπίπεδο, τότε, όπως και στην απλή παλινδρόμηση, το κατάλληλο κριτήριο είναι το κριτήριο ελαχίστων τετραγώνων, σύμφωνα με το οποίο πρέπει να προσδιορίσουμε τις τιμές των συντελεστών

$$\hat{b}_0, \hat{b}_1, \ldots, \hat{b}_k$$

που ελαχιστοποιούν το άθροισμα των τετραγώνων των σφαλμάτων

$$\sum_{i=1}^{n} \hat{e}_i^2 \qquad (3.2.2)$$

b_0: σταθερός όρος παλινδρόμησης.

στην παλινδρόμηση

$$\boxed{Y_i = b_0 X_{i0} + b_1 X_{i1} + \cdots + b_k X_{ik} + e_i, \quad i=1,2,\ldots,n} \qquad (3.2.3)$$

όπου $X_{i0}=1$ για κάθε $i=1,2,\ldots,n$ έτσι ώστε ο όρος b_0 να είναι ο σταθερός όρος της παλινδρόμησης.

Για την εκτίμηση των συντελεστών $b_0, b_1, \ldots, b_k$ της (3.2.3) θα χρησιμοποιήσουμε τη θεωρία των πινάκων η οποία μας επιτρέπει να παρουσιάσουμε συνοπτικά τους υπολογισμούς και τα αποτελέσματά μας. Επίσης, δεν είναι απαραίτητο να εκφράσουμε τις τιμές των $X_1, X_2, \ldots, X_k$ σε αποκλίσεις από τους μέσους τους, όπως κάναμε στην απλή παλινδρόμηση, για να διευκολύνουμε τους υπολογισμούς μας.

Οι n παρατηρήσεις για την (3.2.3) μπορούν να παρουσιαστούν συνοπτικά ως εξής:

μονάδες

$$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}_{n\times 1} = \begin{bmatrix} X_{10} & X_{11} & \ldots & X_{1k} \\ X_{20} & X_{21} & \ldots & X_{2k} \\ \vdots & \vdots & & \vdots \\ X_{n0} & X_{n1} & \ldots & X_{nk} \end{bmatrix}_{n\times(k+1)} \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{bmatrix}_{(k+1)\times 1} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}_{n\times 1}$$

ή

$$\boxed{y = Xb + e} \qquad (3.2.4)$$

όπου

- y είναι το n×1 διάνυσμα των παρατηρήσεων πάνω στην εξαρτημένη μεταβλητή,
- X είναι ο n×(k+1) πίνακας των παρατηρήσεων πάνω στις ανεξάρτητες μεταβλητές,
- b είναι το (k+1)×1 διάνυσμα των συντελεστών που πρέπει να εκτιμήσουμε, και
- e είναι το n×1 διάνυσμα των σφαλμάτων.

Το άθροισμα των τετραγώνων που πρέπει να ελαχιστοποιήσουμε είναι το

$$\boxed{\begin{aligned} F(\hat{b}) &= \sum_{i=1}^{n} \hat{e}_i^2 = \hat{e}'\hat{e} \\ &= (y - X\hat{b})'(y - X\hat{b}) \end{aligned}}$$

(γραμμές x στήλες)

Τετραγωνικές Μορφές: $x'Ax$

$$\frac{\partial xA}{\partial x} = A, \quad \frac{\partial (x'Ax)}{\partial x} = 2Ax; \ A \text{ symmetric}$$

είναι αριθμός. π.χ. 3' = 3.

$$\begin{aligned} &= (y' - \hat{b}'X')(y - X\hat{b}) \\ &= y'y - y'X\hat{b} - \hat{b}'X'y + \hat{b}'X'X\hat{b} \\ &= y'y - 2\hat{b}'X'y + \hat{b}'X'X\hat{b}^{1}. \end{aligned} \qquad (3.2.5)$$

τετραγωνική μορφή.

Οι συνθήκες πρώτης τάξης για την ελαχιστοποίηση της (3.2.5) είναι

$$\frac{\partial F}{\partial \hat{b}} = -2X'y + 2X'X\hat{b} = 0$$

ή

$$X'X\hat{b} = X'y. \qquad (3.2.6)$$

Οι εξισώσεις (3.2.6) ονομάζονται "κανονικές εξισώσεις" και αποτελούν ένα σύστημα (k+1) εξισώσεων με (k+1) αγνώστους και η επίλυση του συστήματος αυτού θα μας δώσει τις εκτιμήτριες $\hat{b}_0, \hat{b}_1, \ldots, \hat{b}_k$ ελαχίστων τετραγώνων

Τύπος. $$\boxed{\hat{b} = (X'X)^{-1}X'y} \qquad (3.2.7)$$ Διάνυσμα εκτιμητριών.

της παλινδρόμησης (3.2.4).

Για να έχει λύση το σύστημα των κανονικών εξισώσεων (3.2.6) πρέπει ο τετραγωνικός πίνακας $X'X$, διαστάσεων (k+1)×(k+1) να είναι πλήρους βαθμού. Αυτό, όπως είναι γνωστό, θα ισχύει μόνο αν ο πίνακας X είναι πλήρους βαθμού δηλαδή αν

$$r(X) = k+1 \qquad (3.2.8)$$

Στην περίπτωση αυτή ο πίνακας $X'X$ είναι θετικός πεπερασμένος, υπάρχει ο αντίστροφός του $(X'X)^{-1}$ και το σύστημα των κανονικών εξισώσεων προσδιορίζει τις εκτιμήτριες (3.2.7) των ελαχίστων τετραγώνων.

Το διάνυσμα των τιμών της Y που εκτιμήσαμε από την παλινδρόμηση είναι:

---

1. Σημειώστε ότι $y'X\hat{b} = (y'X\hat{b})' = \hat{b}'X'y$ διότι το $y'X\hat{b}$ έχει διαστάσεις 1×1 δηλαδή είναι ένας πραγματικός αριθμός.

$$\boxed{\hat{y}=X\hat{b}} \qquad (3.2.9)$$

και το διάνυσμα των καταλοίπων της παλινδρόμησης είναι:

$$\hat{e}=y-\hat{y}$$
$$=y-X\hat{b}\,. \qquad (3.2.10)$$

Θα αποδείξουμε στη συνέχεια μερικές ιδιότητες του "πολυεπιπέδου" ελαχίστων τετραγώνων που θα μας είναι χρήσιμες στα όσα θα εκτεθούν στη συνέχεια:

ΙΔΙΟΤΗΤΕΣ ΠΟΛΥΕΠΙΠΕΔΟΥ Ε.Τ.

(i) $$\boxed{i'X\hat{b}=i'y} \qquad (3.2.11)$$

όπου i είναι το $n\times 1$ διάνυσμα του οποίου όλα τα στοιχεία είναι ίσα με τη μονάδα. Αν γράψουμε πιο αναλυτικά τις κανονικές εξισώσεις (3.2.6):

$$\begin{bmatrix} n & \Sigma X_1 & \Sigma X_2 & \dots & \Sigma X_k \\ \Sigma X_1 & \Sigma X_1^2 & \Sigma X_1X_2 & \dots & \Sigma X_1X_k \\ \Sigma X_2 & \Sigma X_2X_1 & \Sigma X_2^2 & \dots & \Sigma X_2X_k \\ \vdots & \vdots & \vdots & & \vdots \\ \Sigma X_k & \Sigma X_kX_1 & \Sigma X_kX_2 & \dots & \Sigma X_k^2 \end{bmatrix} \begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \hat{b}_2 \\ \vdots \\ \hat{b}_k \end{bmatrix} = \begin{bmatrix} \Sigma Y \\ \Sigma X_1Y \\ \Sigma X_2Y \\ \vdots \\ \Sigma X_kY \end{bmatrix} \qquad (3.2.12)$$

τότε εύκολα παρατηρούμε ότι η (3.2.11) είναι η πρώτη από τις κανονικές εξισώσεις.

(ii) $$\boxed{i'\hat{y}=i'X\hat{b}=i'y} \qquad (3.2.13)$$

όπως προκύπτει από τις (3.2.9) και (3.2.11). Η (3.2.13) εκφράζει αυτό που δείξαμε και στην απλή παλινδρόμηση:

$$\boxed{\bar{Y}_i=\bar{\hat{Y}}_i\,.}$$

ΤΟ ΠΟΛΥΕΠΙΠΕΔΟ ΤΩΝ Ε.Τ. ΠΕΡΝΑ ΠΑΝΤΑ ΑΠΟ ΤΟ ΚΕΝΤΡΟ ΒΑΡΟΥΣ ΤΟΥ ΔΕΙΓΜΑΤΟΣ.

(iii) $$\boxed{i'\hat{e}=i'y-i'\hat{y}=0} \qquad (3.2.14)$$

όπως προκύπτει από τις (3.2.10) και (3.2.13). Η (3.2.14) εκφράζει το ότι το άθροισμα των καταλοίπων της παλινδρόμησης $\Sigma\hat{e}_i$ είναι μηδέν, δηλαδή ο μέσος των $\hat{e}_i$ είναι μηδέν.

Γεωμετρική ερμηνεία του $\hat{y}'\hat{e}=0$.

Τα X κάθετα στο e ($X'e=0$ ή $X\perp e$). Τα $\hat{y}$ ∈ στο πολυεπίπεδο που ορίζουν τα X, άρα κι αυτά κάθετα στο e.

(iv) $$\begin{aligned} X'\hat{e} &= X'y-X'\hat{y} && [\text{λόγω της } (3.2.10)] \\ &= X'y-X'X\hat{b} && [\text{λόγω της } (3.2.9)] \\ &= 0 && [\text{λόγω της } (3.2.6)]. \end{aligned} \qquad (3.2.15)$$

Η (3.2.15) διαβεβαιώνει ότι οι στήλες του πίνακα X είναι διανύσματα ορθογώνια προς το διάνυσμα των καταλοίπων $\hat{e}$ της παλινδρόμησης. $X\perp\hat{e} \Rightarrow X'e=0$

(v) $\hat{y}'\hat{e}=\hat{b}'X'\hat{e}=0$ [λόγω των (3.2.9) και (3.2.15)] (3.2.16)

δηλαδή τα διανύσματα $\hat{y}$ και $\hat{e}$ είναι ορθογώνια. $\hat{y}\perp\hat{e} \Rightarrow \hat{y}'\hat{e}=0.$

3.3. ΤΟ ΣΤΑΤΙΣΤΙΚΟ ΥΠΟΔΕΙΓΜΑ ΤΗΣ ΠΟΛΛΑΠΛΗΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

Για την προσαρμογή του πολυεπιπέδου ελαχίστων τετραγώνων

$$\hat{y}=X\hat{b}$$

δε χρησιμοποιήσαμε μέχρι τώρα καμιά υπόθεση, εκτός από το ότι ο πίνακας X είναι πλήρους βαθμού.

Αν επιθυμούμε να προχωρήσουμε σε στατιστική επαγωγή για τις αληθείς τιμές των παραμέτρων $b_0, b_1, \dots, b_k$ στον πληθυσμό, τότε απαιτούνται ανάλογες υποθέσεις με εκείνες που κάναμε στην απλή παλινδρόμηση. Οι υποθέσεις αυτές διατυπώνονται, με τη βοήθεια των πινάκων και των διανυσμάτων, ως εξής:

ΑΣΘΕΝΕΙΣ ΥΠΟΘΕΣΕΙΣ.

$\boxed{(\text{υ}.1)}$ $$y=Xb+e$$

Η υπόθεση αυτή εκφράζει το πιστεύω του ερευνητή ότι η αληθής σχέση που συνδέει την εξαρτημένη μεταβλητή Y με τις ανεξάρτητες μεταβλητές $X_1, X_2, \dots, X_k$ είναι γραμμική.

$\boxed{(\text{υ}.2)}$ $$E(e)=0$$

Σύμφωνα με την υπόθεση αυτή, οι τυχαίες μεταβλητές $e_1, e_2, \dots, e_n$ έχουν μέσο μηδέν. Η ανάλογη υπόθεση για τις μεταβλητές $Y_1, Y_2, \dots, Y_n$ είναι η

$$E(y)=Xb.$$

$\boxed{(\upsilon.3)}$ $$C(e)=\sigma^2 I_n$$

όπου I_n είναι ο $n\times n$ μοναδιαίος πίνακας. Αναλυτικά η υπόθεση αυτή γράφεται:

$$\Sigma e_i e_j/n \qquad C(e)=E(ee')=\begin{bmatrix} \sigma^2 & 0 & \dots & 0 \\ 0 & \sigma^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \sigma^2 \end{bmatrix}$$

και αποτελεί συντομογραφία των δύο επιμέρους υποθέσεων:

(υ.3α) $V(e_i)=E(e_i^2)=\sigma^2, \quad i=1,2,\dots,n$

(υ.3β) $C(e_i e_j)=E(e_i e_j)=0, \quad i,j=1,2,\dots,n, \ (i\neq j).$

Σύμφωνα με την (υ.3α) τα σφάλματα $e_1,e_2,\dots,e_n$ έχουν την ίδια σταθερή διακύμανση σ^2, ενώ, σύμφωνα με την (υ.3β) τα σφάλματα e_i και e_j είναι ασυσχέτιστα για κάθε $i,j=1,2,\dots,n (i\neq j)$.

Η αντίστοιχη υπόθεση για τις μεταβλητές $Y_1,Y_2,\dots,Y_n$ είναι:

$$C(y)=\sigma^2 I_n .$$

$\boxed{(\upsilon.4)}$ $$r(X)=k+1 \quad \text{∃ η κ σταθερά}$$

Η υπόθεση αυτή είναι αναγκαία για να έχει λύση το σύστημα των κανονικών εξισώσεων και να ορίζονται οι εκτιμήτριες $\hat{b}$ ελαχίστων τετραγώνων. (Τεχνική υπόθεση που στην πράξη πληρείται πάντα)

$\boxed{(\upsilon.5)}$ Ο πίνακας X παραμένει σταθερός σε επανειλημμένα δείγματα. Σύμφωνα με την υπόθεση αυτή οι ανεξάρτητες μεταβλητές $X_1,X_2,\dots,X_k$ δεν είναι τυχαίες μεταβλητές αλλά οι τιμές τους παραμένουν σταθερές σε επανειλημμένα δείγματα.

$\boxed{(\upsilon.6)}$ Οι μεταβλητές $Y,X_1,X_2,\dots,X_k$ έχουν μετρηθεί χωρίς σφάλμα. Τη σημασία της υπόθεσης αυτής θα εξετάσουμε στα επόμενα.

Οι υποθέσεις (υ.1) έως και (υ.6) αποτελούν το σύνολο των ασθενών υποθέσεων για το στατιστικό υπόδειγμα της πολ-

(1) $y=Xb \Rightarrow X'y=X'Xb \Rightarrow X'y=(X'X)b \Rightarrow$
$\Rightarrow (X'X)^{-1}X'y=\hat{b} \Rightarrow \hat{b}=(X'X)^{-1}X'y.$

λαπλής παλινδρόμησης. Στην περίπτωση που το δείγμα μας είναι μικρό, θα απαιτηθεί, αργότερα, όπως και στην απλή παλινδρόμηση η ισχυρή υπόθεση:

$\boxed{(\upsilon.7)}$ $$e\sim N(0,\sigma^2 I_n)$$ ΙΣΧΥΡΗ ΥΠΟΘΕΣΗ (περίπτωση μικρού δείγματος)

δηλαδή ότι τα σφάλματα $e_1,e_2,\dots,e_n$ είναι ένα σύνολο ανεξάρτητων τυχαίων μεταβλητών που ακολουθούν την κανονική κατανομή με μέσο μηδέν και σταθερή διακύμανση σ^2 (τυχαίο δείγμα). Η αντίστοιχη υπόθεση για τις μεταβλητές $Y_1,Y_2,\dots,Y_n$ είναι η

$$y\sim N(Xb,\sigma^2 I_n).$$

3.4. Η ΜΑΘΗΜΑΤΙΚΗ ΕΛΠΙΔΑ ΚΑΙ Ο ΠΙΝΑΚΑΣ ΔΙΑΚΥΜΑΝΣΕΩΝ - ΣΥΝΔΙΑΚΥΜΑΝΣΕΩΝ ΤΩΝ ΕΚΤΙΜΗΤΡΙΩΝ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ. ΤΟ ΘΕΩΡΗΜΑ ΤΩΝ GAUSS-MARKOV

Για την εκτιμήτρια $\hat{b}$ ελαχίστων τετραγώνων του υποδείγματος της πολλαπλής γραμμικής παλινδρόμησης θα δείξουμε ότι κάτω από το σύνολο των ασθενών υποθέσεων ισχύουν τα εξής:

$$\boxed{E(\hat{b})=b}$$
$$\boxed{C(\hat{b})=\sigma^2(X'X)^{-1}.}$$

Λόγω της (υ.5) η εκτιμήτρια $\hat{b}$ είναι γραμμική συνάρτηση των $Y_1,Y_2,\dots,Y_n$, άρα και των $e_1,e_2,\dots,e_n$:

$$\hat{b}=(X'X)^{-1}X'y \tag{3.4.1}$$

$$=(X'X)^{-1}X'(Xb+e) \ [\text{λόγω της } (\upsilon.1)]$$

$$=\underbrace{(X'X)^{-1}X'X}_{I_{k+1}}b+(X'X)^{-1}X'e \Rightarrow \hat{b}=b+Ge. \qquad G \rightarrow (k+1)\times n \tag{3.4.2}$$

όπου ο $\overset{\text{Hat matrix}}{G}=(X'X)^{-1}X'$, είναι ένας σταθερός πίνακας. Για τον πίνακα G ισχύουν τα εξής:

$$GX=(X'X)^{-1}X'X=I_{k+1} \tag{3.4.3}$$

$$GG'=(X'X)^{-1}X'X(X'X)^{-1}=(X'X)^{-1}. \tag{3.4.4}$$

Άρα

$$E(\hat{b})=b+GE(e)$$
$$=b \qquad [\text{λόγω της } (υ.2)] \qquad (3.4.5)$$

και

$$C(\hat{b})=E(\hat{b}-b)(\hat{b}-b)'$$ outer product.
$$=E(Gee'G') \qquad [\text{λόγω της } (3.4.2)]$$
$$=GE(ee')G'$$
$$=G(\sigma^2 I)G' \qquad [\text{λόγω της } (υ.3)]$$
$$=\sigma^2 GG'$$
$$=\sigma^2(X'X)^{-1} \qquad [\text{λόγω της } (3.4.4)]$$
$$=\sigma^2 V \qquad (3.4.6)$$

όπου

$$V=(X'X)^{-1}=qq' \qquad V \to (k+1)\times(k+1)$$

Είναι φανερό ότι τα διαγώνια στοιχεία $\sigma^2 v_{ii}$ του πίνακα $\sigma^2 V$ είναι οι διακυμάνσεις των εκτιμητριών b_i, $i=0,1,2,\dots,k$, ενώ τα μη διαγώνια στοιχεία $\sigma^2 v_{ij}$ είναι οι συνδιακυμάνσεις των εκτιμητριών $\hat{b}_i$ και $\hat{b}_j$ $(i\neq j)$. Ακόμα, είναι φανερό ότι ο πίνακας $\sigma^2 V$ είναι συμμετρικός.

Και στην περίπτωση της πολλαπλής γραμμικής παλινδρόμησης ισχύει το

ΘΕΩΡΗΜΑ GAUSS-MARKOV

"Θεώρημα των Gauss-Markov": Στο σύνολο των γραμμικών και αμερόληπτων εκτιμητριών του διανύσματος b, η εκτιμήτρια $\hat{b}$ των ελαχίστων τετραγώνων είναι η άριστη (τα στοιχεία της έχουν τη μικρότερη διακύμανση). (3.4.7)

απόδειξη

Σύμφωνα με την (3.4.1) η εκτιμήτρια $\hat{b}$ των ελαχίστων τετραγώνων είναι γραμμική συνάρτηση των $Y_1, Y_2, \dots, Y_n$, και σύμφωνα με την (3.4.5) είναι και αμερόληπτη εκτιμήτρια του διανύσματος b των παραμέτρων στον πληθυσμό.

Έστω $\bar{b}$ μια άλλη γραμμική και αμερόληπτη εκτιμήτρια του διανύσματος b.

Αφού η $\bar{b}$ είναι γραμμική θα έχουμε:

$$\bar{b}=Hy$$
$$=H(Xb+e)$$
$$=HXb+He. \qquad (3.4.8)$$

Ο πίνακας των σταθμιστών H μπορεί να γραφτεί:

$$H=G+D \qquad (3.4.9)$$

όπου $G=(X'X)^{-1}X'$ και D ένας μη μηδενικός σταθερός πίνακας διαστάσεων $(k+1)\times n$.

Αφού η $\bar{b}$ είναι αμερόληπτη εκτιμήτρια του b πρέπει να ισχύει:

$$E(\bar{b})=b. \qquad (3.4.10)$$

Αλλά, από την (3.4.8) εύκολα προκύπτει ότι:

$$E(\bar{b})=HXb+HE(e)$$
$$=HXb$$

και για να ισχύει η (3.4.10) πρέπει ο πίνακας H να είναι τέτοιος ώστε

$$HX=I \qquad (3.4.11)$$

δηλαδή

$$HX=(G+D)X$$
$$=I+DX \qquad (GX=I)$$
$$=I$$

και αυτό θα ισχύει μόνο αν

$$DX=0. \qquad (3.4.12)$$

Στην περίπτωση αυτή η (3.4.8) γράφεται:

$$\bar{b}=b+He \qquad (3.4.13)$$

* $GD' = (X'X)^{-1}X'D' = (X'X)^{-1}(DX)'$. Αλλά $DX=0$ οπότε η $GD'=0$

$DG' = DX(X'X)^{-1} = 0 \cdot (X'X)^{-1} \Rightarrow DG'=0$

και ο πίνακας συνδιακυμάνσεων της εκτιμήτριας $\tilde{b}$ είναι τώρα:

$$
\begin{aligned}
C(\tilde{b}) &= E(\tilde{b}-b)(\tilde{b}-b)' \\
&= E(Hee'H') && [\text{λόγω της } (3.4.13)] \\
&= HE(ee')H' && [\text{λόγω της } (υ.2)] \\
&= \sigma^2 HH' \\
&= \sigma^2(G+D)(G'+D') && [\text{λόγω της } (3.4.9)] \\
&= \sigma^2(GG'+GD'+DG'+DD') \\
&= \sigma^2(X'X)^{-1}+\sigma^2(DD') && [\text{λόγω της } (3.4.12)]\ ^{*} \\
&= C(\hat{b})+\sigma^2(DD'). && (3.4.14)
\end{aligned}
$$

Επειδή ο πίνακας D είναι μη μηδενικός, ο τετραγωνικός πίνακας DD' είναι μη αρνητικός, δηλαδή τα διαγώνια στοιχεία του d_{ii} είναι μη αρνητικά. Άρα για τη διακύμανση κάθε στοιχείου της $\tilde{b}$ θα ισχύει: Τα διαγώνια στοιχεία του πίνακα [illegible] είναι οι διακυμάνσεις.

$$V(\tilde{b}_i)=V(\hat{b}_i)+\sigma^2 d_{ii}, \quad i=0,1,2,\ldots,k$$

ή Αμεροληψία.

$$V(\tilde{b}_i) \geqslant V(\hat{b}_i), \quad i=0,1,2,\ldots,k,$$

και αυτό αποδεικνύει το θεώρημα.

3.5. Η ΣΥΝΕΠΕΙΑ ΤΩΝ ΕΚΤΙΜΗΤΡΙΩΝ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

Στην προηγούμενη παράγραφο αποδείξαμε ότι η εκτιμήτρια $\hat{b}$ των ελαχίστων τετραγώνων είναι, για πεπερασμένα δείγματα, αμερόληπτη εκτιμήτρια του διανύσματος b. Κάτω από το σύνολο των ασθενών υποθέσεων και την πρόσθετη υπόθεση ότι ο πίνακας $\frac{1}{n}(X'X)$ συγκλίνει προς ένα μη ιδιάζοντα πίνακα, δηλαδή

(υ.8) $$\lim_{n\to\infty} \frac{1}{n}(X'X)=Q, \quad Q \text{ μη ιδιάζων}, \qquad (3.5.1)$$

μπορούμε να αποδείξουμε ότι η εκτιμήτρια $\hat{b}$ είναι και συνεπής εκτιμήτρια του διανύσματος b των αληθών παραμέτρων στον πληθυσμό.

Για να αποδείξουμε τη συνέπεια της εκτιμήτριας $\hat{b}$ αρκεί να δείξουμε ότι:

$$\lim_{n\to\infty} E(\hat{b}_i)=b_i, \quad i=0,1,2,\ldots,n \qquad (3.5.2)$$

και

$$\lim_{n\to\infty} V(\hat{b}_i)=0, \quad i=0,1,2,\ldots,n \qquad (3.5.3)$$

Δείξαμε στην προηγούμενη παράγραφο ότι για κάθε μέγεθος δείγματος n ισχύει

$$E(\hat{b})=b$$

και προφανώς αυτό θα ισχύει και για $n\to\infty$. Άρα η (3.5.2) ισχύει.

Λόγω των (3.4.6) και (3.5.1)

$$
\begin{aligned}
\lim_{n\to\infty} C(\hat{b}) &= \lim_{n\to\infty} \sigma^2(X'X)^{-1} \\
&= \lim_{n\to\infty} \frac{\sigma^2}{n}\left[\frac{1}{n}(X'X)^{-1}\right] \\
&= \lim_{n\to\infty} \frac{\sigma^2}{n} \cdot \lim_{n\to\infty} \frac{1}{n}(X'X)^{-1} \\
&= 0\, Q^{-1} \\
&= 0. \qquad \text{ΣΥΝΕΠΕΙΑ.}
\end{aligned}
$$

Άρα τα διαγώνια στοιχεία του πίνακα $\sigma^2(X'X)^{-1}$ τείνουν στο μηδέν και αυτό αποδεικνύει την (3.5.3).

Οι (3.5.2) και (3.5.3) εξασφαλίζουν ότι η $\hat{b}$ συγκλίνει προς το b κατά μέσον τετράγωνον, άρα και κατά πιθανότητα, δηλαδή

$$\operatorname{plim} \hat{b}=b. \qquad (3.5.4)$$

3.6. ΕΚΤΙΜΗΣΗ ΤΗΣ ΔΙΑΚΥΜΑΝΣΗΣ σ^2 ΤΩΝ ΣΦΑΛΜΑΤΩΝ

Για την εκτίμηση του πίνακα συνδιακυμάνσεων $\sigma^2(X'X)^{-1}$ της εκτιμήτριας $\hat{b}$ απαιτείται μια εκτιμήτρια $\hat{\sigma}^2$ της διακύμανσης σ^2 των σφαλμάτων και η οποία θα προκύψει φυσικά από τη διακύμανση των καταλοίπων $\hat{e}$ της παλινδρόμησης.

Γνωρίζουμε ότι:

$$\hat{e}=y-\hat{y}$$
$$=(Xb+e)-X\hat{b}$$
$$=Xb+e-X(b+Ge) \quad [\text{λόγω της } (3.4.2)]$$
$$=(I-XG)e$$
$$=Me \qquad \hat{e}'=(Me)'=e'M' \qquad (3.6.1)$$

όπου για τον πίνακα

$$M=I-XG \qquad M \to n\times n$$
$$=I-X(X'X)^{-1}X' \qquad (3.6.2)$$

εύκολα μπορούμε να δείξουμε ότι:

$$M=M', \text{ συμμετρικός και} \qquad (3.6.3)$$
$$M^2=M, \text{ αυτοδύναμος.} \qquad (3.6.4)$$

Συνεπώς, το άθροισμα των τετραγώνων των καταλοίπων της παλινδρόμησης είναι:

$$\Sigma\hat{e}_i^2=\hat{e}'\hat{e}$$
$$=e'M'Me \qquad M'M=MM=M^2=M$$
$$=e'Me \quad [\text{λόγω των } (3.6.3) \text{ και } (3.6.4)] \qquad (3.6.5)$$

τετραγωνική μορφή (δηλ. αριθμός)

και

ίχνος (για τετράγωνους πίνακες)

$tr(A)=\sum_i a_{ii}$ άθροισμα των διαγωνίων στοιχείων.

$$E(\hat{e}'\hat{e})=E(e'Me)$$
$$=E\,tr(e'Me) \quad [\text{διότι } e'Me \text{ είναι αριθμός}]$$
$$=E\,tr(Mee') \quad [\text{διότι } tr(AB)=tr(BA)]$$
$$=tr\,E(Mee')$$
$$=tr\,ME(ee')$$
$$=tr\,M(\sigma^2 I) \quad [\text{λόγω της } (\upsilon.3)]$$

Επειδή ο e'Me είναι αριθμός ισούται με το ίχνος του πίνακα.

$$=\sigma^2 tr M$$
$$=\sigma^2 tr[I_n-X(X'X)^{-1}X']$$
$$=\sigma^2\{tr I_n-tr[X(X'X)^{-1}X'] \quad [\text{διότι } tr(A+B)=tr(A)+tr(B)]$$
$$=\sigma^2\{n-tr[X'X(X'X)^{-1}] \quad \text{διότι } tr(AB)=tr(BA)$$
$$=\sigma^2(n-tr I_{k+1})$$
$$=\sigma^2[n-(k+1)]. \qquad (3.6.6)$$

Από την (3.6.5) γίνεται φανερό ότι η

$$\boxed{\hat{\sigma}^2=\frac{\hat{e}'\hat{e}}{n-(k+1)}} \quad \text{εκτιμήτρια του } \sigma^2 \qquad (3.6.7)$$

είναι αμερόληπτη εκτιμήτρια της διακύμανσης σ^2 των σφαλμάτων, άρα η εκτίμηση του πίνακα των συνδιακυμάνσεων της εκτιμήτριας $\hat{b}$ είναι:

$$\text{από τύπο (3.4.5)} \quad \boxed{\begin{aligned} C(\hat{b})&=\hat{\sigma}^2(X'X)^{-1}\\ &=\hat{\sigma}^2 V.\end{aligned}} \qquad (3.6.8)$$

3.7. ΔΙΑΣΤΗΜΑΤΑ ΕΜΠΙΣΤΟΣΥΝΗΣ ΚΑΙ ΕΛΕΓΧΟΙ ΥΠΟΘΕΣΕΩΝ ΓΙΑ ΤΑ ΣΤΟΙΧΕΙΑ ΤΟΥ ΔΙΑΝΥΣΜΑΤΟΣ *b*

Μέχρι τώρα, χρησιμοποιώντας το σύνολο των ασθενών υποθέσεων (υ.1) έως (υ.6), 1 προσδιορίσαμε τη μαθηματική ελπίδα και τον πίνακα συνδιακυμάνσεων της εκτιμήτριας $\hat{b}$, 2 αποδείξαμε το θεώρημα των Gauss-Markov και 3 υπολογίσαμε την εκτιμήτρια $\hat{\sigma}^2$ της διακύμανσης των σφαλμάτων.

Για να προχωρήσουμε στον προσδιορισμό διαστημάτων εμπιστοσύνης και στον έλεγχο υποθέσεων για τις αληθείς τιμές των στοιχείων του διανύσματος b στον πληθυσμό, απαιτείται να γνωρίζουμε και το είδος της κατανομής της εκτιμήτριας $\hat{b}$.

Για τον προσδιορισμό της κατανομής της εκτιμήτριας $\hat{b}$ απαιτείται κάποια υπόθεση για το είδος της κατανομής των σφαλμάτων e. Για πεπερασμένα δείγματα εισάγεται η ισχυρή υπόθεση (υ.7):

Ισχυρή υπόθεση για μικρά δείγματα.

$$e\sim N(0,\sigma^2 I). \qquad (3.7.1)$$

Αν το μέγεθος του δείγματος είναι μεγάλο δεν απαιτείται η ισχυρή υπόθεση της κανονικότητας για την κατανομή των σφαλμάτων. Σύμφωνα με το κεντρικό οριακό θεώρημα για να ισχύει η (3.7.1) αρκεί τα σφάλματα να είναι ανεξάρτητα και να ακολουθούν την ίδια κατανομή με μέσο μηδέν και διακύμανση σ^2.

Επειδή η εκτιμήτρια $\hat{b}$ είναι γραμμική μορφή των σφαλμάτων e:

$$\hat{b}=b+Ge \qquad (3.7.2)$$

εύκολα προκύπτει ότι

$$\hat{b}\sim N[b,G(\sigma^2 I)G']$$ η απόδειξη γίνεται σελ. 70.

$$\sim N[b,\sigma^2 GG']$$

$$\sim N[b,\sigma^2 (X'X)^{-1}]$$

$$\sim N(b,\sigma^2 V). \qquad (3.7.3)$$ V → σελ. 70

Γνωρίζοντας τώρα την κατανομή της εκτιμήτριας $\hat{b}$ μπορούμε να προχωρήσουμε στη στατιστική επαγωγή για το διάνυσμα b των αληθών τιμών των παραμέτρων $b_0, b_1, \ldots, b_k$, στον πληθυσμό.

Η εργασία μας θα διευκολυνθεί πολύ αν θεωρήσουμε τη γενικότερη περίπτωση της στατιστικής επαγωγής για κάποιο γραμμικό μετασχηματισμό του διανύσματος b. Σ'αυτό θα μας βοηθήσει η σχετική θεωρία του παραρτήματος Β την οποία ο σπουδαστής πρέπει να έχει μελετήσει πριν προχωρήσει στην ανάλυση που ακολουθεί.

Ας αναζητήσουμε την κατανομή της εκτιμήτριας

$$\hat{c}=R\hat{b} \qquad (3.7.4)$$

ή αναλυτικά

$$\begin{bmatrix} \hat{c}_1 \\ \hat{c}_2 \\ \vdots \\ \hat{c}_r \end{bmatrix} = \begin{bmatrix} r_{10} & r_{11} & r_{12} & \cdots & r_{1k} \\ r_{20} & r_{21} & r_{22} & \cdots & r_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ r_{r0} & r_{r1} & r_{r2} & \cdots & r_{rk} \end{bmatrix} \begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \vdots \\ \hat{b}_k \end{bmatrix}$$

όπου R είναι ένας πίνακας σταθερών, βαθμού $r \leqslant k+1$. Εύκολα παρατηρούμε ότι:

i) αν $R=I_{k+1}$ τότε $\hat{c}=\hat{b}$,

ii) αν $R=[1\ 0\ 0 \ldots 0]$ τότε $\hat{c}=\hat{b}_0$ και γενικά αν $R=[00\ldots010\ldots0]$ τότε $\hat{c}=\hat{b}_i$, αν η μονάδα βρίσκεται στη στήλη $(i+1)$, όπου $i+1 \leqslant k+1$.

iii) Αν $R=\begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix}$, με $m<k+1$, τότε $\hat{c}'=[\hat{b}_0, \hat{b}_1, \ldots, \hat{b}_m]$, δηλαδή η $\hat{c}$ έχει ως στοιχεία της τα πρώτα m στοιχεία της $\hat{b}$.

Γίνεται έτσι φανερό ότι η στατιστική επαγωγή για κάθε μιά από τις $b_0, b_1, \ldots, b_k$ χωριστά, ή για μερικές από αυτές συγχρόνως, ή για όλες συγχρόνως μπορεί να προκύψει από τη γενική περίπτωση της στατιστικής επαγωγής για το διάνυσμα $c=Rb$ αρκεί να δώσουμε στον πίνακα R την κατάλληλη μορφή. Λόγω της (Β.1.30, (iv)) του παραρτήματος Β θα ισχύει

$$\boxed{E(\hat{c})=Rb \quad \text{και} \quad C(\hat{c})=\sigma^2 RVR'}$$

και μπορούμε να αποδείξουμε (κάτω από το σύνολο των ασθενών υποθέσεων) ότι, στο σύνολο των γραμμικών και αμερόληπτων εκτιμητριών του διανύσματος $c=Rb$, η εκτιμήτρια $\hat{c}=R\hat{b}$ είναι η άριστη.

Ακόμα, από τις (3.7.3) και (Β.2.4) του παραρτήματος Β, εύκολα προκύπτει ότι

$$\hat{c}\sim N(Rb, \sigma^2 RVR'). \qquad (3.7.5)$$

Σύμφωνα με τη (Β.2.5) του παραρτήματος Β, το κανονικό διάνυσμα $\hat{c}=R\hat{b}$ μπορεί να τυποποιηθεί αν από αυτό αφαιρέσουμε το μέσο του Rb και στη συνέχεια πολλαπλασιάσουμε τη διαφορά από αριστερά με τον πίνακα K που είναι τέτοιος ώστε $K'K=(\sigma^2 RVR')^{-1}$. Άρα το διάνυσμα

$$z=K(R\hat{b}-Rb)$$

$$=K[R(X'X)^{-1}X'y-Rb]$$

$$=K[R(X'X)^{-1}X'(Xb+e)-Rb]$$
$$=K[Rb+R(X'X)^{-1}X'e-Rb]$$
$$=KQe \qquad (3.7.6)$$

όπου $Q=R(X'X)^{-1}X'$, είναι ένα τυποποιημένο κανονικό διάνυσμα:

$$z\sim N(0,I) \qquad (3.7.7)$$

και σύμφωνα με την (B.2.10) του παραρτήματος B, η τετραγωνική μορφή

$$z'z=e'Q'K'KQe$$
$$=e'Ne\sim \chi^2_r \qquad (3.7.8)$$

όπου $N=Q'K'KQ$ είναι εύκολο να αποδειχθεί ότι είναι ένας αυτοδύναμος πίνακας βαθμού r.

Σύμφωνα με την (3.6.5)

$$\Sigma e_i^2=\hat{e}'\hat{e}$$
$$=e'Me$$

όπου ο $M=I-X(X'X)^{-1}X'$ είναι αυτοδύναμος πίνακας βαθμού $n-(k+1)$ (διότι ο βαθμός ενός αυτοδύναμου πίνακα είναι ίσος με το ίχνος του). Επειδή το διάνυσμα e είναι ένα σφαιρικό κανονικό διάνυσμα:

$$e\sim N(0,\sigma^2 I)$$

και ο πίνακας M είναι αυτοδύναμος βαθμού $n-(k+1)$, σύμφωνα με την (B.2.8) του παραρτήματος B, η τετραγωνική μορφή

$$\hat{e}'\hat{e}=e'Me\sim\sigma^2\chi^2_{n-(k+1)}. \qquad (3.7.9)$$

Αν τώρα λάβουμε υπόψη ότι

$$\frac{\hat{e}'\hat{e}}{n-(k+1)}=\hat{\sigma}^2$$

τότε εύκολα προκύπτει ότι η στατιστική

$$\upsilon^2=\frac{\hat{e}'\hat{e}}{n-(k+1)}\cdot\frac{n-(k+1)}{\sigma^2}$$
$$=\frac{\hat{\sigma}^2[n-(k+1)]}{\sigma^2}\sim\chi^2_{n-(k+1)}. \qquad (3.7.10)$$

Ακόμα, επειδή οι αυτοδύναμοι πίνακες N και M είναι τέτοιοι ώστε:

$$NM=Q'K'KQ[I-X(X'X)^{-1}X']$$
$$=Q'K'K[Q-QX(X'X)^{-1}X']$$
$$=Q'K'K[Q-R(X'X)^{-1}X'X(X'X)^{-1}X']$$
$$=Q'K'K[Q-R(X'X)^{-1}X']$$
$$=Q'K'K[Q-Q],\ \text{διότι}\ Q=R(X'X)^{-1}X',$$
$$=0, \qquad (3.7.11)$$

σύμφωνα με την (B.2.10) του παραρτήματος B, ο λόγος

$$F=\left(\frac{z'z}{r}\right)\Big/\left[\frac{\upsilon^2}{n-(k+1)}\right]$$
$$=\frac{(R\hat{b}-Rb)'K'K(R\hat{b}-Rb)/r}{\hat{\sigma}^2/\sigma^2}$$
$$=\frac{(R\hat{b}-Rb)'(\sigma^2RVR')^{-1}(R\hat{b}-Rb)/r}{\hat{\sigma}^2/\sigma^2}$$
$$=\frac{1}{r\hat{\sigma}^2}(R\hat{b}-Rb)'(RVR')^{-1}(R\hat{b}-Rb)\sim F^r_{n-(k+1)} \qquad (3.7.12)$$

δηλαδή θα ακολουθεί την κατανομή F με r και $n-(k+1)$ βαθμούς ελευθερίας.

Η (3.7.12) είναι η βασική στατιστική την οποία θα χρησιμοποιήσουμε για τον προσδιορισμό διαστημάτων εμπιστοσύνης και για τον έλεγχο υποθέσεων για τα στοιχεία του διανύσματος c. Ακόμα, δίνοντας την κατάλληλη μορφή στον πίνακα R, μπορούμε να προσδιορίσουμε διαστήματα εμπιστοσύνης και να προβούμε σε ελέγχους υποθέσεων για τα στοιχεία του διανύσματος b καθώς και για οποιουσδήποτε γραμμικούς συνδυασμούς των στοιχείων του b.

i) Διάστημα εμπιστοσύνης για το διάνυσμα $c=Rb$

Από τους πίνακες της κατανομής $F^{r}_{n-(k+1)}$ μπορούμε π.χ. να προσδιορίσουμε την τιμή $F_{0,05}$ δεξιά της οποίας βρίσκεται το 5% της κατανομής. Αν επιθυμούμε να προσδιορίσουμε το διάστημα εμπιστοσύνης μέσα στο οποίο βρίσκεται, με πιθανότητα 95%, η αληθής τιμή του διανύσματος $c=Rb$ τότε πρέπει να ισχύει:

$$\Pr\left[\frac{1}{r\hat{\sigma}^2}(R\hat{b}-Rb)'(RVR')^{-1}(R\hat{b}-Rb) \leqslant F_{0,05}\right]=0.95 \qquad (3.7.13)$$

και από την ανισότητα μέσα στις αγκύλες εύκολα προκύπτει ότι το 95% διάστημα εμπιστοσύνης για το διάνυσμα $c=Rb$ είναι το

$$\frac{1}{r\hat{\sigma}^2}(R\hat{b}-Rb)'(RVR')^{-1}(R\hat{b}-Rb) \leqslant F_{0,05}. \qquad (3.7.14)$$

Το διάστημα εμπιστοσύνης (3.7.14), ως τετραγωνική μορφή βαθμού r, ορίζει ένα ελλειψοειδές στο χώρο των r διαστάσεων μέσα στο οποίο βρίσκεται, με πιθανότητα 95%, η αληθής τιμή του διανύσματος $c=Rb$.

Αν επιθυμούμε τώρα να ελέγξουμε τη γενική υπόθεση

$$H_0:\ Rb=h$$

κατά της

$$H_1:\ Rb\neq h$$

όπου h είναι ένα γνωστό διάνυσμα με r στοιχεία. δεν έχουμε παρά να θέσουμε στην (3.7.14) $Rb=h$. Στο αριστερό μέλος της (3.7.14) όλες οι ποσότητες είναι τώρα γνωστές. Αν λοιπόν η τιμή του αριστερού μέλους της (3.7.14) είναι μεγαλύτερη της $F_{0,05}$ των πινάκων για r και $n-(k+1)$ βαθμούς ελευθερίας, τότε απορρίπτουμε την H_0 και δεχόμαστε ότι $CR\neq h$, ενώ αν η τιμή του αριστερού μέλους της (3.7.14) είναι μικρότερη της $F_{0,05}$ τότε δεν μπορούμε να απορρίψουμε την H_0.

ii) Διάστημα εμπιστοσύνης για το διάνυσμα b

Για να προσδιορίσουμε το 95% διάστημα εμπιστοσύνης για το διάνυσμα b αρκεί, στις (3.7.13) και (3.7.14) να θέσουμε $R=I_{k+1}$. Στην περίπτωση αυτή έχουμε

$$\begin{aligned}(RVR')^{-1}&=V^{-1}\\&=X'X\end{aligned}$$

και το 95% διάστημα εμπιστοσύνης για το διάνυσμα b είναι το

$$\frac{1}{(k+1)\hat{\sigma}^2}(\hat{b}-b)'X'X(\hat{b}-b) \leqslant F_{0,05} \qquad (3.7.15)$$

και ορίζει ένα ελλειψοειδές στο χώρο των (k+1) διαστάσεων.

Αν τώρα θέλουμε π.χ. να ελέγξουμε την υπόθεση

$$H_0:\ b=0$$

κατά της

$$H_0:\ b\neq 0$$

τότε, κάτω από την υπόθεση H_0, η (3.7.15) παίρνει τη μορφή:

$$\frac{1}{(k+1)\hat{\sigma}^2}(\hat{b}'X'X\hat{b}) \leqslant F_{0,05} \qquad (3.7.16)$$

και αν η τιμή του αριστερού μέλους της (3.7.16) είναι μεγαλύτερη της $F_{0,05}$ για (k+1) και n-(k+1) βαθμούς ελευθερίας τότε απορρίπτουμε την H_0 και δεχόμαστε την H_1.

iii) Διάστημα εμπιστοσύνης για μερικά στοιχεία του διανύσματος b.

Ας υποθέσουμε ότι ζητείται να προσδιοριστεί ένα από κοινού διάστημα εμπιστοσύνης για m (m < k+1) στοιχεία του διανύσματος b. Χωρίς βλάβη της γενικότητας μπορούμε να αναδιατάξουμε τα στοιχεία του b και τις στήλες του πίνακα X έτσι ώστε τα στοιχεία που μας ενδιαφέρουν να καταλάβουν τις πρώτες m θέσεις του διανύσματος b και οι αντίστοιχες στήλες τις πρώτες m στήλες του πίνακα X. Μετά την αναδιάταξη αυτή το διάνυσμα b και ο πίνακας X μπορούν να γραφτούν ως εξής:

$$b=\begin{bmatrix} b_m \\ b^* \end{bmatrix} \quad \text{και} \quad X=[X_m : X_*].$$

Αν τώρα στον πίνακα R δώσουμε τη μορφή

$$R=[I_m : 0]$$

τότε

$$Rb=b_m$$

και

$$RVR'=R(X'X)^{-1}R'$$
$$=V_m$$

όπου V_m είναι ο τετραγωνικός πίνακας που σχηματίζεται από τις m πρώτες γραμμές και τις m πρώτες στήλες του πίνακα $(X'X)^{-1}$.

Στην περίπτωση αυτή το κοινό διάστημα εμπιστοσύνης για τα στοιχεία του b_m είναι το

$$\frac{1}{m\hat{\sigma}^2}(\hat{b}_m-b_m)'V_m^{-1}(\hat{b}_m-b_m)\leqslant F_{0,05} \qquad (3.7.17)$$

και ορίζει ένα ελλειψοειδές στο χώρο των m διαστάσεων. Η κατανομή F στην (3.7.17) έχει m και n-(k+1) βαθμούς ελευθερίας. Για τον έλεγχο υποθέσεων σχετικά με την αληθή τιμή του διανύσματος b_m στον πληθυσμό εργαζόμαστε όπως και στις προηγούμενες περιπτώσεις.

iv) Διάστημα εμπιστοσύνης για ένα στοιχείο b_i του b.

Στην περίπτωση αυτή δίνουμε στον πίνακα R τη μορφή:

$$R=[0,0,\dots,0,1,0,\dots,0]$$

δηλαδή είναι ένα οριζόντιο διάνυσμα 1×(k+1) που έχει μηδενικά σε όλες τις στήλες του, εκτός από τη στήλη (i+1) στην οποία έχει τη μονάδα. Εύκολα προκύπτει τώρα ότι:

$$R\hat{b}=b_i$$

και

$$RVR'=R(X'X)^{-1}R'$$
$$=\upsilon_{ii}$$

όπου υ_{ii} είναι το (i,i) διαγώνιο στοιχείο του πίνακα $(X'X)^{-1}$. Το διάστημα εμπιστοσύνης για την αληθή τιμή της παραμέτρου b_i στον πληθυσμό είναι:

$$\frac{1}{\hat{\sigma}^2}(\hat{b}_i-b_i)\upsilon_{ii}^{-1}(\hat{b}_i-b_i)\leqslant F_{0,05} \qquad (3.7.18)$$

ή

$$\frac{(\hat{b}_i-b_i)^2}{\hat{\sigma}^2\upsilon_{ii}}\leqslant F_{0,05} \qquad (3.7.19)$$

όπου η κατανομή F έχει 1 και n-(k+1) βαθμούς ελευθερίας. Η (3.7.19) ορίζει ένα ευθύγραμμο τμήμα (διάστημα) στο χώρο της 1 διάστασης (ευθεία των πραγματικών αριθμών).

Γνωρίζουμε ότι η κατανομή $F_{0,05}$ με 1 και n-(k+1) βαθμούς ελευθερίας είναι το τετράγωνο της κατανομής $t_{0,025}$ με n-(k+1) βαθμούς ελευθερίας:

$$F_{0,05}=t^2_{0,025}\ .$$

Άρα η (3.7.19) μπορεί να γραφτεί:

$$\frac{|\hat{b}_i-b_i|}{\hat{\sigma}\sqrt{\upsilon_{ii}}}\leqslant t_{0,025} \qquad (3.7.20)$$

και το 95% διάστημα εμπιστοσύνης για την αληθή τιμή της παραμέτρου b_i στον πληθυσμό είναι το

$$\hat{b}_i\pm\hat{\sigma}\sqrt{\upsilon_{ii}}\cdot t_{0,025} \qquad (3.7.21)$$

και είναι φανερό ότι το γινόμενο $\hat{\sigma}^2\upsilon_{ii}$ είναι η εκτίμηση της διακύμανσης του στοιχείου $\hat{b}_i$ του b. Άρα η (3.7.21) μπορεί να γραφτεί:

$$\hat{b}_i \pm t_{0,025}\sqrt{V(\hat{b}_i)}. \qquad (4.7.22)$$

Για τον έλεγχο υποθέσεων σχετικά με την τιμή της παραμέτρου b_i στον πληθυσμό μπορούμε να χρησιμοποιήσουμε ή την κατανομή F και την (3.7.19) ή την κατανομή t και την (3.7.20).

v) Έλεγχος για την ισχύ γραμμικών περιορισμών σχετικά με τα στοιχεία του b.

Ας υποθέσουμε ότι η συνάρτηση παραγωγής μιας επιχείρησης είναι της μορφής Gobb-Douglas:

$$Q_t = AL_t^{b_1}K_t^{b_2} \qquad (3.7.23)$$

όπου Q_t είναι η παραγόμενη ποσότητα, L_t η εισροή της εργασίας και K_t η εισροή του κεφαλαίου στην περίοδο t, t=1,2,...,n. Αν λογαριθμήσουμε τα δύο μέλη της (3.7.23) έχουμε:

$$Y_t = b_0 + b_1X_{1t} + b_2X_{2t} \qquad (3.7.24)$$

όπου

$$Y_t = \log Q_t$$

$$X_{1t} = \log L_t$$

$$X_{2t} = \log K_t$$

$$b_0 = \log A .$$

Αν διαθέτουμε χρονολογικές σειρές (t=1,2,...,n) για το ύψος της παραγωγής Q_t, για την εργασία L_t και για το κεφάλαιο K_t μπορούμε να εκτιμήσουμε τις παραμέτρους b_0, b_1 και b_2 της (3.7.24) με τη μέθοδο ελαχίστων τετραγώνων και γνωρίζουμε ότι οι παράμετροι b_1 και b_2 εκφράζουν την ελαστικότητα της παραγωγής ως προς την εργασία και το κεφάλαιο αντίστοιχα.

Ας υποθέσουμε τώρα ότι θέλουμε να ελέγξουμε την υπόθεση

$$H_0:\ b_1 + b_2 = 1 \qquad (3.7.25)$$

κατά της

$$H_1:\ b_1 + b_2 \neq 1 \qquad (3.7.26)$$

δηλαδή να ελέγξουμε αν η συνάρτηση παραγωγής της επιχείρησης χαρακτηρίζεται από σταθερή κλίμακα απόδοσης.

Στην περίπτωση αυτή ο πίνακας R θα είναι το διάνυσμα [0,1,1] οπότε:

$$R\hat{b} = [0,1,1]\begin{bmatrix}\hat{b}_0\\ \hat{b}_1\\ \hat{b}_2\end{bmatrix} = \hat{b}_1 + \hat{b}_2$$

και

$$RVR' = [0,1,1](X'X)^{-1}\begin{bmatrix}0\\1\\1\end{bmatrix}$$

$$= \upsilon$$

όπου υ είναι ένας πραγματικός αριθμός.

Κάτω από την υπόθεση H_0 η (3.7.14) γράφεται:

$$\frac{1}{\hat{\sigma}^2\upsilon}(\hat{b}_1 + \hat{b}_2 - 1)^2 \leq F_{0,05} \qquad (3.7.27)$$

όπου η κατανομή F έχει 1 και n-3 βαθμούς ελευθερίας. Αν το αριστερό μέλος της (3.7.27) είναι μεγαλύτερο από την τιμή $F_{0,05}$ των πινάκων, τότε απορρίπτουμε την H_0 και δεχόμαστε την H_1 ενώ αν το αριστερό μέλος της (3.7.27) είναι μικρότερο από την τιμή $F_{0,05}$ των πινάκων δεχόμαστε την H_0.

Αν τώρα θέλουμε να ελέγξουμε και την υπόθεση:

$$H_0:\ b_1 = b_2 \quad \text{ή} \quad b_1 - b_2 = 0$$

κατά της

$$H_1:\ b_1 \neq b_2 \quad \text{ή} \quad b_1 - b_2 \neq 0$$

δηλαδή αν οι ελαστικότητες της παραγωγής ως προς το κεφάλαιο και την εργασία είναι ίσες, τότε θα θέσουμε:

$$R = [0,1,-1]$$

και μπορούμε εύκολα να δείξουμε ότι η (3.7.14) θα πάρει τη μορφή

$$\frac{1}{\hat{\sigma}^2 \upsilon_*}(\hat{b}_1-\hat{b}_2)^2 \leqslant F_{0,05}. \tag{3.7.28}$$

Είναι φανερό ότι αντί των (3.7.27) και (3.7.28) μπορούμε να χρησιμοποιήσουμε τις

$$\frac{|\hat{b}_1+\hat{b}_2-1|}{\hat{\sigma}\sqrt{\upsilon_*}} \leqslant t$$

και

$$\frac{|b_1-b_2|}{\hat{\sigma}\sqrt{\upsilon_*}} \leqslant t$$

όπου η κατανομή t έχει n-3 βαθμούς ελευθερίας.

3.8. ΔΙΑΤΥΠΩΣΗ ΠΡΟΒΛΕΨΕΩΝ ΜΕ ΒΑΣΗ ΤΗΝ ΠΑΛΙΝΔΡΟΜΗΣΗ ΠΟΥ ΕΚΤΙΜΗΣΑΜΕ

Όπως στην απλή παλινδρόμηση, έτσι και στην πολλαπλή παλινδρόμηση, αν μας δοθεί ένα νέο σύνολο τιμών

$$x_0'=[1, x_{01}, x_{02}, \dots, x_{0k}]$$

για τις ανεξάρτητες μεταβλητές $X_0, X_1, X_2, \dots, X_k$ θα αναζητήσουμε:

i) το διάστημα εμπιστοσύνης για το μέσο $\mu_0=E(Y_0/x_0)$ και
ii) το διάστημα εμπιστοσύνης για τη μεμονωμένη τιμή Y_0 της εξαρτημένης μεταβλητής Y.

i) Το διάστημα εμπιστοσύνης για το μέσο μ_0.

Ως εκτίμηση $\hat{\mu}_0$ του μέσου $\mu_0=E(Y_0/x_0)$ της Y για επανειλημένες δοκιμές στις οποίες διατηρούμε σταθερό το διάνυσμα x_0, θα χρησιμοποιήσουμε την εκτίμησή του από την παλινδρόμηση:

$$\begin{aligned}\hat{\mu}_0 &= x_0'\hat{b}\\ &= \hat{b}_0+\hat{b}_1x_{01}+\hat{b}_2x_{02}+\dots+\hat{b}_kx_{0k}.\end{aligned} \tag{3.8.1}$$

Για να εκτιμήσουμε τώρα το διάστημα εμπιστοσύνης για την αληθή τιμή $\mu_0=x_0'b$ του μέσου στον πληθυσμό, αρκεί να θέσουμε στην (3.7.14)

$$R=x_0'. \tag{3.8.2}$$

Έτσι το 95% διάστημα εμπιστοσύνης για το μέσο μ_0 είναι:

$$\frac{1}{\hat{\sigma}^2}(x_0'\hat{b}-x_0'b)'(x_0'Vx_0)^{-1}(x_0'\hat{b}-x_0'b) \leqslant F_{0,05} \tag{3.8.3}$$

και αν λάβουμε υπόψη ότι οι όροι $x_0'\hat{b}-x_0'b$ και $x_0'Vx_0$ είναι πραγματικοί αριθμοί, τότε η (3.8.3) γράφεται

$$\frac{1}{\hat{\sigma}^2}\,\frac{(\hat{\mu}_0-\mu_0)^2}{x_0'Vx_0} \leqslant F_{0,05}, \tag{3.8.4}$$

όπου η κατανομή F έχει 1 και n-(k+1) βαθμούς ελευθερίας. Ακόμα, αν λάβουμε υπόψη ότι

$$t^2_{0,025} = F_{0,05} \tag{3.9.5}$$

τότε η (3.8.4) γράφεται:

$$\frac{1}{\hat{\sigma}}\,\frac{|\hat{\mu}_0-\mu_0|}{\sqrt{x_0'Vx_0}} \leqslant t_{0,025} \tag{3.8.6}$$

όπου η κατανομή t έχει n-(k+1) βαθμούς ελευθερίας. Από την (3.8.6) εύκολα προκύπτει ότι το 95% διάστημα εμπιστοσύνης για το μέσο μ_0 στον πληθυσμό είναι:

$$\hat{\mu}_0 \pm t_{0,025}\,\hat{\sigma}\sqrt{x_0'Vx_0}. \tag{3.8.7}$$

ii) Το διάστημα εμπιστοσύνης για τη μεμονωμένη τιμή Y_0 της Y.

Το διάστημα εμπιστοσύνης για την τιμή Y_0 της Y σε μια μόνο δοκιμή με το διάνυσμα x_0 είναι, όπως είδαμε και στην απλή παλινδρόμηση, το ίδιο με το (3.8.7) εκτός από το ότι στο υπόριζο πρέπει να προσθέσουμε τη μονάδα που αντιπροσωπεύει, πολλαπλασιασμένη επί $\hat{\sigma}^2$, τη διακύμανση της μεμονωμένης παρατήρησης. Άρα το διάστημα εμπιστοσύνης για τη μεμονωμένη τιμή Y_0 είναι:

$$\hat{\mu}_0 \pm t_{0,025}\ \hat{\sigma}\sqrt{x_0' V x_0 + 1} \qquad (3.8.8)$$

όπου η κατανομή t έχει, και πάλι, $n-(k+1)$ βαθμούς ελευθερίας.

Και στην περίπτωση της πολλαπλής παλίνδρόμησης επισημαίνουμε τους κινδύνους για τις προβλέψεις όταν το x_0 απομακρύνεται από το κέντρο βάρους

$$(0,\bar{X}_1,\bar{X}_2,\ldots,\bar{X}_k)$$

των ανεξάρτητων μεταβλητών $X_0,X_1,X_2,\ldots,X_k$.

Το πρόβλημα της διατύπωσης προβλέψεων με βάση την παλινδρόμηση που εκτιμήσαμε δεν εξαντλείται εδώ. Στο θέμα αυτό θα επανέλθουμε αργότερα.

3.9. ΓΕΩΜΕΤΡΙΚΗ ΕΡΜΗΝΕΙΑ ΤΗΣ ΜΕΘΟΔΟΥ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

Τις εκτιμήτριες ελαχίστων τετραγώνων

$$\hat{b}=(X'X)^{-1}X'y \qquad (3.9.1)$$

του γενικού γραμμικού υποδείγματος

$$y=Xb+e \qquad (3.9.2)$$

μπορούμε να υπολογίσουμε και να ερμηνεύσουμε και γεωμετρικά με τη βοήθεια του λογισμού των διανυσμάτων. Χωρίς βλάβη της γενικότητας και για να έχουμε εποπτική αντίληψη της γεωμετρίας θα περιοριστούμε στο χώρο των τριών διαστάσεων θεωρώντας την περίπτωση της απλής παλινδρόμησης:

$$Y_i=b_0+b_1X_i+e_i,\quad i=1,2,3 \qquad (3.9.3)$$

στην οποία πρέπει να εκτιμήσουμε τις παραμέτρους b_0 και b_1 από ένα δείγμα τριών παρατηρήσεων.

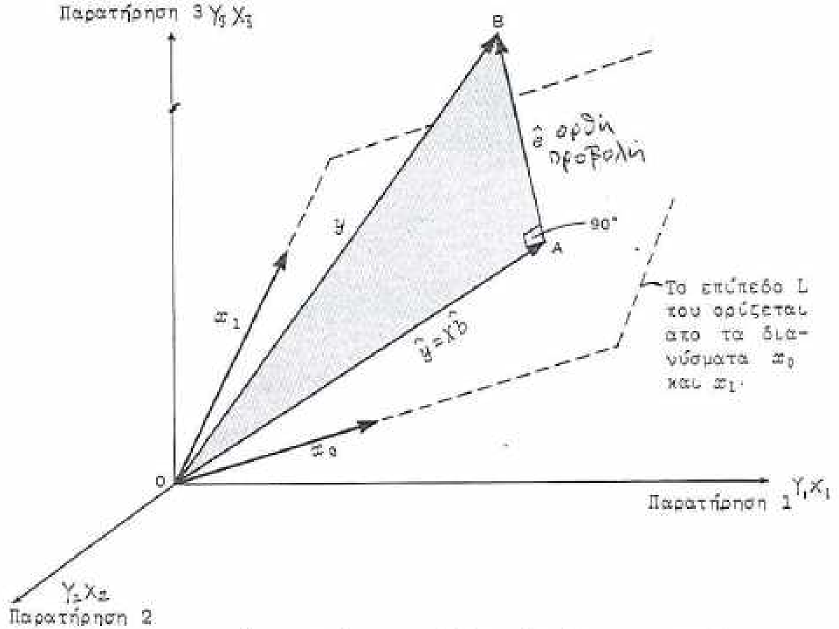
Η (3.9.3) γράφεται με τη μορφή (3.9.2) αν θέσουμε

$$y=\begin{bmatrix}Y_1\\Y_2\\Y_3\end{bmatrix},\quad X=\begin{bmatrix}1 & X_1\\1 & X_2\\1 & X_3\end{bmatrix}=[x_0,x_1],\quad b=\begin{bmatrix}b_0\\b_1\end{bmatrix},\quad e=\begin{bmatrix}e_1\\e_2\\e_3\end{bmatrix}. \qquad (3.9.4)$$

Οι τριάδες

$$y=\begin{bmatrix}Y_1\\Y_2\\Y_3\end{bmatrix},\quad x_0=\begin{bmatrix}1\\1\\1\end{bmatrix},\quad x_1=\begin{bmatrix}X_1\\X_2\\X_3\end{bmatrix} \qquad (3.9.5)$$

ορίζουν στο χώρο των τριών διαστάσεων του Σχήματος (3.1) τα



Σχήμα 3.1: Γεωμετρική ερμηνεία της μεθόδου ελαχίστων τετραγώνων. (Το επίπεδο L καλύπτεται από το σύνολο των γραμμικών συνδυασμών $xb=b_0x_0+b_1x_1$, το διάνυσμα $\hat{y}=x\hat{b}$ είναι η ορθή προβολή του διανύσματος y πάνω στο επίπεδο L, $\hat{e}$ είναι το διάνυσμα των καταλοίπων $y-\hat{y}$ και $\hat{b}$ οι εκτιμήτριες ελαχίστων τετραγώνων.

τρία διανύσματα y, x_0 και x_1. Τα διανύσματα x_0 και x_1 ορίζουν στο χώρο των τριών διαστάσεων ένα επίπεδο L που τα περιέχει και, όπως είναι γνωστό, το επίπεδο L "καλύπτεται" (παράγεται) από το σύνολο των γραμμικών συνδιασμών

$$Xb=[x_0 \quad x_1]\begin{bmatrix} b_0 \\ b_1 \end{bmatrix}=b_0x_0+b_1x_1 \qquad (3.9.6)$$

των διανυσμάτων x_0 και x_1, όπου b είναι ένα οποιοδήποτε διάνυσμα διαστάσεων 2×1. Έτσι, σε κάθε διάνυσμα b αντιστοιχεί ένα σημείο του επιπέδου L που ορίζεται από το γραμμικό συνδιασμό Xb και σε κάθε σημείο του επιπέδου L αντιστοιχεί ένα διάνυσμα b τέτοιο ώστε το Xb να απεικονίζεται στο σημείο αυτό. Αυτό είναι αντίστοιχο του ότι οι εξισώσεις δεν έχουν λύση καθεμιά

Το διάνυσμα y που ορίζεται από τις τρεις παρατηρήσεις πάνω στην εξαρτημένη μεταβλητή Y, κατά κανόνα δεν ανήκει στο επίπεδο L. Το πρόβλημα της "γραμμικής πολλαπλής παλινδρόμησης" είναι να προσδιορίσει το διάνυσμα

$$\hat{b}=\begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \end{bmatrix},$$

έτσι ώστε το διάνυσμα $X\hat{b}$ του επιπέδου L να είναι το πλησιέστερο προς το y.

Όπως γνωρίζουμε, το διάνυσμα $X\hat{b}$ είναι η "ορθή προβολή" του διανύσματος y πάνω στο επίπεδο L. Έτσι, στο σχήμα (3.1), ορίζεται το ορθογώνιο τρίγωνο ΟΑΒ με υποτείνουσα το διάνυσμα y και κάθετες πλευρές τα διανύσματα $X\hat{b}$ και $\hat{e}$. Τα στοιχεία του διανύσματος $\hat{e}$ είναι τα κατάλοιπα της παλινδρόμησης και, σύμφωνα με το λογισμό των διανυσμάτων, είναι η διαφορά των διανυσμάτων y και $X\hat{b}$:

$$\hat{e}=y-X\hat{b}. \qquad (3.9.7)$$

Η αντίστοιχη γεωμετρία στην περίπτωση της πολλαπλής παλινδρόμησης με k+1 ανεξάρτητες μεταβλητές $X_0, X_1, \ldots, X_k$, είναι απλή επέκταση των πιο πάνω εννοιών στο χώρο των n διαστάσεων, όπου n είναι ο αριθμός των παρατηρήσεων δηλαδή το μέγεθος του δείγματος, αλλά δεν μπορούμε να έχουμε εποπτική

αντίληψη της γεωμετρίας του χώρου των n διαστάσεων. Έτσι, στην περίπτωση της πολλαπλής παλινδρόμησης, το y και οι (k+1) στήλες (k+1 <n) $x_0, x_1, \ldots, x_k$ του πίνακα X είναι διανύσματα του χώρου n διαστάσεων. Επειδή έχουμε υποθέσει ότι ο πίνακας X είναι (k+1) βαθμού, τα διανύσματα $x_0, x_1, \ldots, x_k$ ορίζουν έναν "υπόχωρο" L^{k+1} του χώρου L^n των n διαστάσεων που "καλύπτεται" από αυτά, δηλαδή καλύπτεται από το σύνολο των γραμμικών συνδιασμών Xb, όπου $b'=[b_0, b_1, \ldots, b_k]$ είναι ένα 1×(k+1) διάνυσμα πραγματικών αριθμών. Το πρόβλημά μας είναι, και πάλι, να προσδιορίσουμε το διάνυσμα $\hat{b}'=[\hat{b}_0, \hat{b}_1, \ldots, \hat{b}_k]$ έτσι ώστε το διάνυσμα $X\hat{b}$ να είναι η ορθή προβολή του y στον υπόχωρο L^{k+1}. Και στην περίπτωση αυτή τα διανύσματα $\hat{e}$ και $X\hat{b}$ θα είναι οι κάθετες πλευρές του ορθογωνίου τριγώνου με υποτείνουσα το y και το διάνυσμα $\hat{e}$ των καταλοίπων της παλινδρόμησης θα ορίζεται από την (3.9.7).

Το ότι το διάνυσμα των καταλοίπων $\hat{e}$ είναι κάθετο στο διάνυσμα $X\hat{b}$ σημαίνει ότι το εσωτερικό τους γινόμενο πρέπει να μηδενίζεται, δηλαδή

$$\begin{aligned}(X\hat{b})'\hat{e}&=\hat{b}'X'\hat{e}\\ &=0\end{aligned} \qquad (3.9.10)$$

από την οποία προκύπτει ότι

$$X'\hat{e}=0. \quad (\hat{b}\neq 0) \qquad (3.9.11)$$

Η συνθήκη (3.9.11) σε συνδιασμό με την (3.9.7) προσδιορίζει τις εκτιμήτριες ελαχίστων τετραγώνων. Πράγματι, σύμφωνα με τις (3.9.11) και (3.9.7), πρέπει:

$$\begin{aligned}X'\hat{e}&=X'(y-X\hat{b})\\ &=X'y-X'X\hat{b}\\ &=0\end{aligned} \qquad (3.9.12)$$

και η (3.9.12) προσδιορίζει τις "κανονικές εξισώσεις":

$$X'X\hat{b}=X'y. \qquad (3.9.13)$$

Επειδή έχουμε υποθέσει ότι ο πίνακας X είναι πλήρους βαθμού, το σύστημα (3.9.13) έχει μοναδική λύση την

$$\boxed{\hat{b}=(X'X)^{-1}X'y} \qquad (3.9.14)$$

η οποία προσδιορίζει τις εκτιμήτριες ελαχίστων τετραγώνων.

Από το ορθογώνιο τρίγωνο του Σχ. (3.1) γίνεται φανερό ότι το αρχικό διάνυσμα y αναλύεται σε δύο συνιστώσες:

(i) τη συνιστώσα

$$\begin{aligned}\hat{y}&=X\hat{b}\\ &=X(X'X)^{-1}X'y\end{aligned} \qquad (3.9.15)$$

Μέρος που ερμηνεύεται

που ανήκει στο "πολυεπίπεδο" L^{k+1} και εκφράζει το μέρος του y που "ερμηνεύεται" από την παλινδρόμηση, και

(ii) τη συνιστώσα

$$\begin{aligned}\hat{e}&=y-\hat{y}\\ &=y-X(X'X)^{-1}X'y\\ &=[I-X(X'X)^{-1}X']y\\ &=My\end{aligned} \qquad (3.9.16)$$

Ανερμήνευτο μέρος της παλινδρόμησης.

που είναι κάθετη στο "πολυεπίπεδο" L^{k+1} και εκφράζει το μέρος του y που "δεν ερμηνεύεται" από την παλινδρόμηση.

Άρα, σύμφωνα με τη γεωμετρία των διανυσμάτων,

$$\begin{aligned}y&=\hat{y}+\hat{e}\\ &=X\hat{b}+\hat{e}.\end{aligned} \qquad (3.9.17)$$

Από την ανάλυση που προηγήθηκε έγινε φανερό ότι μπορούμε να υπολογίσουμε τις εκτιμήτριες $\hat{b}$ των ελαχίστων τετραγώνων χρησιμοποιώντας μόνο τη γεωμετρία των διανυσμάτων και να φτάσουμε στα ίδια ακριβώς συμπεράσματα που φτάσαμε χρησιμοποιώντας την άλγεβρα των διανυσμάτων και των πινάκων.

3.10. Ο ΣΥΝΤΕΛΕΣΤΗΣ ΠΟΛΛΑΠΛΟΥ ΠΡΟΣΔΙΟΡΙΣΜΟΥ R^2

Συνεχίζοντας τη γεωμετρική ανάλυση της παραγράφου (3.9) μπορούμε να εφαρμόσουμε το Πυθαγόρειο θεώρημα στο ορθογώνιο τρίγωνο ΟΑΒ του σχήματος (3.1) οπότε θα έχουμε:

$$\|y\|^2=\|\hat{y}\|^2+\|\hat{e}\|^2 \qquad (3.10.1)$$

Πυθαγόρειο θεώρημα.

όπου γενικά

$$\|z\|^2=z'z=\Sigma z_i^2 \qquad (3.10.2)$$

είναι το τετράγωνο του μέτρου του διανύσματος z.

Άρα, η (3.10.1) γράφεται:

$$y'y=\hat{y}'\hat{y}+\hat{e}'\hat{e} \qquad (3.10.3)$$

ή

$$\boxed{\Sigma Y_i^2=\Sigma\hat{Y}_i^2+\Sigma\hat{e}_i^2} \qquad (3.10.4)$$

η οποία αναλύει το άθροισμα των τετραγώνων ΣY_i^2 στις δύο ορθογώνιες συνιστώσες $\Sigma\hat{Y}_i^2$ και $\Sigma\hat{e}_i^2$.

Περισσότερο όμως ενδιαφέρον παρουσιάζει η ανάλυση της διακύμανσης της εξαρτημένης μεταβλητής Y. Γνωρίζουμε ότι:

$$\Sigma(Y_i-\bar{Y})^2=\Sigma Y_i^2-n\bar{Y}^2 \qquad (3.10.5)$$

και

$$\Sigma(\hat{Y}_i-\bar{\hat{Y}})^2=\Sigma\hat{Y}_i^2-n\bar{\hat{Y}}^2 \qquad (3.10.6)$$

Αλλά, σύμφωνα με την (3.2.13), $\bar{Y}_i=\bar{\hat{Y}}_i$ και, σύμφωνα με την (3.2.14), $\bar{\hat{e}}_i=0$. Αν λοιπόν από τα δύο μέλη της (3.10.4) αφαιρέσουμε την ποσότητα $n\bar{Y}^2=n\bar{\hat{Y}}^2$, λόγω των (3.10.5) και (3.10.6), θα έχουμε:

ορθογώνιες συνιστώσες

$$\Sigma(Y_i-\bar{Y})^2=\Sigma(\hat{Y}_i-\bar{\hat{Y}})^2+\Sigma\hat{e}_i^2 \qquad (3.10.7)$$

ή

$$\frac{\Sigma(Y_i-\bar{Y})^2}{n-1}=\frac{\Sigma(\hat{Y}_i-\bar{\hat{Y}})^2}{n-1}+\frac{\Sigma\hat{e}_i^2}{n-1} \qquad (3.10.8)$$

από την οποία γίνεται φανερό ότι η "συνολική διακύμανση" της

Y αναλύεται στο "ερμηνευόμενο μέρος" από την παλινδρόμηση που είναι η διακύμανση των $\hat{Y}_i$ και στο "ανερμήνευτο μέρος" που είναι η διακύμανση των καταλοίπων $\hat{e}_i$.

Ο "συντελεστής πολλαπλού προσδιορισμού" ορίζεται στην πολλαπλή παλινδρόμηση ως ο λόγος του μέρους της διακύμανσης της Y που ερμηνεύεται από την παλινδρόμηση προς τη συνολική διακύμανση της Y:

R^2

$$R^2 = \frac{\Sigma(\hat{Y}_i-\bar{\hat{Y}})^2}{\Sigma(Y_i-\bar{Y})^2}$$

$$=1-\frac{\Sigma\hat{e}_i^2}{\Sigma(Y_i-\bar{Y})^2} \tag{3.10.9}$$

και από την (3.10.9) γίνεται φανερό ότι

$$0 \leq R^2 \leq 1. \tag{3.10.10}$$

Όπως και στην απλή παλινδρόμηση, ο συντελεστής R^2 είναι ένα μέτρο της ερμηνευτικής ικανότητας της παλινδρόμησης και όσο η τιμή του πλησιάζει προς τη μονάδα τόσο καλλίτερα το γραμμικό υπόδειγμα που εκτιμήσαμε προσαρμόζεται στα διαθέσιμα στατιστικά στοιχεία.

Παρόλα αυτά πρέπει να είμαστε προσεκτικοί όταν χρησιμοποιούμε το συντελεστή R^2 για να αξιολογήσουμε μια παλινδρόμηση ή για να συγκρίνουμε δύο ή περισσότερες παλινδρομήσεις. Σχετικά πρέπει να επισημάνουμε τα εξής:

Προσοχη

(i) Στην εκτίμηση με τη μέθοδο ελαχίστων τετραγώνων, όταν εισάγουμε μια ή περισσότερες πρόσθετες ερμηνευτικές μεταβλητές το άθροισμα των τετραγώνων των καταλοίπων ελαττώνεται. Η ιδιότητα αυτή είναι καθαρά μαθηματική και δεν εξαρτάται από το αν οι ερμηνευτικές μεταβλητές που εισάγουμε είναι σχετικές με την αιτιώδη σχέση που εκφράζεται με την παλινδρόμηση. Κατά συνέπεια, όταν εισάγονται στην παλινδρόμηση πρόσθετες ερμηνευτικές μεταβλητές ο συντελεστής R^2 αυξάνεται.

(ii) Γενικά, αν αντί της Y χρησιμοποιήσουμε ως εξαρτη-

μένη μεταβλητή κάποιο γραμμικό συνδιασμό της Y και μιας ή περισσότερων από τις ερμηνευτικές μεταβλητές $X_1, X_2, \ldots, X_k$, τότε το μεν άθροισμα $\Sigma\hat{e}_i^2$ δε μεταβάλλεται ενώ το άθροισμα $\Sigma(Y_i-\bar{Y})^2$ -άρα και ο συντελεστής R^2- μεταβάλλεται.

Ας υποθέσουμε π.χ. ότι η εξίσωση της ζήτησης ενός αγαθού είναι:

$$Q_t=b_0+b_1P_t+b_2Q_{t-1}+e_t \tag{3.10.11}$$

όπου Q_t είναι η ζητούμενη ποσότητα και P_t η τιμή στην περίοδο t. Ας υποθέσουμε ακόμα ότι ο ερευνητής ενδιαφέρεται να ερμηνεύσει όχι τη ζήτηση αλλά τις μεταβολές της ζήτησης του αγαθού. Στην περίπτωση αυτή μπορεί να αφαιρέσει και από τα δύο μέλη της (3.10.11) την ποσότητα Q_{t-1} οπότε προκύπτει η εξίσωση:

$$Q_t-Q_{t-1}=b_0+b_1P_t+b_3Q_{t-1}+e_t \tag{3.10.12}$$

όπου $b_3=b_2-1$. Συγκρίνοντας τις εξισώσεις (3.10.11) και (3.10.12), εύκολα διαπιστώνουμε ότι έχουν την ίδια ερμηνευτική ικανότητα (οι συντελεστές b_0 και b_1 είναι ίδιοι και $b_3=b_2-1$) και ότι δίνουν το ίδιο άθροισμα καταλοίπων $\Sigma\hat{e}_t^2$. Εντούτοις η διακύμανση της εξαρτημένης μεταβλητής στις δύο εξισώσεις είναι διαφορετική (κατά κανόνα οι μεταβολές της ζήτησης έχουν μικρότερη διακύμανση από τη ζήτηση) με αποτέλεσμα ο συντελεστής R^2, όπως προκύπτει από την (3.10.9), να είναι διαφορετικός (στην περίπτωσή μας μικρότερος).

Παρατηρούμε λοιπόν ότι, ενώ οι δύο παλινδρομήσεις, χρησιμοποιώντας τα ίδια στατιστικά στοιχεία, θα δώσουν ταυτόσημες πληροφορίες για τη ζήτηση του αγαθού, εντούτοις η πρώτη θα έχει μεγαλύτερο R^2 από τη δεύτερη. Γενικά λοιπόν μπορούμε να πούμε ότι "μια υψηλή τιμή του R^2 δείχνει πάντα μία καλή προσαρμογή, αλλά μια χαμηλή τιμή του R^2 δε σημαίνει απαραίτητα ότι η παλινδρόμηση είναι ακατάλληλη".

Σε παρόμοια συμπεράσματα οδηγούμαστε και στην περίπτωση που επιχειρούμε να συγκρίνουμε παλινδρομήσεις που έχουν τις ίδιες ερμηνευτικές μεταβλητές αλλά η εξαρτημένη μεταβλητή έχει διαφορετική συναρτησιακή μορφή. Π.χ. στις παλινδρομήσεις

$$Y_t=b_0+b_1X_{1t}+b_2X_{2t}+e_t$$

$$\log Y_t=c_0+c_1X_{1t}+c_2X_{2t}+u_t$$

η εξειδίκευση των υποδειγμάτων, η ερμηνεία των συντελεστών, τα κατάλοιπα καθώς και ο υπολογισμός του συντελεστή R^2 είναι τελείως διαφορετικά και δεν παρέχουν καμιά κοινή βάση για τη σύγκριση των δύο εξισώσεων με τη βοήθεια του R^2.

Από την ανάλυση που προηγήθηκε γίνεται φανερό ότι ο συντελεστής R^2 μπορεί να χρησιμοποιηθεί μόνο για τη σύγκριση παλινδρομήσεων που έχουν την ίδια εξαρτημένη μεταβλητή. Και επειδή ο συντελεστής R^2 αυξάνεται όταν προσθέτουμε νέες ερμηνευτικές μεταβλητές, απαιτείται ο επιπλέον περιορισμός, οι παλινδρομήσεις να έχουν τον ίδιο αριθμό ερμηνευτικών μεταβλητών.

Στα πλαίσια αυτά μπορούμε να επισημάνουμε αρκετές περιπτώσεις στις οποίες ο συντελεστής R^2 είναι πράγματι χρήσιμος οδηγός. Ας θεωρήσουμε π.χ. την απλή συνάρτηση παραγωγής Gobb-Douglas:

$$\log Q_t=b_0+b_1\log K_t+b_2L_t+e_t$$

στην οποία το κεφάλαιο K_t έχει μετρηθεί σωστά, αλλά υπάρχουν αρκετοί δείκτες για τη μέτρηση της εργασίας L_t, τους οποίους δε μπορούμε να αξιολογήσουμε θεωρητικά. Σε μια τέτοια περίπτωση μπορούμε να προχωρήσουμε σε εμπειρική αξιολόγηση των δεικτών για τη μέτρηση της εργασίας εκτιμόντας χωριστή παλινδρόμηση για κάθε δείκτη και επιλέγοντας την παλινδρόμηση με τον υψηλότερο συντελεστή R^2. Και πάλι όμως πρέπει να τονίσουμε ότι, η διαδικασία αυτή οδηγεί σε σωστά συμπεράσματα μόνο αν η συναρτησιακή μορφή του υποδείγματος έχει εξειδικευθεί σωστά και όλες οι άλλες μεταβλητές είναι σωστά ορισμένες.

Ένα άλλο παράδειγμα σωστής χρησιμοποίησης του συντελεστή R^2 βρίσκουμε στην περίπτωση της εμπειρικής επιλογής της

κατάλληλης χρονικής υστέρισης μιας ερμηνευτικής μεταβλητής. Ας υποθέσουμε π.χ. ότι το επιτόκιο r προσδιορίζεται διαχρονικά από την εξίσωση:

$$r_t=b_0+b_1M_{t-s}+b_2Y_t+b_3P_t+e_t$$

όπου M είναι η προσφορά χρήματος, Y_t το ακαθάριστο εθνικό εισόδημα και P_t ένας δείκτης για τη μέτρηση του πληθωρισμού. Ο εμπειρικός προσδιορισμός της κατάλληλης χρονικής υστέρισης για την προσφορά χρήματος μπορεί να γίνει με την εκτίμηση παλινδρομήσεων για διαφορετικές τιμές του $s=0,1,2,...$ και την επιλογή της παλινδρόμησης με τον υψηλότερο συντελεστή R^2 και πάλι όμως με την προϋπόθεση ότι η εξειδίκευση του υποδείγματος είναι σωστή. Στην εμπειρική ανάλυση θα συναντήσουμε πολλά τέτοια παραδείγματα ορθής χρησιμοποίησης του συντελεστή R^2.

Ακόμα πρέπει να επισημάνουμε ότι αν τα στατιστικά στοιχεία είναι χρονολογικές σειρές τότε οι τιμές του R^2 είναι συνήθως υψηλές διότι οι εξαρτημένες και οι ανεξάρτητες μεταβλητές περιέχουν κοινές διαχρονικές τάσεις. Απεναντίας, όταν τα στατιστικά στοιχεία είναι διαστρωματικά οι τιμές του R^2 τείνουν να είναι χαμηλές λόγω των μεγάλων διαφορών που παρουσιάζει η συμπεριφορά των ιδιωτών και της έλλειψης κοινών τάσεων στη διαμόρφωση της συμπεριφοράς τους. Έτσι, τιμές του R^2 μεγαλύτερες από 0.6 θεωρούνται ικανοποιητικές για διαστρωματικές αναλύσεις ενώ στην περίπτωση διαχρονικών αναλύσεων αναμένονται τιμές του R^2 γύρω στο 0.9.

Ας επανεξετάσουμε τώρα την περίπτωση της εισαγωγής νέων ερμηνευτικών μεταβλητών σε μια παλινδρόμηση και, συγκεκριμένα, ας θεωρήσουμε τις παλινδρομήσεις:

$$Y_i=b_0+b_1X_{1i}+e_{1t}$$

$$Y_i=c_0+c_1X_{1i}+c_2X_{2i}+e_{2t}$$

τις οποίες δε μπορούμε να συγκρίνουμε με τη βοήθεια του R^2 διότι η δεύτερη παλινδρόμηση, αναγκαία, θα έχει μεγαλύτερο

R^2 από την πρώτη. Αλλά, ενώ το άθροισμα των τετραγώνων των καταλοίπων $\Sigma\hat{e}^2_{2t}$ της δεύτερης παλινδρόμησης είναι, αναγκαία, μικρότερο από το αντίστοιχο άθροισμα $\Sigma\hat{e}^2_{1t}$ της πρώτης παλινδρόμησης (αυτό είναι "κέρδος" διότι θα έχουμε μικρότερες εκτιμήσεις για τη διακύμανση $\hat{\sigma}^2$ των σφαλμάτων και για τις διακυμάνσεις των συντελεστών c_0, c_1, c_2, με συνέπεια τα διαστήματα εμπιστοσύνης να είναι "στενότερα" και οι προβλέψεις πιο αξιόπιστες) εντούτοις θα έχουμε μείωση των "βαθμών ελευθερίας" του δείγματος από n-2 σε n-3 διότι θα εκτιμήσουμε ένα ακόμα συντελεστή (η απώλεια βαθμών ελευθερίας είναι "ζημία" διότι συνεπάγεται υψηλότερες θεωρητικές τιμές για τις κατανομές t και F με αποτέλεσμα τη διεύρυνση των διαστημάτων εμπιστοσύνης και τη μείωση της αξιοπιστίας των προβλέψεων).

Εύκολα λοιπόν τίθεται το ερώτημα κατά πόσο με την εισαγωγή μιας ή περισσότερων νέων ερμηνευτικών μεταβλητών το "κέρδος" από τη μείωση του αθροίσματος των τετραγώνων των καταλοίπων "ισοφαρίζεται" από τη "ζημία" της απώλειας των αντίστοιχων βαθμών ελευθερίας.

Για να απαντήσουμε στο ερώτημα αυτό χρησιμοποιούμε, στη γενική περίπτωση, τη στατιστική $\frac{\Sigma(\varepsilon_i-\bar{\varepsilon})^2}{n-(k+1)} = \frac{\Sigma\varepsilon_i^2}{n-(k+1)}$

Διακύμανση καταλοίπων

$$V(\hat{e}) = \frac{\sum_{i=1}^{n} \hat{e}_i^2}{n-(k+1)} \qquad (3.10.13)$$

όπου k είναι ο αριθμός των ανεξάρτητων μεταβλητών $X_1, X_2, \ldots, X_k$.

Όταν εισάγονται νέες μεταβλητές στην παλινδρόμηση τότε ο αριθμητής της (3.10.13) αναγκαία ελαττώνεται, αλλά αυτό δε σημαίνει ότι ελαττώνεται αναγκαία και η τιμή του λόγου $V(\hat{e})$ διότι ελαττώνεται συγχρόνως και ο παρονομαστής με την αύξηση της τιμής του k. Έτσι, η στατιστική $V(\hat{e})$, λαμβάνει υπόψη και τους βαθμούς ελευθερίας ενώ αυτό δε συμβαίνει με το συντελεστή R^2.

Με τη βοήθεια της στατιστικής $V(\hat{e})$ μπορούμε να ορίσουμε ένα νέο συντελεστή, ανάλογο με τον R^2, ως εξής:

$$\bar{R}^2 = 1 - \frac{V(\hat{e})}{V(y)}$$

$$= 1 - \frac{(\Sigma\hat{e}_i^2)/[n-(k+1)]}{\Sigma(Y_i-\bar{Y})^2/(n-1)} \qquad (3.10.14)$$

όπου $V(y)=\Sigma(Y_i-\bar{Y})^2/(n-1)$ είναι η διακύμανση της Y. Επειδή ο παρονομαστής $V(y)$ δεν εξαρτάται από τον αριθμό των ανεξάρτητων μεταβλητών, η τιμή του συντελεστή $\bar{R}^2$ εξαρτάται μόνο από την τιμή του αριθμητή $V(\hat{e})$ και αντίστροφα.

Ο συντελεστής $\bar{R}^2$ είναι γνωστός γενικά με την ονομασία "συντελεστής πολλαπλού προσδιορισμού διορθωμένος ως προς τους βαθμούς ελευθερίας" και είναι φανερό ότι η τιμή του μπορεί και να ελαττωθεί από την εισαγωγή πρόσθετων ερμηνευτικών μεταβλητών ενώ, αντίθετα, η τιμή του R^2 αναγκαία αυξάνεται.

Βέβαια, αυτό δε σημαίνει ότι, γενικά, η παλινδρόμηση που δίνει τη μικρότερη τιμή στο λόγο $V(\hat{e})$ -άρα τη μεγαλύτερη τιμή στο συντελεστή $\bar{R}^2$- είναι αναγκαία η καλλίτερη. Και αυτό γιατί η απόφαση για την εισαγωγή ή όχι μιας ανεξάρτητης μεταβλητής πρέπει να λαμβάνεται με θεωρητικά κριτήρια -οικονομική θεωρία- και όχι με το εμπειρικό κριτήριο της τιμής του $\bar{R}^2$.

Γενικά μπορούμε να πούμε ότι αν η εισαγωγή μιας ή περισσότερων ερμηνευτικών μεταβλητών αυξάνει την τιμή του $\bar{R}^2$ τότε η προβλεπτική ικανότητα της παλινδρόμησης αυξάνεται και αντίστροφα. Επομένως αν ο σκοπός του ερευνητή είναι η διατύπωση προβλέψεων τότε πρέπει να επιλέξει την παλινδρόμηση που δίνει τη μεγαλύτερη τιμή για το συντελεστή $\bar{R}^2$. Συμπέρασμα. Πρόβλεψη

Αντίθετα, αν ο σκοπός του ερευνητή είναι ο έλεγχος υποθέσεων για τις αληθείς τιμές των παραμέτρων της παλινδρόμησης στον πληθυσμό, τότε, απαιτούνται αμερόληπτες εκτιμήτριες των παραμέτρων. Στην περίπτωση αυτή, βασική επιδίωξη του ερευνητή πρέπει, όπως θα δούμε παρακάτω, να είναι η σωστή εξειδίκευση του υποδείγματος και η εισαγωγή των ερμηνευτικών μεταβλητών που προσδιορίζει η οικονομική θεωρία ανεξάρτητα από την επίδρασή τους στην τιμή του συντελεστή $\bar{R}^2$.

Από τις σχέσεις (3.10.9) και (3.10.14) προκύπτει ότι:

$$\boxed{1-\bar{R}^2=(1-R^2)\,\frac{n-1}{n-(k+1)}} \qquad (3.10.15)$$

$$(1-\bar{R}^2)(n-1)-k(1-\bar{R}^2)=(1-R^2)(n-1)$$

$$1-\bar{R}^2=1-R^2+\frac{k}{n-1}(1-\bar{R}^2)$$

$$\bar{R}^2=R^2-\frac{k}{n-1}\left[(1-R^2)\frac{n-1}{n-(k+1)}\right] \quad \text{[λόγω της (3.10.15)]}$$

ή, τελικά,

$$\boxed{\bar{R}^2=R^2-(1-R^2)\frac{k}{n-(k+1)}} \qquad (3.10.16)$$

Από τη σχέση (3.10.16) εύκολα προκύπτει ότι

$$\boxed{\bar{R}^2<R^2} \qquad (3.10.17)$$

(εκτός αν $R^2=1$ ή $k=0$) και ότι ο συντελεστής $\bar{R}^2$ μπορεί να πάρει και αρνητική τιμή αρκεί να ισχύει:

$$R^2<(1-R^2)\frac{k}{n-(k+1)}$$

ή

$\bar{R}^2$: αρνητική τιμή →

$$R^2<\frac{k}{n-1}. \qquad (3.10.18)$$

3.11. Η ΑΝΑΛΥΣΗ ΤΗΣ ΔΙΑΚΥΜΑΝΣΗΣ ΣΤΗΝ ΠΟΛΛΑΠΛΗ ΠΑΛΙΝΔΡΟΜΗΣΗ. ΕΛΕΓΧΟΣ ΤΗΣ ΣΤΑΤΙΣΤΙΚΗΣ ΣΗΜΑΝΤΙΚΟΤΗΤΑΣ ΤΗΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

Σύμφωνα με τις εξισώσεις (3.10.7) και (3.10.8) η συνολική διακύμανση της εξαρτημένης μεταβλητής Y αναλύεται σε δύο ορθογώνιες συνιστώσες:

$$\Sigma(Y_i-\bar{Y})^2=\Sigma(\hat{Y}_i-\bar{\hat{Y}})^2+\Sigma\hat{e}_i^2, \quad (\bar{Y}=\bar{\hat{Y}}), \qquad (3.11.1)$$

όπου $\Sigma(\hat{Y}_i-\bar{\hat{Y}})^2$ είναι το μέρος της διακύμανσης της Y που ερμηνεύεται από την παλινδρόμηση (δηλαδή από τις διακυμάνσεις των $X_1, X_2, \ldots, X_k$) και $\Sigma\hat{e}_i^2$ είναι το μέρος της διακύμανσης της Y που παραμένει ανερμήνευτο.

Η (3.11.1) γράφεται

$$\Sigma y_i^2=\Sigma\hat{y}_i^2+\Sigma\hat{e}_i^2 \qquad (3.11.2)$$

αποκλίσεις από τον μέσον.

και οι όροι της μπορούν να υπολογιστούν εύκολα μετά την εκτίμηση της παλινδρόμησης.

ΠΙΝΑΚΑΣ 3.1: ΑΝΑΛΥΣΗ ΤΗΣ ΔΙΑΚΥΜΑΝΣΗΣ ΣΤΗΝ ΠΟΛΛΑΠΛΗ ΠΑΛΙΝΔΡΟΜΗΣΗ

Αιτίες της διακύμανσης της Y	Σχετικό άθροισμα τετραγώνων	Βαθμοί ελευθερίας	Μέσο σφάλμα τετραγώνου	Η στατιστική F*
$X_1, X_2, \ldots, X_k$	$\Sigma\hat{y}_i^2$	k	$\frac{\Sigma\hat{y}_i^2}{k}$	$F^*=\frac{\Sigma\hat{y}_i^2/k}{\Sigma\hat{e}_i^2/[n-(k+1)]}$
Κατάλοιπα	$\Sigma\hat{e}_i^2$	$n-(k+1)$	$\frac{\Sigma\hat{e}_i^2}{n-(k+1)}$	$=\frac{R^2/k}{(1-R^2)/[n-(k+1)]}$
Σύνολο της διακύμανσης της Y	Σy_i^2	$n-1$		Η στατιστική F* συγκρίνεται με τη θεωρητική κατανομή $F^k_{n-(k+1)}$ των πινάκων

Γνωρίζοντας ότι οι αυτοδύναμες τετραγωνικές μορφές $\Sigma\hat{y}_i^2$ και $\Sigma\hat{e}_i^2$ είναι ανεξάρτητες (ορθογώνιες) με k και $n-(k+1)$ βαθμούς ελευθερίας αντίστοιχα, εύκολα συμπεραίνουμε ότι ο λόγος

$$\boxed{F^*=\frac{\Sigma\hat{y}_i^2/k}{\Sigma\hat{e}_i^2/[n-(k+1)]}} \qquad (3.11.3)$$

ακολουθεί την κατανομή F με k και $n-(k+1)$ βαθμούς ελευθερίας.

Η στατιστική F μπορεί να χρησιμοποιηθεί για τον έλεγχο της στατιστικής σημαντικότητας ολόκληρης της παλινδρόμησης, δηλαδή για τον έλεγχο της υπόθεσης

$$H_0:\ b^*=0$$

κατά της υπόθεσης

$$H_1:\ b^*\neq 0$$

όπου

$$b'^*=[b_1, b_2, \ldots, b_k].$$

Κατά τα γνωστά, αν $F^* > F^k_{n-(k+1)}$ απορρίπτουμε την υπόθεση H_0 και δεχόμαστε ότι η ερμηνευτική ικανότητα της παλινδρόμησης είναι στατιστικά σημαντική, δηλαδή δεχόμαστε ότι υπάρχει στατιστικά σημαντική γραμμική σχέση μεταξύ της Y και των $X_1, X_2, \ldots, X_k$, ενώ αν $F^* < F^k_{n-(k+1)}$ δεχόμαστε ότι η γραμμική σχέση μεταξύ της Y και των $X_1, X_2, \ldots, X_k$ δεν είναι στατιστικά σημαντική.

Μπορούμε ακόμα να δείξουμε ότι ο λόγος F εκφράζεται και ως συνάρτηση του συντελεστή πολλαπλού προσδιορισμού R^2:

$$F^* = \frac{\Sigma\hat{y}_i^2/k}{\Sigma\hat{e}_i^2/[n-(k+1)]}$$

$$= \frac{(\Sigma\hat{y}_i^2/\Sigma y_i^2)/k}{(\Sigma\hat{e}_i^2/\Sigma y_i^2)/[n-(k+1)]}$$

$$= \frac{R^2/k}{(1-R^2)/[n-(k+1)]} \cdot \qquad (3.11.4)$$

3.12. ΕΛΕΓΧΟΣ ΓΙΑ ΤΗ ΒΕΛΤΙΩΣΗ ΤΗΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ ΑΠΟ ΤΗΝ ΕΙΣΑΓΩΓΗ ΠΡΟΣΘΕΤΩΝ ΕΡΜΗΝΕΥΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ

Μέχρι τώρα είδαμε ότι το μόνο κατάλληλο μέτρο για να ελέγξουμε τη βελτίωση της ερμηνευτικής ικανότητας της παλινδρόμησης από την εισαγωγή νέων ερμηνευτικών μεταβλητών, είναι ο συντελεστής $\bar{R}^2$.

Η ανάλυση της διακύμανσης της προηγούμενης παραγράφου μας δίνει τη δυνατότητα να διατυπώσουμε ένα ακόμα κριτήριο για τον έλεγχο της βελτίωσης της παλινδρόμησης από την εισαγωγή νέων ερμηνευτικών μεταβλητών.

Ας υποθέσουμε ότι εκτιμήσαμε την παλινδρόμηση

$$\hat{Y}_i = \hat{b}_0 + \hat{b}_1 X_{i1} + \hat{b}_2 X_{i2} + \cdots + \hat{b}_k X_{ik} \qquad (3.12.1)$$

και ότι $\Sigma\hat{y}_i^2$ και $\Sigma\hat{e}_i^2$ είναι, αντίστοιχα, το ερμηνευόμενο και το ανερμήνευτο μέρος της διακύμανσης της Y.

Στη συνέχεια εισάγουμε μερικές ακόμα ερμηνευτικές μεταβλητές έτσι ώστε ο συνολικός αριθμός τους να αυξηθεί σε m (m > k). Η νέα παλινδρόμηση είναι η

ΠΙΝΑΚΑΣ 3.2: Ανάλυση της διακύμανσης για τον έλεγχο της βελτίωσης της παλινδρόμησης όταν εισάγονται νέες ερμηνευτικές μεταβλητές.

Αιτίες της διακύμανσης της Y	Σχετικό άθροισμα τετραγώνων	Βαθμοί ελευθερίας	Μέσο σφάλμα τετραγώνου	Η στατιστική F*
$X_1, X_2, \ldots, X_k$	$\Sigma\hat{y}_i^2$	k	$\frac{\Sigma\hat{y}_i^2}{k}$	$F^* = \frac{\Sigma\hat{y}_i^2/k}{\Sigma\hat{e}_i^2/[n-(k+1)]}$
$X_1, X_2, \ldots, X_k, \ldots, X_m$	$\Sigma\hat{\hat{y}}_i^2$	m	$\frac{\Sigma\hat{\hat{y}}_i^2}{m}$	$F^* = \frac{\Sigma\hat{\hat{y}}_i^2/m}{\Sigma\hat{\hat{e}}_i^2/[n-(m+1)]}$
Πρόσθετη διακύμανση από την εισαγωγή των $X_{k+1}, X_{k+2}, \ldots, X_m$	$\Sigma\hat{\hat{y}}_i^2 - \Sigma\hat{y}_i^2$	m-k	$\frac{\Sigma\hat{\hat{y}}_i^2 - \Sigma\hat{y}_i^2}{m-k}$	$F^* = \frac{(\Sigma\hat{\hat{y}}_i^2 - \Sigma\hat{y}_i^2)/(m-k)}{\Sigma\hat{\hat{e}}_i^2/[n-(m+1)]}$
Κατάλοιπα μετά την εισαγωγή των $X_{k+1}, X_{k+2}, \ldots, X_m$	$\Sigma\hat{\hat{e}}_i^2$	n-(m+1)	$\frac{\Sigma\hat{\hat{e}}_i^2}{n-(k+1)}$	
Σύνολο διακύμανσης της Y	Σy_i^2	n-1		

με k μεταβλητές

με m μεταβλητές

$$\hat{Y}_i=\hat{\alpha}_0+\hat{\alpha}_1X_{i1}+\cdots+\hat{\alpha}_kX_{ik}+\cdots+\hat{\alpha}_mX_{im} \qquad (3.12.2)$$

και οι νέες συνιστώσες της διακύμανσης της Y είναι $\Sigma\hat{y}_i^2$ και $\Sigma\hat{e}_i^2$.

Ο πίνακας για την ανάλυση της διακύμανσης στις δύο παλινδρομήσεις είναι ο 3.2.

Με τις πληροφορίες που έχουμε συγκεντρώσει στον πίνακα 3.2 μπορούμε να προχωρήσουμε στους εξής ελέγχους:

(i) Έλεγχος της στατιστικής σημαντικότητας της παλινδρόμησης (3.12.2).

Ο σχετικός λόγος F^* που θα χρησιμοποιήσουμε είναι:

$$F^*=\frac{\Sigma\hat{y}_i^2/m}{\Sigma\hat{e}_i^2/[n-(m+1)]} \qquad (3.12.3)$$

$$=\frac{R^2/m}{(1-R^2)/[n-(m+1)]} \qquad (3.12.4)$$

ο οποίος θα συγκριθεί με τη θεωρητική τιμή $F_{0.05}$ των πινάκων με m και n-(m+1) βαθμούς ελευθερίας (επίπεδο σημαντικότητας 5%).

(ii) Έλεγχος της βελτίωσης της παλινδρόμησης (3.12.1) από την εισαγωγή των νέων ερμηνευτικών μεταβλητών $X_{k+1}, X_{k+2}, \ldots, X_m$.

Έλεγχος για την βελτίωση της παλινδρόμησης από την εισαγωγή των πρόσθετων μεταβλητών

Ο σχετικός λόγος F^* είναι ο εξής:

$$F^*=\frac{(\Sigma\hat{y}_i^2-\Sigma\hat{y}_i^2)/(m-k)}{\Sigma\hat{e}_i^2/[n-(m+1)]} \qquad (3.12.5)$$

ο οποίος θα συγκριθεί με τη θεωρητική τιμή $F_{0.05}$ των πινάκων με (m-k) και n-(m+1) βαθμούς ελευθερίας.

CHOW TEST

3.13. ΕΛΕΓΧΟΣ ΤΗΣ ΙΣΟΤΗΤΑΣ ΤΩΝ ΣΥΝΤΕΛΕΣΤΩΝ ΜΙΑΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ ΟΤΑΝ ΑΥΤΗ ΕΚΤΙΜΗΘΕΙ ΜΕ ΔΙΑΦΟΡΕΤΙΚΑ ΔΕΙΓΜΑΤΑ

Υποθέτουμε ότι διαθέτουμε δύο δείγματα μεγέθους n_1 και n_2 αντίστοιχα για τις μεταβλητές Y, $X_1, X_2, \ldots, X_k$ και ότι τα

χρησιμοποιούμε χωριστά για την εκτίμηση της παλινδρόμησης της Y πάνω στις $X_1, X_2, \ldots, X_k$. Έχουμε έτσι δύο εκτιμήσεις της ίδιας παλινδρόμησης για δύο διαφορετικές χρονικές περιόδους ή για δύο διαφορετικά δείγματα διαστρωματικών στοιχείων.

Επιθυμούμε να ελέγξουμε αν οι δυο παλινδρομήσεις παρουσιάζουν στατιστικά σημαντική διαφορά, δηλαδή αν η σχέση μεταξύ της Y και των $X_1, X_2, \ldots, X_k$ έχει μεταβληθεί από τη μια χρονική περίοδο στην άλλη ή από το ένα διαστρωματικό δείγμα στο άλλο.

Για τη διεξαγωγή του ελέγχου αυτού ακολουθούμε τα εξής βήματα:

(i) Χρησιμοποιώντας και τα δυο δείγματα κατασκευάζουμε ένα νέο δείγμα μεγέθους n_1+n_2, εκτιμούμε με αυτό την παλινδρόμηση

$$\hat{Y}_i=\hat{\alpha}_0+\hat{\alpha}_1X_{i1}+\cdots+\hat{\alpha}_kX_{ik} \qquad (3.13.1)$$

και υπολογίζουμε το άθροισμα των τετραγώνων των καταλοίπων:

$$\Sigma\hat{e}_i^2=\Sigma(Y_i-\hat{Y}_i)^2 \qquad (3.13.2)$$

το οποίο έχει $(n_1+n_2)-(k+1)$ βαθμούς ελευθερίας.

(ii) Εκτιμούμε την παλινδρόμηση (3.13.1) χρησιμοποιώντας χωριστά τα δυο δείγματα και υπολογίζουμε το άθροισμα των τετραγώνων των καταλοίπων των δυο παλινδρομήσεων:

$$\hat{Y}_{1i}=\hat{b}_0+\hat{b}_1X_{i1}+\cdots+\hat{b}_kX_{ik}, \quad \Sigma\hat{e}_{1i}^2=\Sigma(Y_i-\hat{Y}_{1i})^2 \qquad (3.13.3)$$

με $n_1-(k+1)$ βαθμούς ελευθερίας και

$$\hat{Y}_{2i}=\hat{c}_0+\hat{c}_1X_{i1}+\cdots+\hat{c}_kX_{ik}, \quad \Sigma\hat{e}_{2i}^2=\Sigma(Y_i-\hat{Y}_{2i})^2 \qquad (3.13.4)$$

με $n_2-(k+1)$ βαθμούς ελευθερίας.

(iii) Υπολογίζουμε το άθροισμα:

$$\Sigma\hat{e}_{1i}^2+\Sigma\hat{e}_{2i}^2$$

το οποίο έχει

$$n_1-(k+1)+n_2-(k+1)=(n_1+n_2)-2(k+1)$$

βαθμούς ελευθερίας.

(iv) Υπολογίζουμε τη διαφορά

$$\Sigma\hat{e}_i^2-(\Sigma\hat{e}_{1i}^2+\Sigma\hat{e}_{2i}^2)$$

η οποία έχει

$$[(n_1+n_2)-(k+1)]-[(n_1+n_2)-2(k+1)]=k+1$$

βαθμούς ελευθερίας.

(v) Κατασκευάζουμε το λόγο

Chow test

$$\boxed{F^*=\frac{[\Sigma\hat{e}_i^2-(\Sigma\hat{e}_{1i}^2+\Sigma\hat{e}_{2i}^2)]/(k+1)}{(\Sigma\hat{e}_1^2+\Sigma\hat{e}_2^2)/[(n_1+n_2)-2(k+1)]}} \tag{3.13.5}$$

τον οποίο θα χρησιμοποιήσουμε για να ελέγξουμε την υπόθεση

$$H_0:\ b=a$$

κατά της

$$H_1:\ b\neq a$$

όπου b και a είναι, αντίστοιχα, τα διανύσματα των συντελεστών στις παλινδρομήσεις (3.13.3) και (3.13.4):

Η τιμή του λόγου F^* θα συγκριθεί με τη θεωρητική τιμή της στατιστικής $F_{0.05}$ των πινάκων με $(k+1)$ και $[(n_1+n_2)-2(k+1)]$ βαθμούς ελευθερίας. Αν $F^*>F_{0,05}$ απορρίπτουμε την H_0 και δεχόμαστε ότι υπάρχει στατιστικά σημαντική διαφορά (σε επίπεδο 5%) μεταξύ των παλινδρομήσεων στις δυο διαφορετικές χρονικές περιόδους ή στα δύο διαστρωματικά δείγματα, δηλαδή ότι η γραμμική σχέση μεταξύ της Y και των $X_1,X_2,\ldots,X_k$ έχει μεταβληθεί διαχρονικά ή είναι διαφορετική στα δυο διαστρωματικά δείγματα.

Πρέπει να επισημάνουμε ότι ο έλεγχος που περιγράψαμε δεν είναι σε θέση να προσδιορίσει που οφείλεται η μεταβολή

της γραμμικής σχέσης μεταξύ της Y και των $X_1,X_2,\ldots,X_k$, δηλαδή αν οφείλεται σε παράλληλη μετατόπιση ($b_0\neq c_0$) ή στη μεταβολή των κλίσεων της Y ως προς τις $X_1,X_2,\ldots,X_k$ ($b_i\neq c_i$ για μερικά $i\leq k$). Το θέμα αυτό αντιμετωπίζεται, όπως θα δούμε παρακάτω, με τη βοήθεια των ψευδομεταβλητών ή με την εισαγωγή του χρόνου ως ανεξάρτητης μεταβλητής.

3.14. ΕΛΕΓΧΟΣ ΤΗΣ ΣΤΑΘΕΡΟΤΗΤΑΣ ΤΩΝ ΣΥΝΤΕΛΕΣΤΩΝ ΤΗΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ ΟΤΑΝ ΑΥΞΗΘΕΙ ΤΟ ΜΕΓΕΘΟΣ ΤΟΥ ΔΕΙΓΜΑΤΟΣ

Όταν αυξάνεται το μέγεθος του δείγματος είναι δυνατόν να παρατηρηθούν μεταβολές στη γραμμική σχέση που συνδέει την Y με τις ερμηνευτικές μεταβλητές $X_1,X_2,\ldots,X_k$. Αν τα στοιχεία είναι π.χ. χρονολογικές σειρές, τότε, διάφορα γεγονότα -όπως η μεταβολή του ύψους της φορολογίας, η υποτίμηση του νομίσματος κλπ.- επιφέρουν διαρθωτικές μεταβολές στην οικονομία οι οποίες επηρεάζουν και τη γραμμική σχέση μεταξύ των Y και $X_1,X_2,\ldots,X_k$, μεταβάλοντας είτε τους συντελεστές της είτε τη συναρτησιακή μορφή της. Στην περίπτωση που μεταβάλλονται μόνο οι συντελεστές, τίθεται το ερώτημα αν οι μεταβολές αυτές είναι στατιστικά σημαντικές, οπότε πρέπει να λάβουμε σοβαρά υπόψη τη διαρθρωτική μεταβολή της εξίσωσης.

Αν ο αριθμός των νέων παρατηρήσεων είναι αρκετά μεγαλύτερος από τον αριθμό των προς εκτίμηση παραμέτρων, τότε μπορούμε να χρησιμοποιήσουμε το σύνολο των παρατηρήσεων αυτών ως νέο δείγμα και να διεξαγάγουμε τον έλεγχο που περιγράψαμε στην προηγούμενη παράγραφο.

Αν όμως το σύνολο n_1 των νέων παρατηρήσεων δεν είναι αρκετά μεγάλο, τότε διεξάγουμε τον ακόλουθο έλεγχο:

(i) Εκτιμούμε την παλινδρόμηση χρησιμοποιόντας το σύνολο όλων των παρατηρήσεων -τις n παρατηρήσεις του αρχικού δείγματος και τις n_1 νέες παρατηρήσεις- και υπολογίζουμε το άθροισμα των τετραγώνων των καταλοίπων της παλινδρόμησης:

$$\hat{Y}_{1i}=\hat{a}_0+\hat{a}_1X_{1i}+\cdots+\hat{a}_kX_{ik}$$

$$\Sigma\hat{e}_{1i}^2=\Sigma(Y_{1i}-\hat{Y}_{1i})^2, \quad \text{με} \quad (n+n_1)-(k+1) \text{ βαθμούς ελευθερίας.}$$

(ii) Εκτιμούμε ξανά την παλινδρόμηση χρησιμοποιόντας μόνο το αρχικό δείγμα και υπολογίζουμε και πάλι το άθροισμα των τετραγώνων των καταλοίπων:

$$\hat{Y} = \hat{b}_0 + \hat{b}_1 X_{i1} + \cdots + \hat{b}_k X_{ik}$$

$\Sigma \hat{e}_i^2 = \Sigma (Y_i - \hat{Y}_i)^2$, με $n-(k+1)$ βαθμούς ελευθερίας.

(iii) Υπολογίζουμε τη διαφορά:

$$\Sigma \hat{e}_{i1}^2 - \Sigma \hat{e}_i^2$$

καθώς και τους αντίστοιχους βαθμούς ελευθερίας:

$$[(n+n_1)-(k+1)]-[n-(k+1)]=n_1$$

(iv) Σχηματίζουμε το λόγο:

$$F^* = \frac{(\Sigma \hat{e}_{i1}^2 - \Sigma \hat{e}_i^2)/n_1}{\Sigma \hat{e}_i^2/[n-(k+1)]} \qquad (3.14.1)$$

ο οποίος αποδεικνύεται ότι ακολουθεί την κατανομή F με n_1 και $n-(k+1)$ βαθμούς ελευθερίας.

(v) Ελέγχουμε την υπόθεση

$$H_0: \alpha = b$$

κατά της

$$H_1: \alpha \neq b$$

Αν $F^* > F_{0,05}$ των πινάκων τότε απορρίπτουμε την υπόθεση H_0 και δεχόμαστε ότι έχει επέλθει στατιστικά σημαντική διαρθρωτική αλλαγή στη γραμμική σχέση που συνδέει την Y με τις $X_1, X_2, \ldots, X_k$.

3.15. ΟΙ ΣΥΝΤΕΛΕΣΤΕΣ ΑΠΛΗΣ ΣΥΣΧΕΤΙΣΗΣ, ΜΕΡΙΚΗΣ ΣΥΣΧΕΤΙΣΗΣ ΚΑΙ ΟΙ ΕΚΤΙΜΗΤΡΙΕΣ ΕΛΑΧΙΣΤΩΝ ΤΕΤΡΑΓΩΝΩΝ

Γνωρίζουμε ότι αν η εξαρτημένη μεταβλητή Y ερμηνεύεται γραμμικά από τις ανεξάρτητες μεταβλητές $X_1, X_2, \ldots, X_k$, τότε

με τους συντελεστές "απλής συσχέτισης" (μηδενικής τάξης) r_{yx_i}, $i=1,2,\ldots,k$, μετρούμε την αναλογία της διακύμανσης της Y που ερμηνεύεται μόνο από τη μεταβλητή X_i, $i=1,2,\ldots,k$. Γνωρίζουμε ακόμα ότι, ο συντελεστής πολλαπλού προσδιορισμού $R^2_{y \cdot x_1 x_2 \ldots x_k}$ προσδιορίζει το μέρος της διακύμανσης της Y που ερμηνεύεται από όλες τις μεταβλητές $X_1, X_2, \ldots, X_k$ συγχρόνως.

Αλλά στην πολλαπλή παλινδρόμηση ενδιαφερόμαστε να έχουμε απαντήσεις και σε ερωτήματα όπως τα ακόλουθα:

(i) Τι μέρος της διακύμανσης της Y ερμηνεύεται από τη μεταβλητή X_j μετά την εισαγωγή της μεταβλητής X_i, $i,j=1,2,\ldots,k$, $i \neq j$;

(ii) Τι μέρος της διακύμανσης της Y ερμηνεύεται από τη μεταβλητή X_j μετά την εισαγωγή των μεταβλητών X_i και X_ℓ, $i,j,\ell=1,2,\ldots,k$, $i \neq \ell \neq j$;

(iii) Τι μέρος της διακύμανσης της Y ερμηνεύεται από τη μεταβλητή X_j μετά την εισαγωγή λ από τις υπόλοιπες (k-1) μεταβλητές;

Την απάντηση στο πρώτο ερώτημα δίνει, όπως είναι γνωστό, ο συντελεστής "μερικής συσχέτισης πρώτης τάξης" $r_{yx_j \cdot x_i}$ που μετρά την συσχέτιση των Y και X_j αφού έχουμε απαλείψει την επίδραση της X_i πάνω στις δύο αυτές μεταβλητές. Οι αντίστοιχες απαντήσεις στα ερωτήματα (ii) και (iii) δίνονται από τους συντελεστές "μερικής συσχέτισης δεύτερης τάξης" $r_{yx_j \cdot x_i x_l}$ και "μερικής συσχέτισης ℓ-τάξης" $r_{yx_j \cdot x_1 x_2 \ldots x_\ell}$.

Οι συντελεστές μερικής συσχέτισης παίζουν σημαντικό ρόλο στο αν θα εισαγάγουμε ή όχι νέες ερμηνευτικές μεταβλητές σε ένα υπόδειγμα. Αν π.χ. ο συντελεστής r_{yx_2} είναι πολύ υψηλός ενώ ο συντελεστής $r_{yx_2 \cdot x_1}$ είναι πολύ χαμηλός, αυτό σημαίνει ότι, ενώ μόνη της η X_2 ερμηνεύει πολύ καλά την Y, αν εισαχθεί και η X_1 τότε η X_2 δε βοηθά πλέον σημαντικά στην ερμηνεία της Y. Στην περίπτωση αυτή το ότι η X_2 ερμηνεύει μόνη της ικανοποιητικά την Y οφείλεται όχι στο ότι η X_2 επιδρά στη διαμόρφωση της Y αλλά στο ότι και η Y και η X_2 διαμορφώνονται από την X_1 και συνεπώς δεν υπάρχει λόγος να διατηρήσουμε την X_2 στην παλινδρόμηση.

Ας υποθέσουμε τώρα ότι διαθέτουμε δυο ερμηνευτικές μεταβλητές X_1 και X_2 και πρέπει -εμπειρικά- να αποφασίσουμε αν θα χρησιμοποιήσουμε μια από αυτές ή και τις δυο για την ερμηνεία της Y. Αρχικά εξετάζουμε τους συντελεστές απλής συσχέτισης r_{yx_1} και r_{yx_2} και εισάγουμε πρώτα τη μεταβλητή που παρουσιάζει τη μεγαλύτερη συσχέτιση με την Y. Αν αυτή είναι η X_1 τότε υπολογίζουμε στη συνέχεια το συντελεστή μερικής συσχέτισης $r_{yx_2 \cdot x_1}$ και εισάγουμε και την X_2 μόνο αν ο συντελεστής αυτός είναι υψηλός, ενώ, αν είναι η X_2, εισάγουμε και την X_1 μόνο αν ο $r_{yx_1 \cdot x_2}$ είναι υψηλός. Μερικές φορές μπορούμε να χρησιμοποιήσουμε ένα συνδιασμό των μεταβλητών X_1 και X_2 αντί των δυο μεταβλητών χωριστά. Αν π.χ. οι μεταβλητές X_1 και X_2 αντιπροσωπεύουν, αντίστοιχα, την ειδικευμένη και την ανειδίκευτη εργασία που απαιτείται στη διαδικασία της παραγωγής Y ενός προϊόντος και οι συντελεστές r_{yx_1} και r_{yx_2} είναι και οι δυο υψηλοί ενώ οι συντελεστές μερικής συσχέτισης είναι και οι δυο πολύ χαμηλοί, αυτό σημαίνει ότι ο χωρισμός της εργασίας σε ειδικευμένη και ανειδίκευτη δε βοηθά στην ερμηνεία της παραγωγής Y. Στην περίπτωση αυτή μπορούμε να χρησιμοποιήσουμε ως ερμηνευτική μεταβλητή το άθροισμα X_1+X_2 ή κάποιο άλλο συνολικό μέτρο της απασχόλησης.

Υπάρχουν βέβαια και περιπτώσεις που η σειρά της εισαγωγής των μεταβλητών καθορίζεται εύκολα. Π.χ. αν θέλουμε να ερμηνεύσουμε την Y με ένα πολυώνυμο της X:

$$Y_i=\alpha_0+\alpha_1 X_i+\alpha_2 X_i^2+\cdots+e_i$$

τότε εισάγουμε πρώτα την X_i και στη συνέχεια τις X_i^2, X_i^3 κλπ. και σταματούμε όταν η μεταβολή του συντελεστή $\bar{R}^2$ είναι αμελητέα ή αρνητική.

Θα εξετάσουμε τώρα τις σχέσεις που συνδέουν τους συντελεστές απλής και μερικής συσχέτισης με τους συντελεστές του γενικού γραμμικού υποδείγματος:

$$Y_i=b_0+b_1 X_{i1}+b_2 X_{i2}+\cdots+b_k X_{ik}+e_i. \qquad (3.15.1)$$

Ας συμβολίσουμε με

$$R=\begin{bmatrix} r_{yy} & r_{y1} & r_{y2} & \cdots & r_{yk} \\ r_{1y} & r_{11} & r_{12} & \cdots & r_{1k} \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ \cdot & \cdot & \cdot & & \cdot \\ r_{ky} & r_{k1} & r_{k2} & \cdots & r_{kk} \end{bmatrix}$$

τον πίνακα των συντελεστών απλής συσχέτισης (μηδενικής τάξης) μεταξύ των μεταβλητών $Y, X_1, X_2, \ldots, X_k$, όπου βέβαια $r_{ii}=1$ $i=y,1,2,\ldots,k$, και ο πίνακας R είναι συμμετρικός.

Θα δείξουμε ότι οι εκτιμήτριες ελαχίστων τετραγώνων του γενικού γραμμικού υποδείγματος (3.15.1), καθώς και οι συντελεστές μερικής συσχέτισης, μπορούν να εκφράστούν με τη βοήθεια των "συμπαραγόντων" (co-factors) του πίνακα R.

Εκφράζοντας τις μεταβλητές $Y, X_1, X_2, \ldots, X_k$ σε αποκλίσεις από τους μέσους τους, η εκτίμηση ελαχίστων τετραγώνων της (3.15.1) γράφεται:

$$\hat{y}_i=\hat{b}_1 x_{i1}+\hat{b}_2 x_{i2}+\cdots+\hat{b}_k x_{ik} \qquad (3.15.2)$$

στην οποία οι συντελεστές $\hat{b}_i$, $i=1,2,\ldots,k$, έχουν προκύψει από την ελαχιστοποίηση του αθροίσματος των τετραγώνων:

$$\sum_{i=1}^{n}(y_i-\hat{y}_i)^2. \qquad (3.15.3)$$

Η ελαχιστοποίηση του αθροίσματος των τετραγώνων (3.15.3) οδηγεί στις κανονικές εξισώσεις (παραλείπουμε το δείκτη i για να διευκολύνουμε τους συμβολισμούς):

$$\begin{aligned} &\hat{b}_1\Sigma x_1^2+\hat{b}_2\Sigma x_1 x_2+\cdots+\hat{b}_k\Sigma x_1 x_k=\Sigma x_1 y \\ &\hat{b}_1\Sigma x_2 x_1+\hat{b}_2\Sigma x_2^2+\cdots+\hat{b}_k\Sigma x_2 x_k=\Sigma x_2 y \qquad (3.15.4) \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ &\hat{b}_1\Sigma x_k x_1+\hat{b}_2\Sigma x_k x_2+\cdots+\hat{b}_k\Sigma x_k^2=\Sigma x_k y. \end{aligned}$$

Οι κανονικές εξισώσεις μπορούν να γραφούν με τη μορφή[1]

1. Π.χ. η πρώτη εξίσωση γράφεται:

$$\begin{aligned}
&\hat{b}_1 r_{11} s_1+\hat{b}_2 r_{12} s_2+\dots+\hat{b}_k r_{1k} s_k=r_{1y} s_y\\
&\hat{b}_1 r_{21} s_1+\hat{b}_2 r_{22} s_2+\dots+\hat{b}_k r_{2k} s_k=r_{2y} s_y \qquad (3.15.5)\\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\\
&\hat{b}_1 r_{k1} s_1+\hat{b}_2 r_{k2} s_2+\dots+\hat{b}_k r_{kk} s_k=r_{ky} s_y
\end{aligned}$$

όπου $s_i=\sqrt{\frac{\Sigma x_i^2}{n}}$ είναι η δειγματική τυπική απόκλιση της μεταβλητής X_i.

Λύνοντας τώρα το σύστημα (3.15.5) με τη μέθοδο του Gramer, υπολογίζουμε τις εκτιμήτριες $\hat{b}_i$ ελαχίστων τετραγώνων συναρτήσει των r_{ij}, και των s_i Π.χ. για τη $\hat{b}_1$ βρίσκουμε

$$\hat{b}_1=\frac{\begin{vmatrix} r_{1y}s_y & r_{12}s_2 & \dots & r_{1k}s_k\\ r_{2y}s_y & r_{22}s_2 & \dots & r_{2k}s_k\\ \vdots & \vdots & & \vdots\\ r_{ky}s_y & r_{k2}s_2 & \dots & r_{kk}s_k \end{vmatrix}}{\begin{vmatrix} r_{11}s_1 & r_{12}s_2 & \dots & r_{1k}s_k\\ r_{21}s_1 & r_{22}s_2 & \dots & r_{2k}s_k\\ \vdots & \vdots & & \vdots\\ r_{k1}s_1 & r_{k2}s_2 & \dots & r_{kk}s_k \end{vmatrix}}$$

$$=\frac{s_y s_2 s_3 \dots s_k \mathcal{R}_{y1}}{s_1 s_2 s_3 \dots s_k \mathcal{R}_{yy}}$$

$$=\frac{s_y}{s_1}\frac{\mathcal{R}_{y1}}{\mathcal{R}_{yy}}$$

$$\hat{b}_1\frac{\Sigma x_1^2}{\sqrt{\Sigma x_1^2}\sqrt{\Sigma x_1^2}}\frac{\sqrt{\Sigma x_1^2}}{\sqrt{n}}+\hat{b}_2\frac{\Sigma x_1 x_2}{\sqrt{\Sigma x_1^2}\sqrt{\Sigma x_2^2}}\frac{\sqrt{\Sigma x_2^2}}{\sqrt{n}}+\dots+\hat{b}_k\frac{\Sigma x_1 x_k}{\sqrt{\Sigma x_1^2}\sqrt{\Sigma x_k^2}}\frac{\sqrt{\Sigma x_k^2}}{\sqrt{n}}=$$

$$=\frac{\Sigma x_1 y}{\sqrt{\Sigma x_1^2}\sqrt{\Sigma y^2}}\frac{\sqrt{\Sigma y^2}}{\sqrt{n}}$$

και γενικά

$$\hat{b}_i=-\frac{s_y}{s_i}\frac{\mathcal{R}_{yi}}{\mathcal{R}_{yy}},\quad i=1,2,\dots,k, \qquad (3.15.6)$$

όπου $\mathcal{R}_{ij}$, $i,j=y,1,2,\dots,k$, είναι ο "συμπαράγων" (cofactor) του στοιχείου r_{ij}, $i,j=y,1,2,\dots,k$ του πίνακα R, δηλαδή η ορίζουσα του πίνακα που προκύπτει από τον R, αν αφαιρεθεί η i γραμμή και η j στήλη, με πρόσημο $(-1)^{i+j}$.

Αντικαθιστώντας τις εξισώσεις (3.15.6) στην (3.15.2) προκύπτει η εξίσωση:

$$\frac{\mathcal{R}_{yy}}{s_y}\hat{y}+\frac{\mathcal{R}_{y1}}{s_1}x_1+\frac{\mathcal{R}_{y2}}{s_2}x_2+\dots+\frac{\mathcal{R}_{yk}}{s_k}x_k=0. \qquad (3.15.7)$$

Το άθροισμα των τετραγώνων των καταλοίπων της παλινδρόμησης μπορεί να γραφεί ως:

$$\Sigma\hat{e}_i^2=\Sigma(y_i-\hat{y}_i)^2$$

$$=\Sigma\left(y_i+\frac{s_y}{s_1}\frac{\mathcal{R}_{y1}}{\mathcal{R}_{yy}}x_{i1}+\frac{s_y}{s_2}\frac{\mathcal{R}_{y2}}{\mathcal{R}_{yy}}x_{i2}+\dots+\frac{s_y}{s_k}\frac{\mathcal{R}_{yk}}{\mathcal{R}_{yy}}x_{ik}\right)^2$$

και μετά από πράξεις

$$=\frac{ns_y^2|R|}{\mathcal{R}_{yy}}. \qquad (3.15.8)$$

Άρα, ο συντελεστής πολλαπλού προσδιορισμού $R^2_{y\cdot1,2\dots k}$ παίρνει τη μορφή

$$R^2_{y\cdot1,2\dots k}=1-\frac{\Sigma\hat{e}_i^2}{\Sigma y_i^2}$$

$$=1-\frac{|R|}{\mathcal{R}_{yy}}. \qquad (3.15.9)$$

Οι εξισώσεις (3.15.6) και (3.15.9) εκφράζουν τις εκτιμήτριες ελαχίστων τετραγώνων και το συντελεστή πολλαπλού προσδιορισμού συναρτήσει των στοιχείων του πίνακα R, δηλαδή συναρτήσει των συντελεστών απλής συσχέτισης μεταξύ των $Y, X_1, X_2, \dots, X_k$.

Οι αντίστοιχες εκφράσεις για τους συντελεστές μερικής συσχέτισης πρώτης τάξης δίνονται από τη σχέση[1]:

$$r_{yi\cdot 1,2\ldots i-1,i+1,\ldots k} = -\frac{\mathcal{R}_{yi}}{\sqrt{\mathcal{R}_{yy}\mathcal{R}_{ii}}}.$$

3.16. ΟΙ ΣΥΝΤΕΛΕΣΤΕΣ «ΒΗΤΑ» ΚΑΙ ΕΛΑΣΤΙΚΟΤΗΤΕΣ

Οι συντελεστές "βήτα" χρησιμοποιούνται στην περίπτωση που επιθυμούμε να συγκρίνουμε τη σχετική σπουδαιότητα των ανεξάρτητων μεταβλητών σε μια παλινδρόμηση ανεξάρτητα από τις μονάδες μέτρησης που χρησιμοποιούμε για τη μέτρηση των μεταβλητών της. Για τον προσδιορισμό των συντελεστών βήτα εκτιμούμε, αντί της αρχικής παλινδρόμησης, την παλινδρόμηση στην οποία οι τιμές της εξαρτημένης μεταβλητής Y και των ανεξάρτητων μεταβλητών $X_1, X_2, \ldots, X_k$, έχουν "τυποποιηθεί" δηλαδή έχει αφαιρεθεί από αυτές ο μέσος τους και έχουν διαιρεθεί με την τυπική τους απόκλιση.

Από την παλινδρόμηση

$$Y_i = b_0 + b_1X_{i1} + b_2X_{i2} + \cdots + b_kX_{ik} + e_i. \qquad (3.16.1)$$

προκύπτει ότι

$$\bar{Y} = b_0 + b_1\bar{X}_1 + b_2\bar{X}_2 + \cdots + b_k\bar{X}_k. \qquad (3.16.2)$$

Αν από την (3.16.1) αφαιρέσουμε την (3.16.2) και, στη συνέχεια, διαιρέσουμε και τα δύο μέλη με την τυπική απόκλιση s_y της Y, εύκολα προκύπτει το "τυποποιημένο" γραμμικό υπόδειγμα

$$\frac{Y_i-\bar{Y}}{s_y} = b_1\frac{s_1}{s_y}\cdot\frac{X_{i1}-\bar{X}_1}{s_1} + \cdots + b_k\frac{s_k}{s_y}\,\frac{X_{ik}-\bar{X}_k}{s_k} + \frac{e_i}{s_y} \qquad (3.16.3)$$

1. Βλέπε σχετικά J. Johnston, *Econometric methods*, Second Edition, Mc Graw-Hill, New York 1972, pp. 134-135. Βλέπε επίσης Κ. Δρακάτου, "Μαθήματα Οικονομετρίας" Μέρος Πρώτο: Μέθοδοι, Εκδόσεις Παπαζήση, Αθήνα 1971, σελ. 20-31 για την αναλυτική διατύπωση των πιο πάνω σχέσεων στην περίπτωση της πολλαπλής παλινδρόμησης με δύο ερμηνευτικές μεταβλητές.

ή

$$\frac{Y_i-\bar{Y}}{s_y} = b_1^*\frac{X_{i1}-\bar{X}_1}{s_1} + b_2^*\frac{X_{i2}-\bar{X}_2}{s_2} + \cdots + b_k^*\frac{X_{ik}-\bar{X}_k}{s_k} + e_i^*. \qquad (3.16.4)$$

Από τις δυο αυτές εξισώσεις γίνεται φανερό ότι οι εκτιμήσεις των συντελεστών βήτα συνδέονται με τις εκτιμήσεις των συντελεστών της αρχικής παλινδρόμησης με τη σχέση:

$$\hat{b}_j^* = \hat{b}_j\frac{s_j}{s_y}, \quad j=1,2,\ldots,k. \qquad (3.16.5)$$

Στην παλινδρόμηση (3.16.4) όλες οι μεταβλητές είναι εκφρασμένες σε μονάδες τυπικής απόκλισης και συνεπώς οι τιμές τους είναι απαλλαγμένες από τις μονάδες μέτρησης των μεταβλητών αυτών ενώ αυτό δε συμβαίνει στην αρχική παλινδρόμηση όπου οι μεταβλητές είναι εκφρασμένες σε διαφορετικές μονάδες μέτρησης. Έτσι αν π.χ. $\hat{b}_j^*=0.6$ αυτό σημαίνει ότι αν η X_j μεταβληθεί κατά μια τυπική απόκλιση τότε η Y θα μεταβληθεί κατά 0.6 τυπικές αποκλίσεις.

Πρέπει να επισημάνουμε το ότι ο συντελεστής βήτα για το σταθερό όρο δεν ορίζεται διότι ο σταθερός όρος απαλείφεται κατά τη διαδικασία της τυποποίησης[1].

Πολύ συχνά στην ποσοτική οικονομική ανάλυση ενδιαφερόμαστε επίσης να προσδιορίσουμε την επίδραση πάνω στην εξαρτημένη μεταβλητή από μια "ποσοστιαία" μεταβολή κάποιας ανεξάρτητης μεταβλητής. Όπως γνωρίζουμε η επίδραση αυτή εκφράζεται από την "ελαστικότητα" της Y ως προς την αντίστοιχη ανεξάρτητη μεταβολή. Π.χ. η ελαστικότητα της Y ως προς τη X_j ορίζεται ως ο λόγος της ποσοστιαίας μεταβολής της Y δια της ποσοστιαίας μεταβολής της X_j. Γενικά, οι ελαστικότητες δεν παραμένουν σταθερές αλλά μεταβάλλονται όταν μετρούνται σε διαφορετικά σημεία κατά μήκος της παλινδρόμησης. Οι ελαστι- ΕΛΑΣΤΙΚΟΤΗΤΕΣ

1. Οι συντελεστές βήτα ή αλλιώς "οι συντελεστές του τυποποιημένου γραμμικού υποδείγματος" συζητούνται με περισσότερες λεπτομέρειες στο βιβλίο του D. Aigner, *Basic Econometrics*, Prentice-Hall, Englewood Cliffs, 1971, pp. 72-80.

κότητες που δίνονται από τον υπολογιστή υπολογίζονται συνήθως στους μέσους των Y και X_j. Π.χ. η εκτίμηση της ελαστικότητας της Y ως προς τη X_j "γύρω από τους μέσους $\bar{Y}$ και $\bar{X}_j$" είναι:

$$\boxed{\begin{aligned}\hat{E}_j&=\frac{\partial\hat{Y}}{\partial X_j}\cdot\frac{\bar{X}_j}{\bar{Y}}\\&=\hat{b}_j\frac{\bar{X}_j}{\bar{Y}}.\end{aligned}} \qquad (3.16.6)$$

Οι ελαστικότητες μπορούν να πάρουν οποιαδήποτε θετική ή αρνητική τιμή και είναι απαλλαγμένες από τις μονάδες μέτρησης των μεταβλητών. Π.χ. αν $\hat{E}_j=3$ αυτό σημαίνει ότι, γύρω από τους μέσους $\bar{Y}$ και $\bar{X}_j$, αν η $\bar{X}_j$ αυξηθεί (μειωθεί) κατά 1% τότε η Y θα αυξηθεί (μειωθεί) κατά 3%. Αντίθετα αν $\hat{E}_j=-2$ αυτό σημαίνει ότι αν η X_j αυξηθεί (μειωθεί) κατά 1% τότε η Y θα μειωθεί (αυξηθεί) κατά 2%. Γενικά μεγάλες απόλυτες τιμές της ελαστικότητας σημαίνουν ότι η εξαρτημένη μεταβλητή αντιδρά έντονα στις μεταβολές της αντίστοιχης ανεξάρτητης μεταβλητής.

3.17. ΟΙ ΕΚΤΙΜΗΤΡΙΕΣ ΜΕΓΙΣΤΗΣ ΠΙΘΑΝΟΦΑΝΕΙΑΣ ΤΩΝ ΠΑΡΑΜΕΤΡΩΝ ΤΟΥ ΚΛΑΣΙΚΟΥ ΓΡΑΜΜΙΚΟΥ ΥΠΟΔΕΙΓΜΑΤΟΣ

Κάτω από το σύνολο των ασθενών υποθέσεων (υ.1) έως (υ.6) και της ισχυρής υπόθεσης (υ.7) και σύμφωνα με τη θεωρία του παραρτήματος Β, η συνάρτηση πιθανοφάνειας y/X είναι η

$$L=(2\pi\sigma^2)^{-\frac{n}{2}}e^{-\frac{1}{2\sigma^2}(y-Xb)'(y-Xb)} \qquad (3.17.1)$$

ή σε λογαριθμική μορφή:

$$Q=\ln L=-\frac{n}{2}\ln 2\pi-\frac{n}{2}\ln\sigma^2-\frac{1}{2\sigma^2}(y-Xb)'(y-Xb). \qquad (3.17.2)$$

Οι συνθήκες πρώτης τάξης για τη μεγιστοποίηση της (3.17.2) είναι:

$$\frac{\partial Q}{\partial b}=\frac{1}{2\sigma^2}2X'(y-X\tilde{b})=0 \qquad (3.17.3)$$

$$\begin{aligned}\frac{\partial Q}{\partial\sigma^2}&=-\frac{n}{2}\frac{1}{\tilde{\sigma}^2}+\frac{1}{2\tilde{\sigma}^4}(y-X\tilde{b})'(y-X\tilde{b})\\&=-\frac{n}{2}\frac{1}{\tilde{\sigma}^2}+\frac{1}{2\tilde{\sigma}^4}\hat{e}'\hat{e} \qquad (3.17.4)\\&=0\end{aligned}$$

και η λύση τους δίνει τις εκτιμήτριες μέγιστης πιθανοφάνειας του κλασικού γραμμικού υποδείγματος:

$$\tilde{b}=(X'X)^{-1}X'y$$

και

$$\tilde{\sigma}^2=\frac{1}{n}\hat{e}'\hat{e}=\frac{\Sigma\hat{e}_i}{n}.$$

Εύκολα διαπιστώνουμε ότι οι εκτιμήτριες $\tilde{b}$ της μέγιστης πιθανοφάνειας ταυτίζονται με τις εκτιμήτριες $\hat{b}$ των ελαχίστων τετραγώνων. Η εκτιμήτρια $\tilde{\sigma}^2$ της διακύμανσης των σφαλμάτων δεν είναι αμερόληπτη εκτιμήτρια της σ^2, είναι όμως συνεπής. Η αμερόληπτη εκτιμήτρια είναι, όπως είδαμε, η

$$\hat{\sigma}^2=\frac{\Sigma\hat{e}_i^2}{n-(k+1)}.$$

3.18. Η ΕΚΤΙΜΗΣΗ ΤΟΥ ΚΛΑΣΙΚΟΥ ΓΡΑΜΜΙΚΟΥ ΥΠΟΔΕΙΓΜΑΤΟΣ ΣΤΗΝ ΠΕΡΙΠΤΩΣΗ ΠΟΥ ΔΙΑΘΕΤΟΥΜΕ ΣΥΜΠΛΗΡΩΜΑΤΙΚΕΣ ΠΛΗΡΟΦΟΡΙΕΣ

Για την εκτίμηση των παραμέτρων του κλασικού γραμμικού υποδείγματος

$$y=Xb+e$$

χρησιμοποιήσαμε, μέχρι τώρα, μόνο τις πληροφορίες του δείγματος για την εξαρτημένη και τις ανεξάρτητες μεταβλητές. Σε πολλές όμως εφαρμογές είναι δυνατό να έχουμε, εκτός από το δείγμα $[y,X]$, και συμπληρωματική πληροφόρηση για τα άγνωστα στοιχεία του διανύσματος b ή και για τη διακύμανση σ^2 των σφαλμάτων και το πρόβλημα που αντιμετωπίζεται από τον ερευνητή είναι ο τρόπος αξιοποίησης των πληροφοριών αυτών για τη βελτίωση των εκτιμήσεων και των συμπερασμάτων της στατιστικής επαγωγής του κλασικού γραμμικού υποδείγματος.

Οι μορφές της συμπληρωματικής πληροφόρησης μπορεί να είναι οι εξής:

α) Πληροφορίες για μεμονωμένες παραμέτρους ή για γραμμικούς συνδιασμούς παραμέτρων, που είναι γνωστές με βεβαιότητα και μπορούν να εκφραστούν ως ακριβείς γραμμικοί περιορισμοί.

β) Πληροφορίες για μεμονωμένες παραμέτρους ή για γραμμικούς συνδιασμούς παραμέτρων, που δεν είναι ακριβείς αλλά "στοχαστικές" και μπορούν να εκφραστούν ως γραμμικοί στοχαστικοί περιορισμοί.

Εκτός από αυτές, υπάρχουν και άλλες μορφές συμπληρωματικής πληροφόρησης όπως, γραμμικοί ανισωτικοί περιορισμοί για τα στοιχεία του διανύσματος b της μορφής $Rb \geqslant r$, πληροφόρηση για την εκ των προτέρων κατανομή των στοιχείων του b (Prior density functions), κλπ. αλλά η εξέτασή τους ξεφεύγει από τα πλαίσια του εισαγωγικού αυτού εγχειριδίου[1].

α) Ακριβής συμπληρωματική πληροφόρηση.

Εδώ θα εξετάσουμε την αξιοποίηση των πληροφοριών της μορφής (α).

Ακριβή πληροφόρηση για τις παραμέτρους ενός υποδείγματος μπορούμε να έχουμε σε πολλές περιπτώσεις. Η πληροφόρηση αυτή μπορεί να είναι είτε εμπειρική, δηλαδή να προέρχεται από άλλες έρευνες για το ίδιο θέμα, είτε θεωρητική δηλαδή να προέρχεται από την οικονομική θεωρία. Π.χ. κατά την

1. Ενδιαφέρουσα σύνοψη των μεθόδων αξιοποίησης των διαφόρων μορφών πληροφόρησης, καθώς και των ιδιοτήτων των εκτιμητριών που προκύπτουν, θα βρει ο ενδιαφερόμενος αναγνώστης στο βιβλίο των G. Judge, W. Griffitus, R.-C. Hill and T.-C. Lee, *The Theory and Practice of Econometrics*, Wiley, New York 1980, Chapter 3, και ειδικότερη ανάλυση στο βιβλίο των G.Judge and M. Bock, *The Statistical Implications of Pre-test and Stein-Rule Estimators in Econometrics*, North Holland, Amsterdam 1978.

εκτίμηση μιας συνάρτησης παραγωγής είναι δυνατόν να γνωρίζουμε αν η επιχείρηση εργάζεται κάτω από συνθήκες σταθερής κλίμακας απόδοσης. Ακόμη, κατά την εκτίμηση συναρτήσεων ζήτησης, η οικονομική θεωρία επιβάλει τους περιορισμούς της ομογένειας, της αθροιστικότητας κλπ.

Η τέτοιου είδους συμπληρωματική πληροφόρηση εκφράζεται με ακριβείς γραμμικούς περιορισμούς, για μερικά ή όλα τα στοιχεία του διανύσματος b, δηλαδή

$$Rb=r \qquad (3.18.1)$$

όπου R είναι ένας γνωστός (J×K) πίνακας, r είναι ένα γνωστό (J×1) διάνυσμα και J είναι ο αριθμός των σπμπληρωματικών πληροφοριών που διαθέτουμε. Π.χ. αν για τα στοιχεία του διανύσματος b γνωρίζουμε ότι $b_2=0{,}7$, $b_1=b_3$ και $b_1+b_2+\cdots+b_K=1$ τότε η (3.18.1) γράφεται αναλυτικά ως εξής:

$$\begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 1 & 0 & -1 & \dots & 0 \\ & & & & \\ 1 & 1 & 1 & \dots & 1 \end{bmatrix} \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_K \end{bmatrix} = \begin{bmatrix} 0{,}7 \\ 0 \\ \\ 1 \end{bmatrix}$$

όπου J=3.

Επειδή οι πληροφορίες της μορφής (3.18.1) είναι ακριβείς, δηλαδή δεν παρουσιάζουν μεταβολές από δείγμα σε δείγμα, θα πρέπει να θεωρηθούν ως "περιορισμοί" κατά τη διαδικασία εκτίμησης. Άρα, στην προκείμενη περίπτωση, πρέπει να ελαχιστοποιήσουμε το άθροισμα των τετραγώνων

$$f(b)=(y-Xb)'(y-Xb) \qquad (3.18.2)$$

κάτω από τους περιορισμούς:

$$Rb-r=0. \qquad (3.18.3)$$

Η ενδεδειγμένη μέθοδος για την επίλυση του προβλήματος είναι η μέθοδος των πολλαπλασιαστών του Lagrange η οποία δίνει τη λύση[1]:

$$\hat{b}^{*}=\hat{b}+(X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}(r-R\hat{b}) \qquad (3.18.4)$$
$$=\hat{b}+VR'(RVR')^{-1}(r-R\hat{b}) \qquad (3.18.5)$$

όπου $V=(X'X)^{-1}$, που διαφέρει από τη γνωστή εκτιμήτρια $\hat{b}$, των ελαχίστων τετραγώνων χωρίς περιορισμούς, κατά το γραμμικό συνδιασμό

$$VR'(RVR')^{-1}(r-R\hat{b}) \qquad (3.18.6)$$

του διανύσματος $(r-R\hat{b})$.

Το διάνυσμα $\hat{b}^{*}$ έχει μέσο (να δειχτεί ως άσκηση)

$$E(\hat{b}^{*})=b+VR'(RVR')^{-1}(r-Rb) \qquad (3.18.7)$$
$$=b+VR'(RVR')^{-1}d \qquad (3.18.8)$$

και είναι αμερόληπτη εκτιμήτρια του διανύσματος b τότε και μόνο τότε αν $r-Rb=d=0$, δηλαδή αν οι περιορισμοί (3.18.3) είναι ακριβείς. Αντίθετα, αν οι περιορισμοί (3.18.3) δεν είναι σωστοί, $(d\neq 0)$ τότε η εκτιμήτρια $\hat{b}^{*}$ είναι μεροληπτική και η μεροληπτικότητα είναι γραμμική συνάρτηση του σφάλματος $d=r-Rb$ των περιορισμών.

Ο πίνακας διακυμάνσεων συνδιακυμάνσεων του διανύσματος $\hat{b}^{*}$ είναι (να δειχτεί επίσης ως άσκηση):

$$E[\hat{b}^{*}-E(\hat{b}^{*})][\hat{b}^{*}-E(\hat{b}^{*})]'=\sigma^{2}[V-VR'(RVR')^{-1}RV]$$
$$=\sigma^{2}[V-C] \qquad (3.18.9)$$

όπου $C=VR'(RVR')^{-1}RV$ είναι ένας θετικός ημιπεπερασμένος πίνακας. Αυτό, μεταξύ άλλων, σημαίνει ότι τα διαγώνια στοιχεία του πίνακα διακυμάνσεων-συνδιακυμάνσεων του $\hat{b}^{*}$ είναι μικρότερα ή ίσα από τα αντίστοιχα διαγώνια στοιχεία του πίνακα διακυμάνσεων-συνδιακυμάνσεων του διανύσματος $\hat{b}$ των ε-

1. Βλέπε A.S. Goldberger, *Econometric Theory*, Wiley, New York 1964 pp. 257-258 για τους αναλυτικούς υπολογισμούς.

λαχίστων τετραγώνων χωρίς περιορισμούς. Συνεπώς η εκτιμήτρια $\hat{b}^{*}$, εφόσον είναι αμερόληπτη, είναι η άριστη μέσα στην κλάση των αμερόληπτων εκτιμητριών που είναι γραμμικές και ως προς το διάνυσμα y (του δείγματος) και ως προς το διάνυσμα r (της συμπληρωματικής πληροφόρησης).

Αν οι περιορισμοί δεν είναι σωστοί δηλαδή αν

$$E(r-R\hat{b})=d\neq 0 \qquad (3.18.10)$$

τότε η εκτιμήτρια $\hat{b}^{*}$ είναι μεροληπτική και έχει μέσο σφάλμα τετραγώνου (risk matrix):

$$\mathcal{R}(b,\hat{b}^{*})=E[(\hat{b}^{*}-b)(\hat{b}^{*}-b)']$$
$$=\sigma^{2}(V-C)+VR'(RVR')^{-1}dd'(RVR')^{-1}RV \qquad (3.18.11)$$

το οποίο είναι ίσο με τον πίνακα διακυμάνσεων - συνδιακυμάνσεων του $\hat{b}^{*}$, συν τον πίνακα που προκύπτει από τον πολλαπλασιασμό των διανυσμάτων της μεροληπτικότητας.

Είναι λοιπόν φανερό από τις σχέσεις (3.18.9) και (3.18.11) ότι αν οι συμπληρωματικές πληροφορίες είναι σωστές τότε, η χρησιμοποίησή τους σε συνδυασμό με την πληροφόρηση του δείγματος, θα οδηγήσει σε αμερόληπτες εκτιμήτριες που θα είναι πιο αποτελεσματικές από τις εκτιμήτριες ελαχίστων τετραγώνων χωρίς περιορισμούς. Βέβαια, στις εφαρμογές, ποτέ δεν μπορούμε να είμαστε βέβαιοι ότι η συμπληρωματική πληροφόρηση είναι ακριβής και συνεπώς ο κίνδυνος να οδηγηθούμε σε εκτιμήσεις όχι μόνο μεροληπτικές, αλλά και λιγότερο αποτελεσματικές από τις εκτιμήτριες Ε.Τ. χωρίς περιορισμούς, είναι σοβαρός. Η συγκριτική μελέτη της αποτελεσματικότητας των εκτιμητριών $\hat{b}$ και $\hat{b}^{*}$ για διάφορες τιμές του σφάλματος d της συμπληρωματικής πληροφόρησης απασχόλησε πολλούς ερευνητές οι οποίοι χρησιμοποίησαν διάφορα κριτήρια για τη σύγκρισή τους. Π.χ. οι Judge and Bock[1] χρησιμοποίησαν το κριτήριο

1. G.G. Judge and M.E. Bock, *The Statistical Implications of Pre-test and Stein-Rule Estimators in Econometrics*, ο.π., σελ. 29.

$$E(\hat{b}-b)(\hat{b}-b)'-E(\hat{b}^*-b)(\hat{b}^*-b)' \qquad (3.18.12)$$

και έδειξαν ότι οι εκτιμήτριες $\hat{b}^*$ θα είναι πιο αποτελεσματικές από τις $\hat{b}$ αν

$$\frac{d'(RVR')^{-1}d}{\sigma^2} \leqslant 1. \qquad (3.18.13)$$

β) Στοχαστική συμπληρωματική πληροφόρηση

Στις οικονομετρικές εφαρμογές οι συμπληρωματικές πληροφορίες κατά κανόνα δεν είναι ακριβείς όπως στην παράγραφο 3.18, αλλά στοχαστικές. Οι συμπληρωματικές πληροφορίες είναι, συνήθως, εκτιμήσεις μερικών στοιχείων του διανύσματος b, καθώς και των διακυμάνσεων-συνδιακυμάνσεών τους, που έχουν προκύψει από άλλες οικονομετρικές έρευνες ή, ακόμη, πληροφορίες ότι ορισμένα στοιχεία του διανύσματος b βρίσκονται ανάμεσα σε ορισμένα όρια[1].

Τέτοιες συμπληρωματικές πληροφορίες εκφράζονται με στοχαστικούς γραμμικούς περιορισμούς της μορφής:

$$r=Rb+v \qquad (3.18.14)$$

1. Π.χ. αν διαθέτουμε την πληροφόρηση ότι για τα στοιχεία b_1 και b_2 του διανύσματος b ισχύει, με πιθανότητα 95%, ότι $0<b_1<1$ και $\frac{1}{4}<b_2<\frac{3}{4}$, τότε, αν χρησιμοποιήσουμε τον κανόνα "του διπλάσιου της διακύμανσης", τα διαστήματα για τις τιμές των b_1 και b_2 είναι $\frac{1}{2}\pm 2\sqrt{\frac{1}{16}}$ και $\frac{1}{2}\pm 2\sqrt{\frac{1}{64}}$ αντίστοιχα και συνεπώς μπορούμε να γράψουμε

$$\frac{1}{2}=b_1+v_1, \quad E(v_1)=0, \quad E(v_1^2)=\frac{1}{16}$$

$$\frac{1}{2}=b_2+v_2, \quad E(v_2)=0, \quad E(v_2^2)=\frac{1}{64}$$

Άρα

$$r=\begin{bmatrix}\frac{1}{2}\\ \frac{1}{2}\end{bmatrix}, \quad R=\begin{bmatrix}1 & 0 & 0 & \dots & 0\\ 0 & 1 & 0 & \dots & 0\end{bmatrix}, \quad v=\begin{bmatrix}v_1\\ v_2\end{bmatrix}, \quad W=\begin{bmatrix}\frac{1}{16} & 0\\ 0 & \frac{1}{64}\end{bmatrix}$$

όπου και πάλι R είναι ένας ($J\times K$) γνωστός πίνακας, r είναι ένα ($J\times 1$) γνωστό διάνυσμα και v είναι ένα ($J\times 1$) μη παρατηρήσιμο διάνυσμα που ακολουθεί την κανονική κατανομή με μέσο d και πίνακα διακυμάνσεων-συνδιακυμάνσεων (χωρίς βλάβη της γενικότητας) ίσο με $\sigma^2 W$ (W, γνωστός πίνακας).

Δοθέντων των πληροφοριών του τύπου αυτού οι Theil-Goldberger[1] υπέθεσαν ότι τα σφάλματα e του κλασικού γραμμικού υποδείγματος $y=Xb+e$ είναι ανεξάρτητα από το διάνυσμα v της (3.19.1) και έγραψαν το στατιστικό υπόδειγμα που περιέχει την πληροφόρηση του δείγματος [y,X] και τη συμπληρωματική πληροφόρηση (3.18.14) ως εξής:

$$\begin{bmatrix}y\\ r\end{bmatrix}=\begin{bmatrix}X\\ R\end{bmatrix}b+\begin{bmatrix}e\\ v\end{bmatrix} \qquad (3.18.15)$$

όπου το διάνυσμα [e',v']' ακολουθεί την πολυκανονική κατανομή με

$$E\begin{bmatrix}e\\ v\end{bmatrix}=\begin{bmatrix}0\\ d\end{bmatrix} \quad \text{και} \quad E\begin{bmatrix}e\\ v\end{bmatrix}\begin{bmatrix}e\\ v\end{bmatrix}'=\sigma^2\begin{bmatrix}I & 0\\ 0 & W\end{bmatrix}=\sigma^2 W_*. \qquad (3.18.16)$$

Αν ο πίνακας διακυμάνσεων-συνδιακυμάνσεων (3.18.16) είναι γνωστός τότε η AGLS[2] εκτιμήτρια του b είναι η

$$\tilde{b}=\left[[X'R']W_*^{-1}\begin{bmatrix}X\\ R\end{bmatrix}\right]^{-1}[X'R']W_*^{-1}\begin{bmatrix}y\\ r\end{bmatrix}$$

$$=(\sigma^{-2}X'X+\sigma^{-2}R'W^{-1}R)^{-1}(\sigma^{-2}X'y+\sigma^{-2}R'W^{-1}r) \qquad (3.18.17)$$

με μέσο

$$E(\tilde{b})=b+[X'X+R'W^{-1}R]^{-1}R'W^{-1}d, \qquad (3.18.18)$$

πίνακα διακυμάνσεων-συνδιακυμάνσεων

1. H. Theil and A.S. Goldbarger, *On Pure and Mixed Statistical Estimation in Economics*, *International Economic Review*, Vol. 2 (1961), pp. 65-78.

2. Για τις εκτιμήτριες AGLS βλέπε σχετικά Κεφ. 4, παραγρ.

$$C(\bar{b})=\sigma^2[X'X+R'W^{-1}R]^{-1} \tag{3.18.19}$$

$$=\sigma^2 D^{-1} \tag{3.18.20}$$

και μέσο σφάλμα τετραγώνου

$$E[(\bar{b}-b)(\bar{b}-b)']=\sigma^2 D^{-1}+D^{-1}R'W^{-1}dd'W^{-1}RD^{-1}$$

$$=C(\bar{b})+(\text{πίνακας μεροληπτικότητας})^2.$$

Αν οι στοχαστικοί περιορισμοί (3.18.14) είναι αμερόληπτοι, δηλαδή,

$$E=[r-Rb]=E(v)=d=0 \tag{3.18.21}$$

τότε, όπως προκύπτει από την (3.18.18), η εκτιμήτρια $\bar{b}$ είναι αμερόληπτη και άριστη μέσα στην κλάση των γραμμικών και αμερόληπτων εκτιμητριών που αξιοποιούν, συγχρόνως την πληροφόρηση του δείγματος, και τη συμπληρωματική στοχαστική πληροφόρηση (3.18.14).

Κεφάλαιο **4**

Έλεγχοι των στατιστικών υποθέσεων του κλασικού γραμμικού υποδείγματος

4.1. ΕΙΣΑΓΩΓΗ

Στο προηγούμενο κεφάλαιο είδαμε ότι, κάτω από τις υποθέσεις (υ.1) έως (υ.7), μπορούμε να εφαρμόσουμε τη μέθοδο ελαχίστων τετραγώνων στο κλασικό γραμμικό υπόδειγμα, να αποδείξουμε ότι οι εκτιμήτριες ελαχίστων τετραγώνων είναι οι άριστες γραμμικές αμερόληπτες εκτιμήτριες, να προσδιορίσουμε διαστήματα εμπιστοσύνης για τις αληθείς τιμές των παραμέτρων στους πληθυσμούς, να ελέγξουμε διάφορες υποθέσεις σχετικά με τις αληθείς τιμές των παραμέτρων στους πληθυσμούς και να διατυπώσουμε προβλέψεις.

Η ορθότητα συνεπώς των αποτελεσμάτων και η ισχύς των συμπερασμάτων της πολλαπλής παλινδρόμησης στην εκτίμηση ποσοτικών οικονομικών σχέσεων, εξαρτάται απόλυτα από την ισχύ των υποθέσεων (υ.1) έως (υ.7). Πρέπει λοιπόν να είμαστε σε θέση:

i) να ελέγξουμε αν οι υποθέσεις αυτές ισχύουν πράγματι σε κάθε συγκεκριμένη εφαρμογή του γραμμικού υποδείγματος.

ii) να γνωρίζουμε ποιες είναι οι επιπτώσεις στις ιδιότητες των εκτιμητριών ελαχίστων τετραγώνων αν δεν ισχύουν μια ή περισσότερες από τις υποθέσεις αυτές, και

iii) να εισαγάγουμε τον κατάλληλο μετασχηματισμό του γραμμικού υποδείγματος ή την κατάλληλη μέθοδο για την αντιμετώπιση των προβλημάτων εκτίμησης στην περίπτωση που δεν ισχύει κάποια από τις υποθέσεις αυτές.

Πριν προχωρήσουμε στην αναλυτική εξέταση των υποθέσεων (υ.1) έως (υ.7) ας δούμε σε ποια από τα στάδια της στατιστικής επαγωγής του προηγούμενου κεφαλαίου είναι απαραίτητη κάθε μια από τις υποθέσεις αυτές:

1. Υπολογισμός των εκτιμητριών ελαχίστων τετραγώνων $\hat{b}=(X'X)^{-1}X'y$ 2. Υπολογισμός των συντελεστών R^2 και $\bar{R}^2$	(υ.1), (υ.4)
3. Απόδειξη της αμεροληψίας των εκτιμητριών ελαχίστων τετραγώνων	(υ.1),(υ.2), (υ.4),(υ.5) και (υ.6)
4. Υπολογισμός του πίνακα διακυμάνσεων-συνδιακυμάνσεων των εκτιμητριών ελαχίστων τετραγώνων. 5. Απόδειξη του θεωρήματος των Gauss-Markov 6. Υπολογισμός της εκτιμήτριας $\hat{\sigma}^2$ της διακύμανσης των σφαλμάτων και του πίνακα συνδιακυμάνσεων $\hat{\sigma}^2(X'X)^{-1}$	(υ.1) έως (υ.6)
7. Απόδειξη της συνέπειας των εκτιμητριών ελαχίστων τετραγώνων	(υ.1) έως (υ.6) και (υ.8)
8. Προσδιορισμός διαστημάτων εμπιστοσύνης, έλεγχοι υποθέσεων, διατύπωση προβλέψεων (σε μικρά δείγματα)	(υ.1) έως (υ.7).

Δυστυχώς, στην εκτίμηση ποσοτικών οικονομικών σχέσεων, ποτέ δε θα είμαστε σε θέση να έχουμε "άμεση" επαλήθευση των πιο πάνω υποθέσεων. Για να ελέγξουμε άμεσα την ισχύ π.χ. της υπόθεσης (υ.3α), δηλαδή της σταθερότητας της σ^2, έπρεπε να διαθέτουμε πληροφορίες για όλους τους πληθυσμούς των $Y, X_1, X_2, \ldots, X_k$. Όμως, από τη φύση των οικονομικών προβλημάτων, δεν είμαστε σε θέση να έχουμε στη διάθεσή μας όλους τους πληθυσμούς των $Y, X_1, X_2, \ldots, X_k$, αλλά, και αν μπορούσαμε να το πράξουμε, η διαπίστωση της ισχύος ή όχι των υποθέσεων θα ήταν άχρηστη γιατί, αν διαθέταμε όλες τις πληροφορίες για τους πληθυσμούς, δε θα χρειαζόταν να προχωρήσουμε σε στατιστική επαγωγή.

Στην πράξη λοιπόν θα περιοριστούμε σε "έμμεσους ελέγχους" για την ισχύ των υποθέσεων αυτών όχι στους πληθυσμούς αλλά στο συγκεκριμένο δείγμα που θα χρησιμοποιήσουμε για την εκτίμηση μιας οικονομικής σχέσης.

4.2. (υ.1): Η ΥΠΟΘΕΣΗ ΤΗΣ ΓΡΑΜΜΙΚΟΤΗΤΑΣ

Όταν μας δοθεί ένα συγκεκριμένο δείγμα παρατηρήσεων για τις μεταβλητές $Y, X_1, \ldots, X_k$, μπορούμε πάντα να προσαρμόσουμε στις παρατηρήσεις του δείγματος το πολυεπίπεδο ελαχίστων τετραγώνων $\hat{y}=X\hat{b}$ και να υπολογίσουμε το συντελεστή R^2 (ή $\bar{R}^2$) έστω και αν η πραγματική σχέση που συνδέει τις μεταβλητές $Y, X_1, X_2, \ldots, X_k$ είναι μη γραμμική ή δεν είναι αυτή ακριβώς η γραμμική σχέση που έχουμε αποφασίσει να εκτιμήσουμε. Και για την εκτίμηση του πολυεπιπέδου ελαχίστων τετραγώνων και για τον υπολογισμό του R^2 δεν απαιτείται καμιά από τις υποθέσεις (υ.2) έως (υ.7) εκτός από την (υ.4) που είναι απαραίτητη για την επίλυση του συστήματος των κανονικών εξισώσεων και την (υ.6) της οποίας τη σημασία θα εξετάσουμε αργότερα.

ΜΑΘΗΜΑ-ΤΙΚΗ — ΤΕΧΝΙΚΗ Σκοπιά

Στη συνέχεια, αν ισχύουν οι ασθενείς υποθέσεις (υ.1) έως (υ.6) μπορούμε να δείξουμε ότι το πολυεπίπεδο ελαχίστων τετραγώνων είναι το άριστο και να υπολογίσουμε τους μέσους και τα τυπικά σφάλματα των εκτιμητριών $\hat{b}_0, \hat{b}_1, \ldots, \hat{b}_k$. Αν επιπλέον ισχύει, σε μικρά δείγματα, και η ισχυρή υπόθεση (υ.7) τότε μπορούμε να προχωρήσουμε στον υπολογισμό των γνωστών λόγων t και F για να εξετάσουμε αν οι επιμέρους εκτιμήτριες είναι στατιστικά σημαντικές ή αν ολόκληρο το πολυεπίπεδο ελαχίστων τετραγώνων είναι στατιστικά σημαντικό.

Γίνεται λοιπόν φανερό ότι το μόνο "εμπειρικό" κριτήριο που μας είναι χρήσιμο για να εξετάσουμε το βαθμό προσαρμογής του πολυεπιπέδου ελαχίστων τετραγώνων "στις παρατηρήσεις του δείγματος" είναι ο συντελεστής πολλαπλού προσδιορισμού R^2 (ή $\bar{R}^2$) αφού τα συμπεράσματα από τους λόγους t και F εξαρτώνται από την ισχύ των υποθέσεων (υ.1) έως (υ.7). Έτσι, αν ο συντελεστής R^2 (ή $\bar{R}^2$) είναι υψηλός[1] και τα πρόσημα των εκτιμητριών $\hat{b}_0, \hat{b}_1, \ldots, \hat{b}_k$ είναι σύμφωνα με τις επιταγές της οικονομικής θεωρίας τότε, σε πρώτη προσέγγιση, μπορούμε να πούμε ότι η "προσαρμογή" είναι καλή και να προχωρήσουμε στην εξέταση της ισχύος των άλλων υποθέσεων και στη συνέχεια στη στατιστική επαγωγή. ΟΙΚΟΝΟΜΙΚΗ ΘΕΩΡΙΑ ↓

Στο σημείο αυτό πρέπει να τονίσουμε ότι ο μοναδικός "a priori" σύμβουλός μας για το είδος της σχέσης που συνδέει τις μεταβλητές Y και $X_1, X_2, \ldots, X_k$ είναι η Οικονομική θεωρία: Η θεωρητική ανάλυση είναι εκείνη που προσδιορίζει βασικά αν το υπόδειγμα θα είναι γραμμικό ή όχι καθώς και το ποιες είναι οι ανεξάρτητες μεταβλητές που επηρεάζουν τη διαμόρφωση της εξαρτημένης μεταβλητής. Όμως, η οικονομική θεωρία προσδιορίζει ακριβείς σχέσεις μεταξύ των οικονομικών μεταβλητών και γνωρίζουμε ότι, όσο προσεκτικά και αν εφαρμόσουμε την οικονομική θεωρία και όσο προσεκτικά και αν συγκεντρώσουμε τα στατιστικά στοιχεία, πάντα θα έχουμε κάποιες αποκλίσεις από την ακριβή σχέση (που τις ονομάσαμε σφάλματα e_i) και οι ιδιότητές τους θα εξαρτώνται απόλυτα από το πόσο προσεκτικά προσδιορίσαμε το υπόδειγμα που θα εκτιμήσουμε και από το πόσο προσεκτικά συγκεντρώσαμε τις παρατηρήσεις του δείγματός μας. ΣΦΑΛΜΑΤΑ ΕΞΕΙΔΙΚΕΥΣΗΣ

Τα κυριότερα σφάλματα εξειδίκευσης που μπορούμε να διαπράξουμε κατά τον προσδιορισμό του υποδείγματος που θα εκτιμήσουμε είναι τα ακόλουθα:

(1) Να εκτιμήσουμε ένα γραμμικό υπόδειγμα ενώ η πραγματική σχέση που συνδέει την Y με τις $X_1, X_2, \ldots, X_k$ δεν είναι γραμμική.

(2) Το υπόδειγμα είναι γραμμικό αλλά έχουμε παραλήψει ορισμένες σημαντικές ερμηνευτικές μεταβλητές ή έχουμε εισαγάγει ορισμένες άσχετες ερμηνευτικές μεταβλητές.

(3) Το υπόδειγμα είναι γραμμικό αλλά η γραμμική σχέση μεταβάλλεται διαχρονικά (παράλληλη μετατόπιση με μεταβολή του σταθερού όρου b_0 ή μεταβολή των κλίσεων $b_1, b_2, \ldots, b_k$).

Εξετάζουμε αναλυτικά τα προβλήματα εκτίμησης σε κάθε μια από τις περιπτώσεις αυτές.

1. Με τις επιφυλάξεις που έχουμε διατυπώσει για τους συντελεστές αυτούς.

4.3. ΕΚΤΙΜΗΣΗ ΤΩΝ ΠΑΡΑΜΕΤΡΩΝ ΣΕ ΜΗ ΓΡΑΜΜΙΚΑ ΥΠΟΔΕΙΓΜΑΤΑ

Η οικονομική θεωρία συχνά υποδεικνύει ότι η σχέση μεταξύ της εξαρτημένης και των ανεξάρτητων μεταβλητών είναι μη γραμμική. Αν δεν υπάρχουν πληροφορίες από την οικονομική θεωρία, το διάγραμμα των παρατηρήσεων του δείγματος (κυρίως στην περίπτωση μιας ερμηνευτικής μεταβλητής) είναι συχνά αρκετό για να αντιληφθούμε -εμπειρικά- ότι το υπόδειγμα δεν είναι γραμμικό.

Στην περίπτωση που το υπόδειγμα δεν είναι γραμμικό τότε έχουμε τις εξής διεξόδους:

(α) να προσαρμόσουμε στα στατιστικά στοιχεία απευθείας το μη γραμμικό υπόδειγμα, ή

(β) να μετασχηματίσουμε τις αρχικές παρατηρήσεις ώστε να χρησιμοποιήσουμε τα συμπεράσματα του γραμμικού υποδείγματος.

Εξετάζουμε παρακάτω χωριστά τις δυο αυτές προσεγγίσεις.

α) Απευθείας εκτίμηση μη γραμμικών υποδειγμάτων.

Η απευθείας εκτίμηση των μη γραμμικών υποδειγμάτων γίνεται είτε με τη μέθοδο ελαχίστων τετραγώνων είτε με τη μέθοδο της μέγιστης πιθανοφάνειας. Και στις δύο περιπτώσεις ο υπολογισμός των εκτιμητριών οδηγεί σε πολύπλοκες και συχνά δυσεπίλυτες εξισώσεις ακόμα και στις περιπτώσεις των πιο απλών υποδειγμάτων. Οι εξισώσεις αυτές είναι μη γραμμικές, υψηλού βαθμού, και πέρα από τις δυσκολίες που έχουμε να τις λύσουμε, δεν είναι βέβαιο ότι έχουν πραγματικές λύσεις. Αλλά, ακόμα και αν έχουν πραγματικές λύσεις, αυτές δεν ορίζονται μονοσήμαντα, δηλαδή ενδέχεται να έχουμε πολλές εκτιμήτριες του διανύσματος των αληθών παραμέτρων.

Παρόλα αυτά, κάτω από ορισμένες γενικές συνθήκες που εξασφαλίζουν την ύπαρξη και το μονοσήμαντο των λύσεων, οι εξισώσεις αυτές μπορούν να λυθούν "αριθμητικά" στον υπολογιστή με επαναληπτικές προσεγγιστικές διαδικασίες. Οι μέθοδοι αριθμητικής επίλυσης μη γραμμικών εξισώσεων παρουσιάζονται στην παράγραφο (A.9) σελ. 385 του παραρτήματος A. Περιγράφουμε στη συνέχεια την εφαρμογή των μεθόδων ελαχίστων τετραγώνων και μέγιστης πιθανοφάνειας σε μη γραμμικά υποδείγματα:

i) Η μη γραμμική μέθοδος ελαχίστων τετραγώνων.

Στη μη γραμμική μέθοδο ελαχίστων τετραγώνων έχουμε ένα σύνολο n παρατηρήσεων $Y_1, Y_2, \ldots, Y_n$, όπου

$$Y_i = f_i(b_1, b_2, \ldots, b_k) + e_i \qquad (4.3.1)$$

και ζητάμε τις τιμές των παραμέτρων $b_1, b_2, \ldots, b_k$ που ελαχιστοποιούν το άθροισμα των τετραγώνων των σφαλμάτων:

$$\sum_{i=1}^{n} [Y_i - f_i(b_1, b_2, \ldots, b_k)]^2. \qquad (4.3.2)$$

Οι συναρτήσεις f_i θα περιέχουν φυσικά και ερμηνευτικές μεταβλητές αλλά αυτό δε μας απασχολεί εδώ γιατί η (4.3.2) θα ελαχιστοποιηθεί ως προς τις παραμέτρους $b_1, b_2, \ldots, b_k$. Αν χρησιμοποιήσουμε το συμβολισμό των διανυσμάτων και των πινάκων η (4.3.2) γράφεται:

$$S(b) = [y - f(b)]'[y - f(b)]. \qquad (4.3.3)$$

Η αναζήτηση των εκτιμητριών $\hat{b}$ ελαχίστων τετραγώνων από την (4.3.3) θα γίνει με μια από τις "αριθμητικές" μεθόδους που περιγράφονται στην παράγραφο (A.9) του παραρτήματος A.

Το άθροισμα των τετραγώνων των καταλοίπων θα είναι:

$$\hat{e}'\hat{e} = [y - f(\hat{b})]'[y - f(\hat{b})] = S(\hat{b}) \qquad (4.3.4)$$

και αν υποθέσουμε ότι πληρούνται οι ασθενείς υποθέσεις τότε

η εκτιμήτρια $\hat{b}$ των μη γραμμικών ελαχίστων τετραγώνων ακολουθεί, ασυμπτωματικά, την κανονική κατανομή με μέσο b και πίνακα συνδιακυμάνσεων:

$$C(\hat{b}) = \hat{\sigma}^2 [F(\hat{b})'F(\hat{b})]^{-1} \qquad (4.3.5)$$

όπου

$$\hat{\sigma}^2 = \frac{\hat{e}'\hat{e}}{n} = \frac{S(\hat{b})}{n} \qquad (4.3.6)$$

και

$$F(\hat{b}) = \left.\frac{\partial f(b)}{\partial b}\right|_{\hat{b}}. \qquad (4.3.7)$$

Εκτός από την απευθείας ελαχιστοποίηση της (4.3.3) με μία από τις αριθμητικές μεθόδους είναι δυνατό να ελαχιστοποιήσουμε τη γραμμική προσέγγιση της $f(b)$ γύρω από κάποιο αρχικό σημείο b_0 χρησιμοποιώντας τη μέθοδο των Gauss-Newton. Η μέθοδος αυτή περιγράφεται στο εγχειρίδιο του G.S. Maddala "Econometrics" σελ. 174-175.

ii) Η μέθοδος της Μέγιστης Πιθανοφάνειας.

Αν $\{Y_1, Y_2, \ldots, Y_n\}$ είναι ένα τυχαίο δείγμα από τον πληθυσμό $f(., b)$, τότε, σύμφωνα με τα όσα έχουν εκτεθεί στην παράγραφο (B.4) του παραρτήματος B, η εκτιμήτρια μέγιστης πιθανοφάνειας του διανύσματος b είναι εκείνη που προκύπτει από την επίλυση της εξίσωσης

$$\frac{\partial L(b)}{\partial b} = 0 \qquad (4.3.8)$$

όπου

$$L(b) = \sum_{i=1}^{n} \log f(Y_i, b).$$

Τις περισσότερες φορές η εξίσωση (4.3.8) είναι μη γραμμική, υψηλού βαθμού και πρέπει να λυθεί με επαναληπτικές προσεγγιστικές διαδικασίες. Η επίλυση της (4.3.8), που ισοδυναμεί με την αναζήτηση του μεγίστου της $L(b)$ μπορεί να γίνει με μια από τις μεθόδους που περιγράφονται στην παράγραφο (A.9) του παραρτήματος A.

Αν $\hat{b}$ είναι η εκτιμήτρια μέγιστης πιθανοφάνειας του διανύσματος b, τότε, σύμφωνα με τις (Β.4.6) και (Β.4.8) του παραρτήματος Β, θα είναι συνεπής και θα ισχύει:

$$\sqrt{n}(\hat{b}-b) \sim N[0,-M^{-1}(b)].$$

Αν το μέγεθος του δείγματος είναι ικανοποιητικό τότε η ασυμπτωτική αυτή κατανομή προσεγγίζεται από την

$$\hat{b} \sim N[b,-\frac{1}{n}M^{-1}(\hat{b})] \qquad (4.3.9)$$

από την οποία μπορούμε να προχωρήσουμε κατά τα γνωστά στη στατιστική επαγωγή, χρησιμοποιόντας τους πίνακες της κανονικής κατανομής ή της κατανομής χ^2. Στην περίπτωση των εκτιμητριών μέγιστης πιθανοφάνειας μπορούμε, για τον έλεγχο υποθέσεων, να χρησιμοποιήσουμε και το κριτήριο του "λόγου της πιθανοφάνειας" ("likelihood ratio test"): Ας υποθέσουμε ότι τα k στοιχεία του διανύσματος b διαμερίζονται στα δύο υποσύνολα $b=[b_1\ b_2]$ όπου το b_1 περιέχει r στοιχεία και το b_2 τα υπόλοιπα (k-r) στοιχεία. Αν θέλουμε να ελέγξουμε την υπόθεση $b_1=b_1^*$, τότε υπολογίζουμε πρώτα την εκτιμήτρια μέγιστης πιθανοφάνειας του b χωρίς κανένα περιορισμό και στη συνέχεια την εκτιμήτρια μέγιστης πιθανοφάνειας του b κάτω από τον περιορισμό $b_1=b_1^*$. Αν $\hat{b}$ και $\hat{b}^*$ είναι οι δύο αυτές εκτιμήτριες και ορίσουμε το λόγο της πιθανοφάνειάς τους:

$$\lambda=\frac{L(\hat{b}^*)}{L(\hat{b})} \qquad (3.4.10)$$

αποδεικνύεται ότι η στατιστική $-2\ln\lambda$ ακολουθεί ασυμπτωτικά, την κατανομή χ^2 με r βαθμούς ελευθερίας. Η χρησιμοποίηση του κριτηρίου αυτού για τον έλεγχο υποθέσεων απαιτεί το μέγεθος του δείγματος να είναι αρκετά μεγάλο καθώς και τον διπλό υπολογισμό των εκτιμητριών μέγιστης πιθανοφάνειας με τους περιορισμούς και χωρίς τους περιορισμούς.

β) Εκτίμηση μη γραμμικών υποδειγμάτων με μετασχηματισμό.

Τα κυριώτερα μη γραμμικά υποδείγματα που μπορούμε να τα μετασχηματίσουμε έτσι ώστε να χρησιμοποιήσουμε τα πορίσματα του γραμμικού υποδείγματος είναι τα εξής:

i) Το πολυωνυμικό υπόδειγμα.

Το πολυωνυμικό υπόδειγμα έχει τη μορφή:

$$Y_i=b_0+b_1X_i+b_2X_i^2+\dots+b_kX_i^k+e_i \qquad (4.3.11)$$

και παρά το ότι είναι μη γραμμικό ως προς την ερμηνευτική μεταβλητή X, είναι εντούτοις γραμμικό ως προς τις παραμέτρους $b_0,b_1,\dots,b_k$. Αν θέσουμε:

$$X_i=Z_{1i},X_i^2=Z_{2i},X_i^3=Z_{3i},\dots,X_i^k=Z_{ki} \qquad (4.3.12)$$

τότε το πολυωνυμικό υπόδειγμα γράφεται:

$$Y_i=b_0+b_1Z_{1i}+b_2Z_{2i}+\dots+b_kZ_{ki}+e_i \qquad (4.3.13)$$

και είναι φανερό ότι, αν τα σφάλματα πληρούν τις γνωστές υποθέσεις, τότε αυτό μπορεί να εκτιμηθεί με τη μέθοδο Ε.Τ.Ω. Ως προς τον πίνακα Z αυτός είναι πλήρους βαθμού γιατί, παρά το ότι υπάρχουν ακριβείς σχέσεις μεταξύ των $Z_1,Z_2,\dots,Z_k$ οι σχέσεις αυτές δεν είναι γραμμικές. Σε ποια δύναμη της X θα σταματήσουμε θα εξαρτηθεί από την τιμή του $\bar{R}^2$ (διορθωμένου ως προς τους βαθμούς ελευθερίας). Το πολυωνυμικό υπόδειγμα είναι χρήσιμο στην εκτίμηση του κόστους C συναρτήσει της παραγωγής Q:

$$C_t=b_0+b_1Q_t+b_2Q_t^2+e_t$$

καθώς και άλλων οικονομικών σχέσεων που προσδιορίζονται από εξισώσεις δευτέρου ή ανωτέρου βαθμού ως προς τις ερμηνευτικές μεταβλητές. Η εκτίμηση του ύψους της παραγωγής για το οποίο θα έχουμε το μικρότερο κόστος εκτιμάται από την εξίσωση:

$$\frac{\partial\hat{C}}{\partial Q}=\hat{b}_1+2\hat{b}_2Q=0$$

και είναι ίσο με

$$Q=-\frac{\hat{b}_1}{2\hat{b}_2}.$$

(ii) Το λογαριθμικό υπόδειγμα.

Από την οικονομική θεωρία είναι γνωστή η μη γραμμική συνάρτηση παραγωγής Gobb-Douglas:

$$Q_t=AL_t^{b_1}K_t^{b_2} \qquad (4.3.14)$$

όπου Q_t είναι το ύψος της παραγωγής μιας επιχείρησης κατά την περίοδο t και L_t και K_t η εργασία και το κεφάλαιο που χρησιμοποιήθηκαν στην ίδια περίοδο (μετρημένα κατάλληλα). Εισάγοντας στην (4.3.14) ένα πολλαπλασιαστικό όρο σφάλματος έχουμε το στατιστικό υπόδειγμα

$$Q_t=AL_t^{b_1}K_t^{b_2}u_t \qquad (4.3.15)$$

το οποίο εύκολα διαπιστώνουμε ότι είναι μη γραμμικό και ως προς τις μεταβλητές και ως προς τις παραμέτρους b_1 και b_2.

Λογαριθμίζοντας και τα δύο μέλη της (4.3.15) έχουμε:

$$\log Q_t=\log A+b_1\log L_t+b_2\log K_t+\log u_t \qquad (4.3.16)$$

και θέτοντας

$$\log Q_t=Y_t,\ \log L_t=X_{1t},\ \log K_t=X_{2t},\ \log A=b_0,\ \log u_t=e_t$$

η (4.3.16) γράφεται:

$$Y_t=b_0+b_1X_{1t}+b_2X_{2t}+e_t. \qquad (4.3.17)$$

Γίνεται αμέσως φανερό ότι η (4.3.17) είναι ένα γραμμικό υπόδειγμα και ως προς τις μεταβλητές και ως προς τις παραμέτρους και συνεπώς μπορεί να εκτιμηθεί με τη μέθοδο Ε.Τ. κατά τα γνωστά, αρκεί να ισχύουν οι γνωστές υποθέσεις για τα σφάλματα. Πράγματι, αν οι υποθέσεις αυτές ισχύουν για τα σφάλματα u_t τότε θα ισχύουν και για τα e_t εκτός από την υπόθεση (υ.2) για την οποία υπάρχει κάποια περιπλοκή.

Για να ισχύει στο υπόδειγμα (4.3.17) η υπόθεση (υ.2): $E(e_t)=0$ πρέπει:

$$E(e_t)=E(\log u_t)=\frac{\Sigma\log u_t}{n}=\log\sqrt[n]{u_1u_2\dots u_n}=0 \qquad (4.3.18)$$

ΣΤΑΤ. ΣΕΛ. 102.

δηλαδή ο γεωμετρικός μέσος των σφαλμάτων $u_1,u_2,\dots,u_n$ να είναι η μονάδα*. Επειδή όμως οι διακυμάνσεις και οι συνδιακυμάνσεις των σφαλμάτων στο γραμμικό υπόδειγμα μετρούνται σε αποκλίσεις από τον αριθμητικό μέσο, για το λόγο αυτό, στο υπόδειγμα (4.3.15) γίνεται συνήθως η υπόθεση

$$E(u_t)=1 \qquad (4.3.19)$$

ΠΡΟΒΛΗΜΑΤΑ ΣΤΗΝ ΕΚΤΙΜΗΣΗ b_0.

διότι τότε ισχύει:

$$E(Q_t)=AL_t^{b_1}K_t^{b_2}E(u_t)$$
$$=AL_t^{b_1}K_t^{b_2}.$$

Στην περίπτωση αυτή έχουμε περιπλοκή στην εκτίμηση του σταθερού όρου b_0 στην (4.3.17). Γνωρίζουμε ότι ο γεωμετρικός μέσος ενός δείγματος είναι πάντα μικρότερος από τον αριθμητικό μέσο και το ίδιο ισχύει και για τους λογαρίθμους τους (αφού ο λογαριθμικός μετασχηματισμός είναι μονότονος), δηλαδή,

$$\log(\text{γεωμ. μέσου}) < \log(\text{αριθμ. μέσου}). \qquad (4.3.20)$$

Αλλά από την (4.3.18) προκύπτει ότι:

$$\log(\text{γεωμ. μέσου των } u_t)=E(e_t) \qquad (4.3.21)$$

ενώ από την (4.3.19) προκύπτει ότι:

$$\log(\text{αριθμ. μέσου των } u_t)=\log[E(u_t)]=\log 1=0. \qquad (4.3.22)$$

Από τις (4.3.20), (4.3.21) και (4.3.22) προκύπτει ότι για τα σφάλματα e_t του γραμμικού υποδείγματος (4.3.17) ισχύει:

$$E(e_t)<0, \qquad (4.3.23)$$

*Επειδή $\log 1=0$. Ο γεωμετρικός μέσος είναι το $\sqrt[n]{x_1x_2\dots x_n}$.

δηλαδή δεν πληρούται η βασική υπόθεση (υ.2) του γραμμικού υποδείγματος.

Αν στο δεύτερο μέλος της (4.3.17) προσθέσουμε και αφαιρέσουμε την ποσότητα $E(e_t)$, έχουμε:

$$Y_t=[b_0+E(e_t)]+b_1X_{1t}+b_2X_{2t}+[e_t-E(e_t)]$$

ή

$$Y_t=b_0^*+b_1X_{1t}+b_2X_{2t}+e_t^*. \qquad (4.3.24)$$

Για τα σφάλματα του υποδείγματος (4.3.24) εύκολα προκύπτει ότι:

$$E(e_t^*)=E[e_t-E(e_t)]=E(e_t)-E(e_t)=0,$$

επομένως η (υ.2) πληρούται και το υπόδειγμα μπορεί να εκτιμηθεί με τη μέθοδο Ε.Τ. Το μόνο πρόβλημα που έχουμε είναι ότι ο αντιλογάριθμος της εκτιμήτριας $\hat{b}_0^*$ δεν θα είναι η εκτιμήτρια του συντελεστή A της αρχικής συνάρτησης (4.3.14) αλλά κάποια άλλη ποσότητα, δηλαδή, κάτω από την υπόθεση $E(u_t)=1$ η εκτιμήτρια Ε.Τ. της παραμέτρου A δεν είναι αμερόληπτη.

Συνήθως όμως, στην εκτίμηση του λογαριθμικού υποδείγματος, ενδιαφερόμαστε όχι τόσο για την τιμή της σταθεράς A αλλά για τις τιμές των συντελεστών b_1 και b_2 που είναι οι ελαστικότητες της παραγωγής ως προς την εργασία και το κεφάλαιο:

ελαστικότητες.

$$\hat{b}_1=\frac{\partial(\widehat{\log Q})}{\partial(\log L)} \quad \text{και} \quad \hat{b}_2=\frac{\partial(\widehat{\log Q})}{\partial(\log K)}.$$

Ο λογαριθμικός μετασχηματισμός, εκτός από την εκτίμηση των παραμέτρων της συνάρτησης Gobb-Douglas, χρησιμοποιείται γενικά για την εκτίμηση συναρτήσεων σταθερής ελαστικότητας. Π.χ. μια συνάρτηση ζήτησης με σταθερή ελαστικότητα ως προς το εισόδημα και την τιμή έχει τη μορφή σταθερές ελαστικότητες.

$$Q_t=b_0P_t^{b_1}Y_t^{b_2}u_t$$

όπου Q_t είναι η ζητούμενη ποσότητα και P_t και Y_t η τιμή του

αγαθού και το διαθέσιμο εισόδημα αντίστοιχα. Εύκολα διαπιστώνουμε ότι

$$b_1=\frac{\partial Q}{\partial P}\cdot\frac{P}{Q}=\frac{\partial Q}{Q}\Big/\frac{\partial P}{P}=\frac{\partial(\log Q)}{\partial(\log P)}$$

είναι η σταθερή ελαστικότητα της ζήτησης ως προς την τιμή και b_2 είναι η σταθερή ελαστικότητα της ζήτησης ως προς το διαθέσιμο εισόδημα.

Τα ίδια μπορούμε να παρατηρήσουμε και στην ακόλουθη συνάρτηση σταθερής ελαστικότητας της ζήτησης χρήματος (M_t) ως προς το επιτόκιο (r_t) και το ύψος του εθνικού εισοδήματος (Y_t):

$$M_t=b_0r_t^{b_1}Y_t^{b_2}u_t.$$

(iii) Το ημιλογαριθμικό υπόδειγμα.

Ας υποθέσουμε ότι μια μεταβλητή (π.χ. η ιδιωτική κατανάλωση C) παρουσίασε, σε κάποια χρονική περίοδο, σταθερό ρυθμό αύξησης κατά 100g% το χρόνο. Η εξέλιξη της C δίνεται στον ακόλουθο πίνακα:

t	C
0	$C_0=A$
1	$C_1=A+Ag=A(1+g)$
2	$C_2=A(1+g)+[A(1+g)]g=A(1+g)^2$
⋮	⋮
t	$C_t=A(1+g)^t$.

Γενικά λοιπόν για την περίοδο t θα έχουμε:

$$C_t=AB^tu_t \qquad (4.3.25)$$

όπου $B=1+g$ και το σφάλμα έχει εισαχθεί πολ/κά με την υπόθεση, όπως και στο λογαριθμικό υπόδειγμα, ότι $E(u_t)=1$.

Λογαριθμίζοντας και τα δύο μέλη της (4.3.25) έχουμε:

$$\log C_t=\log A+(\log B)\cdot t+\log u_t \qquad (4.3.26)$$

και αν θέσουμε

$$\log C_t = Y_t,\ t = X_t,\ \log A = b_0,\ \log B = b_1,\ \log u_t = e_t$$

η (4.3.26) γράφεται:

$$Y_t = b_0 + b_1 X_t + e_t \qquad (4.3.27)$$

και η εκτίμησή της μπορεί να γίνει με τη μέθοδο Ε.Τ. έχοντας υπόψη ότι, ως προς την εκτίμηση του σταθερού όρου b_0, θα έχουμε τα προβλήματα που αντιμετωπίσαμε και στον λογαριθμικό σχηματισμό. Ο αντιλογάριθμος της εκτιμήτριας $\hat{b}_1$ θα είναι η εκτιμήτρια της παραμέτρου $B=1+g$ και συνεπώς η εκτιμήτρια της σταθερής g θα είναι η

$$\hat{g} = \hat{B} - 1 = \text{antilog}\,\hat{b}_1 - 1$$

και ο σταθερός ρυθμός μεταβολής (ή ο μέσος ρυθμός μεταβολής της C) θα είναι ίσος με 100g.

Κατά τον ίδιο τρόπο μπορούμε να εκτιμήσουμε και το υπόδειγμα στο οποίο η εξαρτημένη μεταβλητή Y μεταβάλλεται εκθετικά συναρτήσει του χρόνου (όπως π.χ. η εξέλιξη του πληθυσμού, κυρίως σε υποανάπτυκτες χώρες, όταν δεν υπάρχει έλεγχος γεννήσεων, ή η ανεμπόδιστη βιολογική ανάπτυξη ενός πληθυσμού μικροβίων κλπ.):

$$Y_t = Ae^{bt}u_t \quad \text{ή} \quad Y_t = Ae^{-bt}u_t$$

ή σε λογαριθμική μορφή:

$$\ln Y_t = \ln A + b\cdot t + \ln u_t$$

και

$$\ln Y_t = \ln A - b\cdot t + \ln u_t.$$

Γενικά, υποδείγματα της μορφής:

$$\log Y_i = \alpha_0 + \alpha_1 X_i + e_i \quad (\text{ή, ισοδύναμα: } Y_i = AB^{X_i}u_i) \qquad (4.3.28)$$

και

$$Y_i = b_0 + b_1 \log X_i + e_i \quad (\text{ή, ισοδύναμα: } X_i = AB^{Y_i}u_i) \qquad (4.3.29)$$

ονομάζονται "ημιλογαριθμικά υποδείγματα" γιατί μόνο η μία από τις μεταβλητές Y ή X υπεισέρχεται σε λογαριθμική μορφή στο υπόδειγμα.

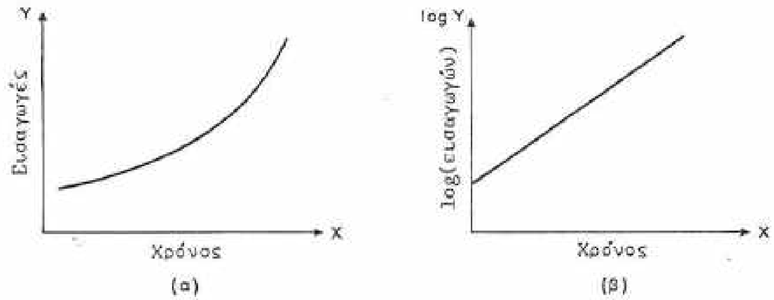
Για το υπόδειγμα (4.3.28) εύκολα προκύπτει ότι:

$$\alpha_1 = \frac{d(\log Y)}{dX} = \frac{1}{Y}\frac{dY}{dX} = \frac{dY}{Y}\Big/dX$$

$$= \frac{\text{Σχετική}^1\ \text{μεταβολή της Y}}{\text{Απόλυτη μεταβολή της X}}.$$

Άρα, σε περιπτώσεις που "απόλυτες" μεταβολές των τιμών της X προκαλούν μεταβολές της Y κατά "σταθερό ποσοστό" το κατάλληλο υπόδειγμα είναι το (4.3.28). Τέτοια υποδείγματα ονομάζονται υποδείγματα "σταθερού ρυθμού ανάπτυξης" και χρησιμοποιούνται για τη διαχρονική μέτριση του "σταθερού ρυθμού ανάπτυξης" ορισμένων μεταβλητών όπως η απασχόληση, ο δείκτης τιμών καταναλωτή, οι εισαγωγές, οι εξαγωγές, η παραγωγικότητα της εργασίας, το ακαθάριστο εθνικό προϊόν κλπ.

Π.χ. αν Y είναι οι εισαγωγές και X ο χρόνος τότε στο σχήμα 4.1(α) δίνεται η γραφική παράσταση της μεταβολής των εισαγωγών συναρτήσει του χρόνου, ενώ στο σχήμα 4.1(β) δίνεται η σχετική (ή αναλογική) μεταβολή των εισαγωγών συναρτήσει του χρόνου.



Σχήμα 4.1: (α) Η μεταβολή των εισαγωγών συναρτήσει του χρόνου
(β) Η σχετική (αναλογική) μεταβολή των εισαγωγών συναρτήσει του χρόνου.

1. Αν η σχετική μεταβολή πολ/στεί επί 100 γίνεται ποσοστιαία μεταβολή.

Αντίστοιχα, στο υπόδειγμα (4.3.29) η σταθερή $\frac{1}{b_1}$ είναι ο λόγος της σχετικής (ή αναλογικής) μεταβολής της Χ συναρτήσει της απόλυτης μεταβολής της Υ:

$$\frac{1}{b_1} = \frac{dX}{X} \Big/ dY$$

$$= \frac{\text{Σχετική μεταβολή της X}}{\text{Απόλυτη μεταβολή της Y}}.$$

Άρα το υπόδειγμα (4.3.29) είναι κατάλληλο για τις περιπτώσεις που σχετικές μεταβολές της ερμηνευτικής μεταβλητής Χ οδηγούν σε απόλυτες μεταβολές της εξαρτημένης μεταβλητής Υ.

Τα υποδείγματα "σταθερού ρυθμού ανάπτυξης" χρησιμοποιούνται και στις περιπτώσεις που ο ρυθμός ανάπτυξης δεν είναι μεν σταθερός αλλά επιθυμούμε να εκτιμήσουμε το "μέσο ρυθμό ανάπτυξης" ενός οικονομικού μεγέθους.

iv) Υπόδειγμα του αντίστροφου μετασχηματισμού.

Υποδείγματα της μορφής

$$Y_i = b_0 + b_1\left(\frac{1}{X_i}\right) + e_i \qquad (4.3.30)$$

ή

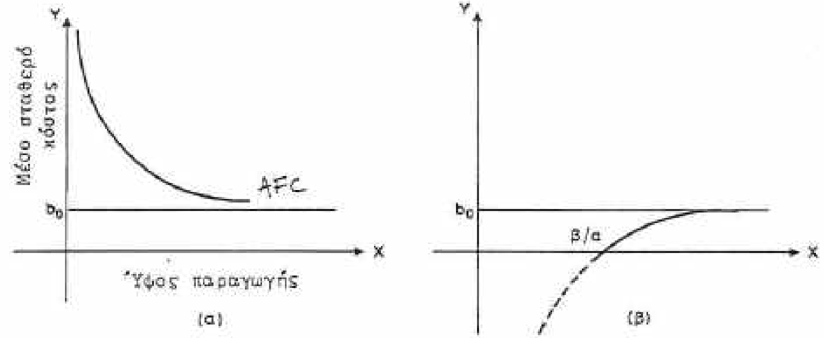
$$Y_i = b_0 - b_1\left(\frac{1}{X_i}\right) + e_i \qquad (4.3.31)$$

είναι γνωστά ως "υποδείγματα αντίστροφων μετασχηματισμών".

Για το υπόδειγμα (4.3.30) εύκολα προκύπτει ότι:

$$\frac{dY}{dX} = -\frac{b_1}{X^2}, \qquad (4.3.32)$$

δηλαδή η κλίση της Υ ως προς Χ (οριακή ροπή) είναι αρνητική και ελαττώνεται απόλυτα καθώς η Χ αυξάνει. Ακόμα διαπιστώνουμε ότι $X \to 0$ καθώς $Y \to \infty$ και ότι $X \to \infty$ καθώς $Y \to b_0$. Άρα η καμπύλη (4.3.30) είναι ο θετικός κλάδος της υπερβολής με ασύμπτωτες τον άξονα των Υ και την ευθεία $Y = b_0$ (Σχήμα 4.2(α)).



Σχήμα 4.2: Υποδείγματα του αντίστροφου μετασχηματισμού.

Με την ίδια ανάλυση για το υπόδειγμα (4.3.31) καταλήγουμε σε μια υπερβολή με ασύμπτωτες τον άξονα των Υ και την ευθεία $Y = b_0$ εκ των κάτω (Σχήμα 4.2(β)).

Ειδικό χαρακτηριστικό των υποδειγμάτων (4.3.30) και (4.3.31) είναι το ότι έχουν ενσωματομένη κάποια ασυμπτωτική (ή οριακή τιμή) προς την οποία τείνει η εξαρτημένη μεταβλητή Υ καθώς η ανεξάρτητη μεταβλητή Χ τείνει στο άπειρο. Η εκτίμηση της ασυμπτωτικής (ή οριακής) αυτής τιμής δίνεται από την εκτιμήτρια του σταθερού όρου b_0. Π.χ. αν Υ είναι το μέσο σταθερό κόστος και Χ είναι το ύψος της παραγωγής, τότε το μέσο σταθερό κόστος ελαττώνεται με την αύξηση της παραγωγής αλλά δεν μπορεί να γίνει μικρότερο από κάποια οριακή τιμή b_0.

Ένα ακόμα σχετικό παράδειγμα είναι το ακόλουθο υπόδειγμα της Κεϋνσιανής ζήτησης χρήματος:

$$r_t = b_0 + b_1\left(\frac{1}{M_t - M_t^*}\right) + e_t$$

όπου

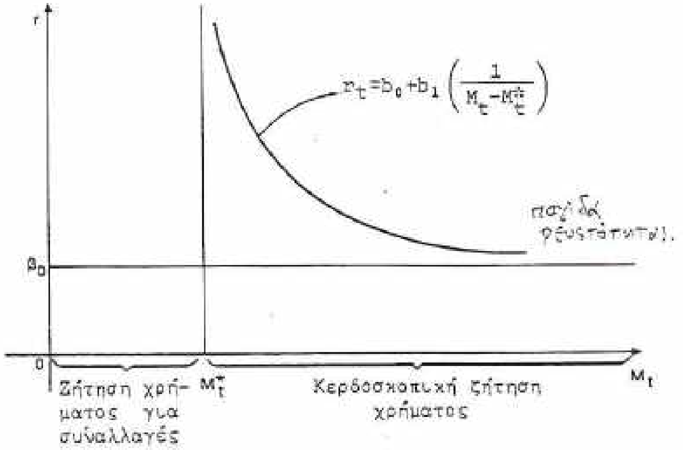
r_t: είναι το επιτόκιο

M_t: η συνολική ζήτηση χρήματος

M_t^*: η ζήτηση χρήματος για συναλλαγές (ανεξάρτητη από το επιτόκιο)

$M_t - M_t^*$: η κερδοσκοπική ζήτηση χρήματος.

Στο υπόδειγμα αυτό ο όρος b_0 αντιπροσωπεύει την παγίδα της ρευστότητας (liquidity trap) (Σχ. 4.3)



Σχήμα 4.3: Η παγίδα της ρευστότητας (liquidity trap).

(v) Άλλα μη γραμμικά υποδείγματα.

Υπάρχουν και άλλα μη γραμμικά υποδείγματα τα οποία μπορούμε να εκτιμήσουμε με τη μέθοδο Ε.Τ. αφού τα μετασχηματίσουμε κατάλληλα. Μερικά από αυτά είναι τα εξής:

(1) $Y = e^{b_0 - \frac{b_1}{X_t}} u_t$ ή, ισοδύναμα, $\ln Y_t = b_0 - b_1\left(\frac{1}{X_t}\right) + \ln u_t$

γνωστό ως υπόδειγμα "λογαριθμικού αντίστροφου μετασχηματισμού".

(2) $$Y_t = b_0 + b_1 t + b_2 \eta\mu\left(\frac{2\pi t}{12}\right) + e_t$$

που είναι ένα υπόδειγμα ανάπτυξης με γραμμική τάση και κυκλικές διακυμάνσεις (να κάνετε τη γραφική παράσταση του υποδείγματος αυτού).

(3) $$Y = b_0 b_1^{t} b_2^{X_t}$$

υπόδειγμα εκθετικής ανάπτυξης (ημιλογαριθμικός μετασχηματισμός).

Στα υποδείγματα αυτά να δώσετε την ερμηνεία των συντελεστών b_0, b_1 και b_2 και να τα μετασχηματίσετε έτσι ώστε να μπορούν να εκτιμηθούν με τη μέθοδο Ε.Τ.

4.4. ΣΥΝΕΠΕΙΕΣ ΑΠΟ ΤΗΝ ΠΑΡΑΛΕΙΨΗ ΣΗΜΑΝΤΙΚΩΝ ΕΡΜΗΝΕΥΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ ΚΑΙ ΑΠΟ ΤΗΝ ΕΙΣΑΓΩΓΗ ΑΣΧΕΤΩΝ ΕΡΜΗΝΕΥΤΙΚΩΝ ΜΕΤΑΒΛΗΤΩΝ ΣΤΟ ΓΡΑΜΜΙΚΟ ΥΠΟΔΕΙΓΜΑ

Οι συνέπειες από την 1 παράλειψη σημαντικών ερμηνευτικών μεταβλητών ή 2 από την εισαγωγή άσχετων ερμηνευτικών μεταβλητών σε μια παλινδρόμηση αποτελούν μέρος του γενικότερου προβλήματος του εσφαλμένου προσδιορισμού του υποδείγματος που επιθυμούμε να εκτιμήσουμε. Το γενικότερο αυτό πρόβλημα θα μας απασχολήσει και σε επόμενα κεφάλαια.

(α) Συνέπειες από την παράλειψη σημαντικών ερμηνευτικών μεταβλητών.

Ας υποθέσουμε ότι η αληθής παλινδρόμηση που ερμηνεύει την Y είναι η

$$y = X_1 b_1 + X_2 b_2 + e \quad \text{Πραγματική} \qquad (4.4.1)$$

και ότι από λάθος ή από αδυναμία συγκέντρωσης των απαραίτητων στατιστικών στοιχείων για τις μεταβλητές X_2 ο ερευνητής εκτιμά το διάνυσμα b_1 από την παλινδρόμηση

$$y = X_1 a_1 + u. \quad \text{Εκτιμώμενη.} \qquad (4.4.2)$$

Η εκτιμήτρια Ε.Τ. από την ελλειπή παλινδρόμηση (4.4.2) είναι, κατά τα γνωστά, η

$$a_1 = (X_1'X_1)^{-1}X_1'y$$
$$= (X_1'X_1)^{-1}X_1'(X_1b_1 + X_2b_2 + e)$$

(διότι η αληθής τιμή του διανύσματος y δίνεται από την παλινδρόμηση (4.4.1))

$$=b_1+(X_1'X_1)^{-1}X_1'X_2b_2+(X_1'X_1)^{-1}X_1'e$$

$$=b_1+\hat{D}b_2+(X_1'X_1)^{-1}X_1'e$$

ο $\hat{D}$ είναι πίνακας

και εύκολα προκύπτει ότι

$$E(\hat{a}_1)=b_1+\hat{D}b_2 \tag{4.4.3}$$

όπου

$$\hat{D}=(X_1'X_1)^{-1}X_1'X_2 \tag{4.4.4}$$

είναι η εκτιμήτρια των συντελεστών των παλινδρομήσεων

$V=(X'X)^{-1}$

$$X_2=DX_1+V \quad \text{βοηθητικές εξισώσεις.} \tag{4.4.5}$$

των μεταβλητών X_1 πάνω στις X_2. Οι εξισώσεις (4.4.5) ονομάζονται συνήθως "βοηθητικές εξισώσεις" (auxiliary equations).

Διαπιστώνουμε λοιπόν ότι η εκτιμήτρια $\hat{a}_1$ του διανύσματος b_1 από την ελλειπή παλινδρόμηση είναι μεροληπτική εκτιμήτρια του αληθούς διανύσματος b_1 στον πληθυσμό. Η μεροληπτικότητα είναι ίση με $\hat{D}b_2$ και θα είναι μηδέν αν:

i) $\hat{D}=0$. Αυτό θα συμβαίνει αν οι μεταβλητές X_1 δεν συσχετίζονται γραμμικά με τις μεταβλητές X_2. Στην περίπτωση αυτή οι μεταβλητές X_1 και X_2 ονομάζονται ορθογώνιες και αυτό εκφράζεται από το ότι $X_1'X_2=0$. Διάλεξη της παλινδρόμησης: $X_{2i}=c+\delta X_{1i}+u_i$

ii) $b_2=0$. δηλαδή αν οι μεταβλητές X_2 δεν συσχετίζονται γραμμικά με την Y.

Πρέπει να τονίσουμε ότι, ακόμα και στην περίπτωση που οι X_1 και X_2 είναι ορθογώνιες, η εκτίμηση του b_1 πρέπει να γίνει από την αληθή παλινδρόμηση γιατί, αν χρησιμοποιηθεί η ελλειπής παλινδρόμηση, τότε τα σφάλματα u θα είναι στην πραγματικότητα ίσα με

$$u=X_2b_2+e$$

και

$$E(u)=X_2b_2\neq 0$$

δηλαδή δεν θα ισχύει η βασική υπόθεση (υ.2) του γραμμικού υποδείγματος, εκτός αν $b_2=0$ που είναι αδύνατο γιατί η αληθής παλινδρόμηση είναι η (4.4.1).

β) Συνέπειες από την εισαγωγή άσχετων ερμηνευτικών μεταβλητών.

Έστω ότι η αληθής παλινδρόμηση είναι η

$$y=X_1b_1+e \tag{4.4.6}$$

από την οποία προκύπτει ότι:

$$\hat{b}_1=(X_1'X_1)^{-1}X_1'y,\quad E(\hat{b}_1)=b_1,\quad C(\hat{b}_1)=\sigma^2(X_1'X_1)^{-1}. \tag{4.4.7}$$

Υποθέστε τώρα ότι ο ερευνητής εισάγει ορισμένες άσχετες μεταβλητές X_2 και εκτιμά την παλινδρόμηση

$$y=X_1a_1+X_2a_2+u. \tag{4.4.8}$$

Η εξίσωση (4.4.8) μπορεί να γραφτεί:

$$y=[X_1\ X_2]\begin{bmatrix}a_1\\a_2\end{bmatrix}+u \tag{4.4.9}$$

και οι "κανονικές εξισώσεις" για την εκτίμηση των a_1 και a_2 από την εξίσωση (4.4.9) είναι οι εξής:

$$\begin{bmatrix}X_1'X_1 & X_1'X_2\\ X_2'X_1 & X_2'X_2\end{bmatrix}\begin{bmatrix}\hat{a}_1\\ \hat{a}_2\end{bmatrix}=\begin{bmatrix}X_1'y\\ X_2'y\end{bmatrix} \tag{4.4.10}$$

από την επίλυση των οποίων προκύπτει (να δειχτεί ως άσκηση) ότι:

$$\hat{a}=(X_1'Q_2X_1)^{-1}X_1'Q_2y \tag{4.4.11}$$

όπου

$$Q_2=I-X_2(X_2'X_2)^{-1}X_2' \quad \text{και} \quad Q_2'=Q_2,\ Q_2^2=Q_2,\ Q_2X_2=0. \tag{4.4.12}$$

Από την (4.4.11) εύκολα προκύπτει (να δειχτεί ως άσκηση) ότι:

$$E(\hat{a}_1)=b_1 \quad \text{και} \quad C(\hat{a}_1)=\sigma^2(X_1'Q_2X_1)^{-1}. \tag{4.4.13}$$

η κανονική είναι $\sigma^2(X'X)^{-1}$

Από τις (4.4.7) και (4.4.13) έχουμε

$$C(\hat{a}_1)-C(\hat{b}_1)=\sigma^2[(X_1'Q_2X_1)^{-1}-(X_1'X_1)^{-1}]. \qquad (4.4.14)$$

Αλλά

$$X_1'Q_2X_1=X_1'[I-X_2(X_2'X_2)^{-1}X_2']X_1$$
$$=X_1'X_1-X_1'X_2(X_2'X_2)^{-1}X_2'X_1. \qquad (4.4.15)$$

Αν θέσουμε

$$X_1'Q_2X_1=B, \quad X_1'X_1=A \quad \text{και} \quad X_1'X_2(X_2'X_2)^{-1}X_2'X_1=C$$

τότε

$$A=B+C$$

και επειδή οι πίνακες A και B είναι θετικοί πεπερασμένοι και ο πίνακας C είναι μη αρνητικός πεπερασμένος, σύμφωνα με την (Α.6.6, ix) του παραρτήματος Α, ο πίνακας

$$B^{-1}-A^{-1}=(X_1'Q_2X_1)^{-1}-(X_1'X_1)^{-1} \qquad (4.4.16)$$

είναι μη αρνητικός πεπερασμένος και από τις (4.4.16) και (4.4.14) είναι φανερό ότι:

$$C(\hat{a}_1)-C(\hat{b}_1)\geqslant 0$$

ή

$$C(\hat{a}_1)\geqslant C(\hat{b}_1) \qquad (4.4.17)$$

στην οποία το ίσον θα ισχύει αν $X_2=0$ οπότε η (4.4.8) θα ταυτίζεται με την αληθή παλινδρόμηση ή αν $X_1'X_2=0$ δηλαδή οι X_1 και X_2 είναι ορθογώνιες.

Από την (4.4.13) προκύπτει ότι η $\hat{a}_1$ εξακολουθεί να εκτιμά αμερόληπτα το αληθές διάνυσμα b_1 παρά την εισαγωγή των άσχετων μεταβλητών X_2, αλλά από την (4.4.17) γίνεται φανερό ότι δεν είναι πλέον η άριστη εκτιμήτρια του b_1 (εκτός αν οι άσχετες μεταβλητές X_2 είναι ορθογώνιες προς τις X_1).

μη αποτελεσματική
εκτιμήτρια
$V(\hat{a}_1) > V(\hat{b}_1)$

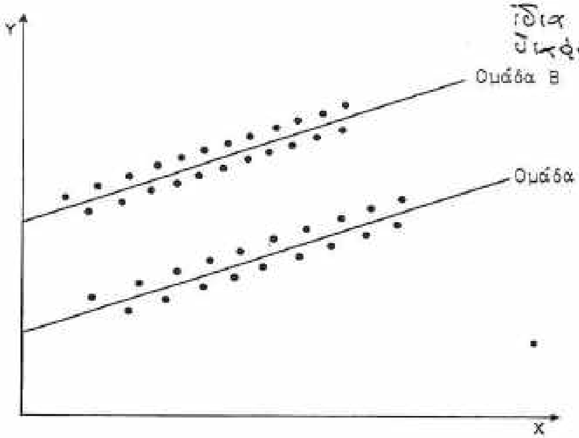
4.5. ΕΚΤΙΜΗΣΗ ΓΡΑΜΜΙΚΩΝ ΥΠΟΔΕΙΓΜΑΤΩΝ ΣΤΑ ΟΠΟΙΑ Η ΓΡΑΜΜΙΚΗ ΣΧΕΣΗ ΜΕΤΑΒΑΛΛΕΤΑΙ ΔΙΑΣΤΡΩΜΑΤΙΚΑ Η ΔΙΑΧΡΟΝΙΚΑ. ΨΕΥΔΟΜΕΤΑΒΛΗΤΕΣ

Πολλές φορές η παράλειψη σημαντικών ερμηνευτικών μεταβλητών οφείλεται στο ότι οι μεταβλητές αυτές είναι ποιοτικές (π.χ. το επίπεδο της μόρφωσης, διαφορές ως προς το φύλο, κοινωνικές διαφορές, κοινωνικές ή πολεμικές αναταραχές στην περίοδο που καλύπτεται από το δείγμα, διαφορές ηλικίας κλπ.) με αποτέλεσμα να μην μπορούμε να τις μετρήσουμε ποσοτικά και να τις εισαγάγουμε στην παλινδρόμηση. Τέτοιες ποιοτικές μεταβολές μπορούν να εισαχθούν στην παλινδρόμηση με τη βοήθεια των "ψευδομεταβλητών" αρκεί η επίδρασή τους να μη μεταβάλλει τη γραμμικότητα του υποδείγματος αλλά ή να το μετατοπίζει παράλληλα (μεταβάλλοντας το σταθερό όρο) ή να μεταβάλλει ορισμένους συντελεστές (μεταβολή ορισμένων οριακών ροπών). Αναφέρουμε παρακάτω μερικά παραδείγματα χρησιμοποίησης των ψευδομεταβλητών για την εισαγωγή ποιοτικών στοιχείων σε μια παλινδρόμηση.

ΠΟΙΟΤΙΚΕΣ
ΜΕΤΑΒΛΗΤΕΣ.

Ας υποθέσουμε ότι η σχέση μεταξύ του εισοδήματος Y και των σπουδών X για δύο ομάδες επαγγελμάτων δίνεται στο σχήμα 4.4.

ίδια κλίση αλλά
διαφορετικό σταθερό
όρο.



Σχήμα 4.4: Σχέση του εισοδήματος Y και των ετών σπουδών X για τις ομάδες Α και Β.

Από το σχήμα γίνεται φανερό ότι οι παλινδρομήσεις για τις δύο ομάδες έχουν σχεδόν την ίδια κλίση αλλά διαφορετικό σταθερό όρο. Αν τα δείγματα για τις δύο ομάδες είναι αρκετά μεγάλα τότε μπορούμε να εκτιμήσουμε τις ακόλουθες χωριστές παλινδρομήσεις:

A ΙΔΙΑ

ΔΙΑΦΟΡΕΤΙΚΗ ΣΤΑΘΕΡΑ

$$Y_i=\alpha_1+bX_i+e_i, \quad \text{για την ομάδα A}$$

$$Y_i=\alpha_2+bX_i+u_i, \quad \text{για την ομάδα B.}$$

① μεγάλα δείγματα.

Αν όμως τα δείγματα είναι μικρά ή αν επιθυμούμε να εκτιμήσουμε μια κοινή παλινδρόμηση για τις δύο ομάδες, τότε χρησιμοποιούμε τα δύο δείγματα ως ένα δείγμα και εκτιμούμε την παλινδρόμηση:

② μικρά δείγματα.

$$Y_i=\alpha_1+(\alpha_2-\alpha_1)D_i+bX_i+e_i$$

όπου

$D_i=0$, αν η παρατήρηση i ανήκει στην ομάδα A
$\quad=1$, αν η παρατήρηση i ανήκει στην ομάδα B.

Εμείς βέβαια θα εκτιμήσουμε την παλινδρόμηση:

$$\hat{Y}_i=\hat{\alpha}_1+\hat{c}D_i+\hat{b}X_i$$

και θα έχουμε υπόψη ότι ο συντελεστής $\hat{c}$ της ψευδομεταβλητής εκτιμά τη διαφορά $\widehat{\alpha_2-\alpha_1}$ των δύο σταθερών όρων εφόσον οι δύο ομάδες έχουν την ίδια κλίση b ως προς τα έτη σπουδών.

B

Ως ένα δεύτερο παράδειγμα ας θεωρήσουμε την ιδιωτική κατανάλωση C ως συνάρτηση του διαθέσιμου εισοδήματος Y και ας υποθέσουμε ότι διαθέτουμε διαχρονικά στοιχεία στα οποία περιλαμβάνεται μια περίοδος μη ομαλή π.χ. πόλεμος. Αν πιστεύουμε ότι η συνάρτηση της κατανάλωσης παρουσιάζει μόνο παράλληλη μετατόπιση προς τα κάτω κατά την περίοδο του πολέμου, ενώ η οριακή ροπή προς κατανάλωση παραμένει σταθερή, τότε μπορούμε να εκτιμήσουμε την παλινδρόμηση:

$$C_t=\alpha_1+(\alpha_2-\alpha_1)D_t+bY_t+e_t$$

όπου

$D_t=1$, για τα έτη πολέμου
$\quad=0$, για τα άλλα έτη.

Αν, ακόμα, πιστεύουμε ότι η συνάρτηση της κατανάλωσης παρουσιάζει παράλληλη μετατόπιση και κατά τα μεταπολεμικά έτη, σε σχέση με την προπολεμική περίοδο και την περίοδο του πολέμου, τότε χρησιμοποιούμε δύο ψευδομεταβλητές:

$$C_t=\alpha_1+(\alpha_2-\alpha_1)D_{1t}+(\alpha_3-\alpha_1)D_{2t}+bY_t+e_t$$

όπου

$D_{1t}=1$, για τα έτη πολέμου
$\quad=0$, αλλού

$D_{2t}=1$, για τα μεταπολεμικά έτη
$\quad=0$, αλλού.

Η εξίσωση που θα εκτιμήσουμε είναι η

$$\hat{C}_t=\hat{\alpha}_1+\hat{c}_1D_{1t}+\hat{c}_2D_{2t}+\hat{b}Y_t$$

όπου

$\hat{\alpha}_1$: είναι η κατανάλωση επιβίωσης για τα προπολεμικά έτη,

$\widehat{\alpha_2-\alpha_1}=\hat{c}_1$: είναι η διαφορά της κατανάλωσης επιβίωσης κατά τα προπολεμικά έτη και κατά τα έτη πολέμου, και

$\widehat{\alpha_3-\alpha_1}=\hat{c}_2$: είναι η διαφορά της κατανάλωσης επιβίωσης κατά τα προπολεμικά και μεταπολεμικά έτη.

Μπορούμε να χρησιμοποιήσουμε τρεις ψευδομεταβλητές D_1, D_2 και D_3 για τις τρεις περιόδους αλλά στην περίπτωση αυτή δεν πρέπει να χρησιμοποιήσουμε σταθερό όρο γιατί τότε ο πίνακας των ανεξάρτητων μεταβλητών:

$$X=\left[\begin{array}{ccccc}
\Sigma\tau. & D_1 & D_2 & D_3 & Y \\
1 & 1 & 0 & 0 & Y_1 \\
1 & 1 & 0 & 0 & Y_2 \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
1 & 1 & 0 & 0 & \vdots \\
 & & & & \\
1 & 0 & 1 & 0 & \vdots \\
1 & 0 & 1 & 0 & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
1 & 0 & 1 & 0 & \vdots \\
 & & & & \\
1 & 0 & 0 & 1 & \vdots \\
1 & 0 & 0 & 1 & \vdots \\
\vdots & \vdots & \vdots & \vdots & \vdots \\
1 & 0 & 0 & 1 & Y_n
\end{array}\right]
\begin{array}{l}
\\ \\ \text{Προπολεμική περίοδος} \\ \\ \\ \\ \\ \text{Περίοδος πολέμου} \\ \\ \\ \\ \\ \text{Μεταπολεμική περίοδος} \\ \\ \\
\end{array}$$

δε θα είναι πλήρους βαθμού αφού κάθε στοιχείο της πρώτης στήλης είναι γραμμικός συνδιασμός (άθροισμα) των αντίστοιχων στοιχείων της δεύτερης, τρίτης και τέταρτης στήλης. Γενικά, αν οι ομάδες είναι λ τότε: 1 ή θα χρησιμοποιήσουμε (λ-1) ψευδομεταβλητές και σταθερό όρο, ή λ ψευδομεταβλητές χωρίς σταθερό όρο.

Ως ένα ακόμα παράδειγμα, ας υποθέσουμε ότι διαθέτουμε διαστρωματικά στοιχεία για την κατανάλωση C και το εισόδημα Y ενός αριθμού οικογενειών και ότι διαθέτουμε επιπλέον τα εξής αναλυτικά στοιχεία:

i) το φύλο του αρχηγού της οικογένειας,

ii) την ηλικία X του αρχηγού της οικογένειας η οποία διακρίνεται σε τρεις κατηγορίες: $X<25$, $25\leqslant X\leqslant 50$, $X>50$,

iii) τη μόρφωση του αρχηγού της οικογένειας επίσης σε τρεις κατηγορίες:

Δεν έχει τελειώσει το Λύκειο,

Είναι απόφοιτος Λυκείου αλλά όχι Πανεπιστημίου.

Είναι πτυχιούχος Πανεπιστημίου.

Στο παράδειγμα αυτό μπορούμε π.χ. να χρησιμοποιήσουμε τις εξής ψευδομεταβλητές:

$D_1=1$, αν ο αρχηγός είναι άνδρας
$=0$, αν είναι γυναίκα
$D_2=1$, αν $X<25$
$=0$, αλλού
$D_3=1$, αν $25\leqslant X\leqslant 50$
$=0$, αλλού
$D_4=1$, αν δεν έχει τελειώσει το Λύκειο
$=0$, αλλού
$D_5=1$, αν είναι απόφοιτος Λυκείου αλλά όχι Πανεπιστημίου
$=0$, αλλού.

Εύκολα παρατηρούμε ότι σε κάθε κατηγορία ο αριθμός των ψευδομεταβλητών είναι κατά ένα μικρότερος από τον αριθμό των περιπτώσεων κάθε κατηγορίας, αν βέβαια θέλουμε να χρησιμοποιήσουμε σταθερό όρο στην εξίσωση:

$$C_i=\alpha+bY_i+c_1D_{1i}+c_2D_{2i}+c_3D_{3i}+c_4D_{4i}+c_5D_{5i}+e_i.$$

Είναι βέβαια φανερό ότι η πιο πάνω ανάλυση ισχύει μόνο στην περίπτωση που οι διάφορες κατηγορίες των καταναλωτών έχουν την ίδια οριακή ροπή προς κατανάλωση b, αλλά διαφορετικό επίπεδο διαβίωσης (διαφορετικό σταθερό όρο). Π.χ. για το υπόδειγμά μας αν ο αρχηγός της οικογένειας είναι γυναίκα, 45 ετών, απόφοιτος Λυκείου τότε η εξίσωση της κατανάλωσης της οικογένειας θα είναι:

$$C=(\alpha+c_3+c_5)+bY.$$

B ΔΙΑΦΟΡΕΣ ΣΤΗΝ ΚΛΙΣΗ ΚΑΙ ΣΤΗΝ ΣΤΑΘΕΡΑ.

Οι ψευδομεταβλητές μπορούν επίσης να χρησιμοποιηθούν όταν πιστεύουμε ότι υπάρχουν διαφορές και στην κλίση των παλινδρομήσεων για τις διάφορες ομάδες. Έτσι, αν

$$Y_{Ai}=\alpha_1+b_1X_{Ai}+e_{Ai},\quad \text{για την ομάδα A}$$

$$Y_{Bi}=\alpha_2+b_2X_{Bi}+e_{Bi},\quad \text{για την ομάδα B}$$

μπορούμε να χρησιμοποιήσουμε την κοινή παλινδρόμηση:

$$Y_i=\alpha_1+(\alpha_2-\alpha_1)D_{1i}+b_1X_i+(b_2-b_1)D_{2i}X_i+e_i$$

όπου

$D_{1i}=1$, αν η παρατήρηση i ανήκει στην ομάδα Β
$=0$, αλλού
$D_{2i}=1$, αν η παρατήρηση i ανήκει στην ομάδα Β
$=0$, αλλού.

Ο συντελεστής της D_1 είναι η διαφορά των σταθερών όρων ενώ ο συντελεστής της D_2 είναι η διαφορά των κλίσεων για τις δύο ομάδες.

Ας υποθέσουμε ακόμα ότι διαθέτουμε στοιχεία για τρεις ομάδες Α, Β, Γ και ότι:

$$Y_{Ai}=\alpha_1+b_1X_{Ai}+e_{Ai}$$
$$Y_{Bi}=\alpha_2+b_1X_{Bi}+e_{Bi}$$
$$Y_{\Gamma i}=\alpha_2+b_2X_{\Gamma i}+e_{\Gamma i}.$$

Τότε μπορούμε να χρησιμοποιήσουμε την ακόλουθη κοινή εξίσωση και για τις τρεις ομάδες:

$$Y_i=\alpha_1+(\alpha_2-\alpha_1)D_{1i}+b_1X_i+(b_2-b_1)D_{2i}X_i+e_i$$

όπου

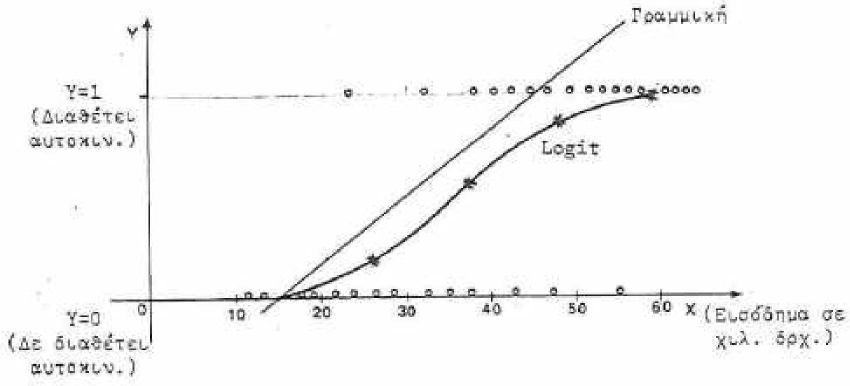
$D_{1i}=1$, για τις παρατηρήσεις των ομάδων Β και Γ
$=0$, αλλού
$D_{2i}=1$, για τις παρατηρήσεις της ομάδας Γ
$=0$, αλλού.

Η εξαρτημένη μεταβλητή Y=1 Y=0

Σε ορισμένες περιπτώσεις είναι δυνατόν και η εξαρτημένη μεταβλητή Y να παίρνει τις τιμές 0 και 1. Αν π.χ. X είναι το εισόδημα μιας οικογένειας και Y=1 ή 0 αν η οικογένεια αυτή διαθέτει ή όχι αυτοκίνητο, τότε το διάγραμμα των παρατηρήσεων του δείγματος θα δίνεται στο σχήμα 4.5.

Έστω ότι η προσαρμογή της ευθείας Ε.Τ. στο δείγμα έδωσε την παλινδρόμηση

$$\hat{Y}=-0,25+0,05X$$



Σχήμα 4.5: Η κατοχή αυτοκινήτου Y συναρτήσει του εισοδήματος X.

από την οποία προκύπτει ότι η αύξηση της X κατά μία μονάδα προκαλεί αύξηση της Y κατά 0.05 μονάδες. Αυτό μπορεί να ερμηνευτεί ως αύξηση της πιθανότητας κατά 0.05 η οικογένεια να έχει αυτοκίνητο. Όμως από το σχήμα 4.5. γίνονται φανερά τα μειονεκτήματα ενός τέτοιου γραμμικού υποδείγματος για τη μέτρηση πιθανοτήτων: α) Η $\hat{Y}$ μπορεί να πάρει τιμές μεγαλύτερες από τη μονάδα και μικρότερες από το μηδέν και β) από το γραμμικό υπόδειγμα προκύπτει ότι σταθερές μεταβολές της X προκαλούν ίσες μεταβολές της $\hat{Y}$ ενώ στην πραγματικότητα, όσο πλησιάζει το 0 ή το 1, σταθερές μεταβολές της X προκαλούν μικρότερες μεταβολές της $\hat{Y}$.

Προβλήμ. τα πιθανότητας.

Τα προβλήματα αυτά αντιμετωπίζονται συνήθως ως εξής: στο κάτω άκρο χρησιμοποιούμε την εξίσωση

$$\ln P=\alpha+bX \quad \text{κάτω άκρο} \qquad (4.5.1)$$

ενώ στο άνω άκρο την εξίσωση

$$\ln(1-P)=\alpha+bX. \quad \text{Άνω άκρο} \qquad (4.5.2)$$

Τις δύο εξισώσεις μπορούμε να τις συγχωνεύσουμε στην ακόλουθη εξίσωση:

$$\ln P-\log(1-P)=\alpha+bX \qquad (4.5.3)$$

ή

$$\ln\left(\frac{P}{1-P}\right)=\alpha+bX. \qquad (4.5.4)$$

Η συγχώνευση αυτή μπορεί να γίνει διότι, όταν βρισκόμαστε στην περιοχή του 0, η μεταβολή της P είναι πολύ μεγαλύτερη από τη μεταβολή της (1-P) και έτσι η χρησιμοποίηση της $\ln\left(\frac{P}{1-P}\right)$ δεν οδηγεί σε σημαντικό λάθος. Π.χ. αν η P μεταβληθεί από 0,01 σε 0,02 τότε η lnP μεταβάλλεται κατά 0,70 ενώ η ln(1-P) μόνο κατά 0,01. Το αντίστροφο συμβαίνει στην περιοχή του 1: αν η P αυξηθεί από 0,98 σε 0,99 τότε η ln(1-P) μεταβάλλεται κατά 0,70 ενώ η lnP μόνο κατά 0,01. Το αρνητικό πρόσημο εισάγεται στην εξίσωση έτσι ώστε οι μεταβολές της ln(1-P) να είναι θετικές και η $\ln\left(\frac{P}{1-P}\right)$ να είναι αύξουσα συνάρτηση της P.

Η εξίσωση (4.5.4) μπορεί να γραφτεί και με τη μορφή

Y=P

$$\boxed{P=\frac{1}{1+e^{-(\alpha+bX)}}} \qquad (4.5.5)$$

logit curve

και ονομάζεται "λογιστική καμπύλη" (logistic curve ή logit curve) και η γραφική της παράσταση δίνεται στο σχήμα 4.5.

Αν χρησιμοποιήσουμε τη μέθοδο Ε.Τ. για την εκτίμηση της (4.5.4) τότε θα αντιμετωπίσουμε σοβαρό πρόβλημα διότι δεν μπορούμε να χρησιμοποιήσουμε τις αρχικές τιμές 0 και 1 της Y (=P). Για να αντιμετωπίσουμε τη δυσκολία αυτή μπορούμε να χωρίσουμε τις παρατηρήσεις του δείγματος σε ομάδες και να υπολογίσουμε τις πραγματικές πιθανότητες κάθε ομάδας που θα είναι, ασφαλώς, διάφορες από το 0 και το 1. Η ομαδοποίηση όμως αυτή εισάγει κάποιο σφάλμα στη μέτρηση της X και, όπως θα διαπιστώσουμε αργότερα, αυτό έχει ως αποτέλεσμα οι εκτιμήτριες Ε.Τ. να είναι μεροληπτικές και ασυνεπείς.

Η εκτίμηση της (4.5.5) μπορεί όμως να γίνει και με τη μέθοδο της μέγιστης πιθανοφάνειας, δηλαδή να επιλέξουμε τις τιμές $\hat{a}$ και $\hat{b}$ των παραμέτρων που μεγιστοποιούν τη συνάρτηση

πιθανοφάνειας με κάποια από τις αριθμητικές μεθόδους που περιγράφονται στην παράγραφο (Α.9) του παραρτήματος Α. Περισσότερα για την εκτίμηση της λογιστικής καμπύλης βλέπε στο βιβλίο των Judge et al[1].

Τις ψευδομεταβλητές μπορούμε ακόμα να χρησιμοποιήσουμε i) για την απαλοιφή της εποχικότητας από τις χρονολογικές σειρές που παρουσιάζουν εποχιακές διακυμάνσεις και ii) την εκτίμηση ποσοτικών σχέσεων ανάμεσα σε μεταβλητές που παρουσιάζουν εποχιακές διακυμάνσεις.

i) Απαλοιφή της εποχικότητας από μια χρονολογική σειρά.

Ας υποθέσουμε ότι διαθέτουμε τριμηνιαίες παρατηρήσεις για n έτη (4n παρατηρήσεις) για τη μεταβλητή Y, η οποία παρουσιάζει εποχιακές διακυμάνσεις στα 4 τρίμηνα και ότι επιθυμούμε να απαλείψουμε τις εποχιακές αυτές διακυμάνσεις από τη χρονολογική σειρά της Y. Γνωρίζουμε από τη στατιστική ότι για το σκοπό αυτό μπορούμε να χρησιμοποιήσουμε κινητούς μέσους ή άλλες μεθόδους. Μπορούμε όμως να απαλείψουμε την εποχικότητα και με τη βοήθεια των ψευδομεταβλητών.

Ας θεωρήσουμε τις 4 ψευδομεταβλητές (μία για κάθε τρίμηνο)

$D_{ij}=1$, αν η παρατήρηση i ανήκει στο j τρίμηνο
$\quad =0$, αλλού

και ας εκτιμήσουμε την παλινδρόμηση:

$$Y_i=b_1D_{i1}+b_2D_{i2}+b_3D_{i3}+b_4D_{i4}+e_i \qquad (4.5.6)$$

ή, σε διανυσματική μορφή,

$$y=Db+e. \qquad (4.5.7)$$

Οι εκτιμήτριες Ε.Τ. της (4.5.7) είναι:

1. G. Judge, W. Griffiths, R.C. Hill, T.C. Lee, *The theory and Practice of Econometrics*, Wiley, New York 1980, pp. 592-609.

$$\hat{b}=(D'D)^{-1}D'y$$

όπου D είναι ο πίνακας:

$$D=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\0&0&0&1\\ \cdots&\cdots&\cdots&\cdots\\0&0&0&1\end{bmatrix}$$

→ n φορές ο μοναδιαίος.
↓
έτη

ενώ τα κατάλοιπα της παλινδρόμησης (4.5.7) είναι: ΠΡΟΣΟΧΗ

$$\begin{aligned}\hat{e}&=y-\hat{y}\\&=y-D\hat{b}\\&=y-D(D'D)^{-1}D'y\\&=[I-D(D'D)^{-1}D']y\\&=My\\&=y^{\alpha},\end{aligned}$$

Η χρονολογική σειρά μετά την απαλοιφή της εποχικότητας είναι τα κατάλοιπα της (4.5.6). !!!

όπου

$$M=I-D(D'D)^{-1}D'.$$

Τα κατάλοιπα της παλινδρόμησης (4.5.7) εκφράζουν, όπως είναι γνωστό, το μέρος της Y που δεν ερμηνεύεται από τις ψευδομεταβλητές D_{ij} (δηλαδή από τις εποχιακές διακυμάνσεις) άρα y^{α} είναι οι τιμές της Y απαλλαγμένες από την εποχικότητα. Πρέπει όμως να παρατηρήσουμε τα εξής: α) από τις ιδιότητες των καταλοίπων, γνωρίζουμε ότι το άθροισμά τους -άρα και ο μέσος τους- είναι μηδέν και δεν μπορούμε να δεχτούμε ότι οι παρατηρήσεις μιας χρονολογικής σειράς, μετά την απαλοιφή της εποχικότητας, θα έχουν μέσο μηδέν. Το πρόβλημα αυτό το αντιμετωπίζουμε προσθέτοντας στις απαλλαγμένες από την εποχικότητα χρονολογικές σειρές το μέσο $\bar{Y}$ των αρχικών παρατηρήσεων. Άρα η απαλλαγμένη από την εποχικότητα χρονολογική σειρά θα αποτελείται από τις παρατηρήσεις:

Παρατηρήσεις

$$\boxed{Y^{*}_{i}=\bar{Y}+Y^{\alpha}_{i},\quad i=1,2,\ldots,4n,}\qquad(4.5.8)$$

και θα έχει τον ίδιο μέσο με την αρχική χρονολογική σειρά.

β) Από τη φύση του πίνακα D είναι εύκολο να διαπιστώσουμε ότι

$$\hat{b}=\begin{bmatrix}\hat{b}_1\\ \hat{b}_2\\ \hat{b}_3\\ \hat{b}_4\end{bmatrix}=\begin{bmatrix}\frac{1}{n}\sum_{i=1}^{n}Y_{i1}\\ \frac{1}{n}\sum_{i=1}^{n}Y_{i2}\\ \frac{1}{n}\sum_{i=1}^{n}Y_{i3}\\ \frac{1}{n}\sum_{i=1}^{n}Y_{i4}\end{bmatrix}\qquad(4.5.9)$$

όπου Y_{ij} είναι η αρχική παρατήρηση της Y που ανήκει στο j τρίμηνο του i έτους. Άρα τα κατάλοιπα Y^{α}_{i} της (4.5.7) εκφράζουν τις αποκλίσεις των αρχικών τιμών της Y από τους μέσους των αντίστοιχων τριμήνων. Αν όμως η χρονολογική σειρά περιέχει, εκτός από τις εποχιακές διακυμάνσεις, και κάποια διαχρονική τάση -γραμμική ή όχι- τότε αυτή θα επηρεάσει τους μέσους των τριμήνων, άρα και την απαλλαγμένη από την εποχικότητα χρονολογική σειρά. Για να αντιμετωπίσουμε το πρόβλημα αυτό πρέπει να απαλλάξουμε από την Y και τη διαχρονική τάση εισάγοντας στην (4.5.6), ως ερμηνευτική μεταβλητή, το χρόνο (στην πρώτη δύναμη αν η τάση είναι γραμμική, στη δεύτερη δύναμη αν είναι δευτέρου βαθμού κλπ.):

ΔΙΑΧΡΟΝΙΚΗ ΤΑΣΗ

$$Y_i=b_1D_{i1}+b_2D_{i2}+b_3D_{i3}+b_4D_{i4}+b_5t+e_i$$

$$Y_i=b_1D_{i1}+b_2D_{i2}+b_3D_{i3}+b_4D_{i4}+b_5t+b_6t^2+u_i \quad \text{κλπ.}$$

και οι παρατηρήσεις της απαλλαγμένης από την εποχικότητα και τη διαχρονική τάση χρονολογικής σειράς θα υπολογίζονται από την (4.5.8).

ii) Εκτίμηση ποσοτικών σχέσεων ανάμεσα σε μεταβλητές που παρουσιάζουν εποχιακές διακυμάνσεις.

Έστω ότι επιθυμούμε να εκτιμήσουμε τη σχέση που υπάρχει ανάμεσα στην κατανάλωση C και το εισόδημα X μιας οικογένειας ή ενός συνόλου οικογενειών χρησιμοποιόντας τριμηνιαίες παρατηρήσεις. Έστω ακόμα ότι C_i και X_i είναι οι αρχικές χρονολογικές σειρές και C_i^* και X_i^* οι αντίστοιχες χρονολογικές σειρές απαλλαγμένες από την εποχικότητα (και τη γραμμική τάση αν υπάρχει).

Μπορούμε να διακρίνουμε τις εξής περιπτώσεις:

α) Η κατανάλωση παρουσιάζει εποχιακές διακυμάνσεις ενώ το εισόδημα όχι. Στην περίπτωση αυτή μπορούμε να εκτιμήσουμε την παλινδρόμηση:

$$C_i^*=b_0+b_1X_i+e_i, \quad i=1,2,\ldots,4n, \qquad (4.5.10)$$

αν ενδιαφερόμαστε να εκτιμήσουμε την απαλλαγμένη από την εποχικότητα κατανάλωση, ή την ∃ σταθερός → 3 dummy.

$$C_i=b_0+b_1X_i+c_1D_{i1}+c_2D_{i2}+c_3D_{i3}+e_i \qquad (4.5.11)$$

αν ενδιαφερόμαστε για την πραγματική κατανάλωση. Σημειώστε ότι στην εξίσωση (4.5.11) έχουμε χρησιμοποιήσει ψευδομεταβλητές μόνο για τα τρία τρίμηνα διότι υπάρχει σταθερός όρος.

β) Η κατανάλωση δεν παρουσιάζει εποχιακές διακυμάνσεις ενώ το εισόδημα παρουσιάζει. Η κατάλληλη παλινδρόμηση είναι τότε η

$$C_i=b_0+b_1X_i^*+e_i. \qquad (4.5.12)$$

γ) Και η κατανάλωση και το εισόδημα παρουσιάζουν εποχιακές διακυμάνσεις αλλά όχι του ίδιου τύπου. Η κατάλληλη παλινδρόμηση στην περίπτωση αυτή είναι η

$$C_i=b_0+b_1X_i^*+c_1D_{i1}+c_2D_{i2}+c_3D_{i3}+e_i. \qquad (4.5.13)$$

Σε πολλές περιπτώσεις -ιδιαίτερα όταν εκτιμούμε σχέσεις ανάμεσα σε μακρο-μεταβλητές διαθέτουμε τις παρατηρήσεις και σε αρχικές τιμές και απαλλαγμένες από την εποχικότητα αλλά η θεωρία δεν είναι σε θέση να μας πληροφορήσει ποιες χρονολογικές σειρές πρέπει να χρησιμοποιήσουμε. Στην

περίπτωση αυτή πολύ χρήσιμα είναι τα συμπεράσματα που απέδειξε ο Lovell[1] και τα οποία συνοψίζονται ως εξής: Έστω ότι επιθυμούμε να εκτιμήσουμε τη γραμμική σχέση ανάμεσα στις μεταβλητές Y και $X_1, X_2, \ldots, X_k$ και ότι y και X είναι αντίστοιχα το διάνυσμα και ο πίνακας των αρχικών παρατηρήσεων ενώ y^α και X^α είναι το διάνυσμα και ο πίνακας των χρονολογικών σειρών που είναι απαλλαγμένες από την εποχικότητα, δηλαδή LOVELL

$$y^\alpha=My \quad \text{και} \quad X^\alpha=MX$$

όπου

$$M=I-D(D'D)^{-1}D'.$$

Ας θεωρήσουμε τις παλινδρομήσεις

$$\begin{aligned} y&=Xc_1+Db_1+e_1 \\ y^\alpha&=X^\alpha c_2+e_2 \qquad (4.5.14) \\ y&=X^\alpha c_3+e_3 \\ y&=X^\alpha c_4+Db_4+e_4. \end{aligned}$$

Ο Lovell απέδειξε ότι

$$c_1=c_2=c_3=c_4 \qquad (4.5.15)$$

δηλαδή αν εκτιμήσουμε την παλινδρόμηση της Y πάνω i) στις $X_1, X_2, \ldots, X_k$ και τις D_1, D_2, D_3, D_4 ή ii) στις $X_1^\alpha, X_2^\alpha, \ldots, X_k^\alpha$ και στις D_1, D_2, D_3, D_4, θα πάρουμε τις ίδιες ακριβώς εκτιμήτριες με εκείνες που θα πάρουμε από την παλινδρόμηση της Y^α πάνω στις $X_1^\alpha, X_2^\alpha, \ldots, X_k^\alpha$.

Είδαμε όμως ότι οι χρονολογικές σειρές $Y^\alpha, X_1^\alpha, X_2^\alpha, \ldots, X_k^\alpha$ παρουσιάζουν τα μειονεκτήματα που αναφέραμε προηγούμενα και δεν είναι σκόπιμο να χρησιμοποιηθούν στην εκτίμηση των πιο

1. M.C. Lovell, "Seasonal Adjustment of Economic Time series and Multiple Regresion Analysis", *Journal of the American Statistical Association*, vol 58 (1963), pp. 993-1010.

πάνω παλινδρομήσεων. Τίθεται λοιπόν το ερώτημα αν θα ισχύουν τα συμπεράσματα του Lovell όταν αντί των $Y^{\alpha}, X_1^{\alpha}, X_2^{\alpha}, \ldots, X_k^{\alpha}$ χρησιμοποιήσουμε τις σωστά απαλλαγμένες από την εποχικότητα μεταβλητές $Y^*, X_1^*, X_2^*, \ldots, X_k^*$. Θεωρητικά, η ισότητα των c_1, c_2, c_3, c_4, δεν μπορεί να αποδειχθεί στην περίπτωση αυτή.

Ο Johnston (Econometric Methods, 2nd Edition, McGraw-Hill, Co, London 1972, σελ. 191-192) αναφέρει ότι σχετικοί πειραματισμοί με διάφορες εξισώσεις του οικονομετρικού υποδείγματος του Cambridge για τη Μεγάλη Βρεττανία έδειξαν ότι η (4.5.15) ισχύει με πολύ καλή προσέγγιση (δεκαδικών ψηφίων) όταν στη χρονική περίοδο που καλύπτεται από το δείγμα η οικονομία παρουσιάζει σταθερό ρυθμό ανάπτυξης χωρίς κυκλικές διακυμάνσεις. Δεν γνωρίζουμε όμως αν τα συμπεράσματα ισχύουν και σε άλλες περιπτώσεις και για το λόγο αυτό κάθε περίπτωση πρέπει να εξετάζεται με την ανάλογη προσοχή. Για μια πληρέστερη διερεύνηση του θέματος των εποχιακών διακυμάνσεων των μεταβλητών στην πολλαπλή παλινδρόμηση ο αναγνώστης μπορεί να συμβουλευτεί την αρθρογραφία στην υποσημείωση 1.

4.6. Η ΥΠΟΘΕΣΗ (υ.2): E(e) = 0

Η υπόθεση (υ.2) σημαίνει ότι οι θετικές και αρνητικές τιμές των σφαλμάτων $e_1, e_2, \ldots, e_n$ έχουν άθροισμα μηδέν. Αν θεωρήσουμε το υπόδειγμα

---

1. D.W. Jorgenson, "Minimum Variance, Linear, Unbiased Seasonal Adjustment of Economic Time Series", *Journal of the American Statistical Association, Vol. 59 (1964)*, pp. 681-724.
M.C. Lovell, "Alternative Axiomatizations of Seasonal Adjustment", *Journal of the American Statistical Association, Vol. 61 (1966)*, pp. 800-802.
G.W. Ladd, "Regression Analyses of Seasonal Data", *Journal of the American Statistical Association, Vol. 69 (1974)*, pp. 402-420.
J.J. Thomas and K.F. Wallis, "Seasonal Variation in Regression Analysis", *Journal of the Royal Statistical Society*, Series A, 1971, pp. 57-72.
K.F. Wallis, "Seasonal Adjustment and Relations between Variables", *Journal of the American Statistical Association Vol. 69 (1974)* pp. 18-31.
C.A. Sims, "Seasonality in Regression", *Journal of the American Statistical Association Vol. 69 (1974)*, pp. 618-626.
J.A. Pesando, "Seasonal Variability in Distributed Lag Models", *Journal of the American Statistical Association Vol. 67 (1972)*, pp. 311-312.

$$y = Xb + e$$

τότε η υπόθεση (υ.2) θα ισχύει αν

$$E(y/X_1, X_2, \ldots, X_k) = Xb.$$

Επομένως η ισχύς της υπόθεσης (υ.2) εξαρτάται από τη γραμμικότητα της σχέσης μεταξύ των Y και $X_1, X_2, \ldots, X_k$ και από το αν η γραμμική αυτή σχέση είναι η πραγματική. Επειδή λοιπόν οι θεωρητικές τιμές των σφαλμάτων $e_1, e_2, \ldots, e_n$ δεν είναι γνωστές είναι αδύνατο να ελέγξουμε απευθείας την ισχύ της υπόθεσης (υ.2). Η ισχύς της θα εξαρτηθεί (i) από το σωστό προσδιορισμό της συναρτησιακής σχέσης ανάμεσα στις Y και $X_1, X_2, \ldots, X_k$ (γραμμική ή μη γραμμική), (ii) από το αν έχουμε εισαγάγει στο υπόδειγμα όλες τις σημαντικές ερμηνευτικές μεταβλητές και (iii) από το αν υπάρχουν συστηματικά σφάλματα στη μέτρηση της Y.

$V(e) = E(e_i^2) = E(ee') = \sigma^2 I$

4.7. Η ΥΠΟΘΕΣΗ (υ.3): E(ee) = σ²I

Η τρίτη υπόθεση του κλασικού γραμμικού υποδείγματος απαιτεί τη σφαιρικότητα των σφαλμάτων και είναι η σύνθεση των δύο επιμέρους υποθέσεων:

$E(e)$
(υ.3α): $E(e_i^2) = \sigma^2$, $i = 1, 2, \ldots, n$ ← ΕΤΕΡΟΣΚΕΔΑΣΤΙΚΟΤΗΤΑ
$E(ee')$
(υ.3β): $E(e_i e_j) = 0$, $i, j = 1, 2, \ldots, n$ $(i \neq j)$. ← ΑΥΤΟΣΥΣΧΕΤΙΣΗ.

Αν δεν ισχύει η (υ.3α) τα σφάλματα ονομάζονται "ετεροσκεδαστικά" (Heteroskedastic disturbances), ενώ αν δεν ισχύει η (υ.3β) τα σφάλματα ονομάζονται "αυτοσυσχετιζόμενα" (autocorrelated disturbances). Αν τα σφάλματα δεν ικανοποιούν την (υ.3), δηλαδή είναι ετεροσκεδαστικά είτε αυτοσυσχετιζόμενα, τότε ονομάζονται γενικά "μη σφαιρικά" σε αντίθεση με τη σφαιρικότητα που εκφράζει η (υ.3). Και η περίπτωση της ετεροσκεδαστικότητας και η περίπτωση της αυτοσυσχέτισης είναι συνηθισμένες στην ποσοτική οικονομική ανάλυση.

Αν αντικαταστήσουμε το μοναδιαίο πίνακα I με ένα $n \times n$ θετικό πεπερασμένο πίνακα W, τότε επιτρέπουμε την ύπαρξη και ετεροσκεδαστικότητας (τα διαγώνια στοιχεία του W δεν είναι όλα ίσα) και αυτοσυσχέτισης (τα μη διαγώνια στοιχεία του W δεν είναι όλα μηδέν).

Ας θεωρήσουμε το υπόδειγμα:

$$y=Xb+e \tag{4.7.1}$$

όπου

$$E(e)=0, \quad E(ee')=\sigma^2 W_{n\times n} \tag{4.7.2}$$

και W είναι ένας γ ν ω σ τ ό ς θετικός πεπερασμένος πίνακας βαθμού n×n (μη ιδιάζων). Όλες οι άλλες υποθέσεις του κλασικού γραμμικού υποδείγματος θεωρούμε ότι ισχύουν. Υποθέτουμε ότι ο πίνακας W είναι γνωστός γιατί δεν είναι δυνατόν να εκτιμήσουμε $n^2+(k+1)$ παραμέτρους (τα n^2 στοιχεία του W και τα k+1 στοιχεία του διανύσματος b) από τους n βαθμούς ελευθερίας του δείγματος. Επομένως απομένουν προς εκτίμηση τα στοιχεία του b και η παράμετρος σ^2.

Από την άλγεβρα των πινάκων [βλ. (Α.6.6.(VII))] γνωρίζουμε ότι για κάθε θετικό πεπερασμένο μη ιδιάζοντα πίνακα W υπάρχει ένας μη ιδιάζων πίνακας K, τέτοιος ώστε να ισχύουν οι σχέσεις

$$K'WK=I \quad \text{και} \quad KK'=W^{-1}. \tag{4.7.3}$$

Ας θεωρήσουμε τον ακόλουθο γραμμικό μετασχηματισμό του υποδείγματος (4.7.1):

$$K'y=K'Xb+K'e \tag{4.7.4}$$

ή

$$y^*=X^*b+e^* \tag{4.7.5}$$

όπου

$$K'y=y^*, \quad K'X=X^* \quad \text{και} \quad K'e=e^*. \tag{4.7.6}$$

Για τα σφάλματα e^* αποδεικνύεται ότι:

$$E(e^*)=E(K'e)=K'E(e)=0 \tag{4.7.7}$$

και

$$E(e^*e^{*\prime})=E(K'ee'K)=K'E(ee')K=\sigma^2(K'WK)=\sigma^2 I. \tag{4.7.8}$$

$K'(\sigma^2 W)K$

Άρα, μπορούμε να εφαρμόσουμε τη μέθοδο Ε.Τ. στο υπόδειγμα (4.7.5) και να προσδιορίσουμε τις εκτιμήτριες

$$\begin{aligned}\tilde{b}&=(X^{*\prime}X^*)^{-1}X^{*\prime}y^*\\&=(X'KK'X)^{-1}X'KK'y\\&=(X'W^{-1}X)^{-1}X'W^{-1}y \rightarrow \text{BLUE}\end{aligned} \tag{4.7.9}$$

Generalized Least Squares; Aitkens.

για το (4.7.5) ή το (4.7.1).

που ονομάζονται "εκτιμήτριες των γενικευμένων ελαχίστων τετραγώνων του Aitkens (Aitkens Generalised Least Squares ή, σε συντομογραφία AGLS).

Κάτι άλλο που πρέπει ακόμα να παρατηρήσουμε είναι ότι οι AGLS εκτιμήτριες $\tilde{b}$ ελαχιστοποιούν το άθροισμα των τετραγώνων των σφαλμάτων του υποδείγματος (4.7.5):

$$e^{*\prime}e^*=e'KK'e=e'W^{-1}e \tag{4.7.10}$$

δηλαδή ελαχιστοποιούν την τετραγωνική μορφή $e'W^{-1}e$ και όχι το άθροισμα των τετραγώνων των σφαλμάτων του υποδείγματος (4.7.1).

Οι εκτιμήτριες $\tilde{b}$ είναι οι άριστες γραμμικές αμερόληπτες εκτιμήτριες (BLUE) για το υπόδειγμα (4.7.5). Πρέπει να δείξουμε ότι είναι επίσης οι άριστες γραμμικές εκτιμήτριες και για το αρχικό υπόδειγμα (4.7.1). Αποδεικνύουμε σχετικά το ακόλουθο

ΘΕΩΡΗΜΑ:

Οι εκτιμήτριες $\tilde{b}$ είναι οι άριστες γραμμικές εκτιμήτριες για τις παραμέτρους του υποδείγματος

$$y=Xb+e$$

ΑΠΟΔΕΙΞΗ:

Αν θέσουμε $H=(X'W^{-1}X)^{-1}X'W^{-1}$, όπου H είναι ένας σταθερός πίνακας, η (4.7.9) γράφεται:

$$\bar{b}=(X'W^{-1}X)^{-1}X'W^{-1}y$$
$$=Hy,$$

άρα οι $\bar{b}$ είναι "γραμμικές" εκτιμήτριες. Εύκολα επίσης προκύπτει ότι:

$$\bar{b}=Hy$$
$$=H(Xb+e)$$
$$=b+He\ , \quad HX=(X'W^{-1}X)^{-1}X'W^{-1}X=I \qquad (4.7.11)$$

και

$$E(\bar{b})=b+HE(e)=b \qquad (4.7.12)$$

άρα οι $\bar{b}$ είναι "αμερόληπτη" εκτιμήτρια του διανύσματος b στον πληθυσμό.

Για τον πίνακα συνδιακυμάνσεων των εκτιμητριών $\bar{b}$ έχουμε

$$C(\bar{b})=E(\bar{b}-b)(\bar{b}-b)'$$
$$=E(Hee'H')$$
$$=HE(ee')H'$$
$$=\sigma^2 HWH' \qquad [\text{δείξτε ότι } HWH'=(X'W^{-1}X)^{-1}]$$
$$=\sigma^2(X'W^{-1}X)^{-1}. \qquad (4.7.13)$$

$C(\bar{b}) =$

Απομένει τώρα να δείξουμε ότι οι εκτιμήτριες $\bar{b}$ είναι οι άριστες μέσα στην κλάση των γραμμικών και αμερόληπτων εκτιμητριών του διανύσματος b.

Έστω λοιπόν μια άλλη γραμμική και αμερόληπτη εκτιμήτρια του b, η

$$b^{+}=Py. \qquad (4.7.14)$$

Τότε

$$E(b^{+})=E(Py)$$
$$=E[P(Xb+e)]$$
$$=PXb$$

και επειδή η b^{+} είναι αμερόληπτη εκτιμήτρια πρέπει $PX=I$ (το οποίο συμβαίνει για τον πίνακα H αφού $HX=I$). Ο πίνακας P μπορεί να γραφτεί ως εξής:

$$P=H+D$$

όπου D είναι ένας σταθερός πίνακας. Συνεπώς

$$PX=HX+DX$$
$$=I+DX$$

και επειδή $PX=I$, εύκολα προκύπτει ότι

$$DX=0 \quad \text{και} \quad DWH'=\underbrace{DX}(X'W^{-1}X)^{-1}=0 \qquad (4.7.15)$$

Άρα

$$C(b^{+})=E(b^{+}-b)(b^{+}-b)'$$
$$=E(Pee'P')$$
$$=\sigma^2(PWP')$$
$$=\sigma^2[(H+D)W(H+D)']$$
$$=\sigma^2(HWH'+\underbrace{HWD'}_{0}+\underbrace{DWH'}_{0}+DWD') \qquad [\text{λόγω της } (4.7.15)]$$
$$=\sigma^2(HWH')+\sigma^2(DWD')$$
$$=C(\bar{b})+\sigma^2 DWD' \qquad (4.7.16)$$

όπου DWD' είναι ένας μη αρνητικός πεπερασμένος πίνακας. Επομένως

$$V(\bar{b}_i)\leqslant V(b_i^{+}),\quad i=1,2,\ldots,k$$

και το ίσον ισχύει τότε και μόνο τότε αν $D=0$ δηλαδή αν $P=H$, και αυτό συμπληρώνει την απόδειξη του θεωρήματος.

Απαιτείται τώρα μια εκτιμήτρια της σ^2. Από το υπόδειγμα (4.7.1) προκύπτει ότι:

$$\bar{e}=y-X\bar{b}$$
$$=y-X(X'W^{-1}X)^{-1}X'W^{-1}y$$

$$=[I-X(X'W^{-1}X)^{-1}X'W^{-1}]y$$

$$=Ly \qquad (4.7.17)$$

όπου

$$L^2=L \quad \text{αλλά} \quad L\neq L',$$

δηλαδή ο πίνακας L είναι αυτοδύναμος αλλά όχι συμμετρικός. Εύκολα διαπιστώνουμε ότι:

$$L=I-XH$$

και

$$\begin{aligned} LX&=X-XHX\\ &=X-X\\ &=0. \end{aligned}$$

Άρα

$$\begin{aligned} \tilde{e}&=Ly\\ &=L(X\tilde{b}+e)\\ &=\underbrace{LX}_{0}\tilde{b}+Le\\ &=Le. \end{aligned}$$

Ας θεωρήσουμε τη μαθηματική ελπίδα της τετραγωνικής μορφής $\tilde{e}'W^{-1}\tilde{e}$:

$$\begin{aligned} E(\tilde{e}'W^{-1}\tilde{e})&=E\mathrm{tr}(W^{-1}\tilde{e}\tilde{e}')\\ &=E\mathrm{tr}(W^{-1}Lee'L')\\ &=\sigma^2\mathrm{tr}(W^{-1}LWL')\\ &=\sigma^2\mathrm{tr}(L'W^{-1}LW)\\ &=\sigma^2\mathrm{tr}(W^{-1}LW), \qquad (L'W^{-1}L=W^{-1}L)^1 \end{aligned}$$

1. $L'W^{-1}L=(I-H'X')W^{-1}(I-XH)$

$$\begin{aligned} &=W^{-1}-H'X'W^{-1}-W^{-1}XH+H'X'W^{-1}XH, \qquad X'W^{-1}XH=X'W^{-1}X(X'W^{-1}X)^{-1}X'W^{-1}=X'W^{-1}\\ &=W^{-1}-H'X'W^{-1}-W^{-1}XH+H'X'W^{-1}\\ &=W^{-1}[I-XH]=W^{-1}L \end{aligned}$$

$$\begin{aligned} &=\sigma^2\mathrm{tr}(LWW^{-1})\\ &=\sigma^2\mathrm{tr}L\\ &=\sigma^2\mathrm{tr}(I-XH), \qquad \mathrm{tr}(XH)=\mathrm{tr}(HX)=\mathrm{tr}I_{k+1}\\ &=\sigma^2[n-(k+1)]. \end{aligned}$$

Άρα η

$$\tilde{\sigma}^2=\frac{\tilde{e}'W^{-1}\tilde{e}}{n-(k+1)} \qquad (4.7.18)$$

είναι η αμερόληπτη εκτιμήτρια της σ^2 και ο πίνακας

$$\tilde{\sigma}^2(X'W^{-1}X)^{-1}$$

είναι η αμερόληπτη εκτιμήτρια του πίνακα συνδιακυμάνσεων των εκτιμητριών $\tilde{b}$.

Φτάσαμε λοιπόν στο εξής γενικό συμπέρασμα: Για το υπόδειγμα

$$y=Xb+e, \quad E(e)=0, \quad E(ee')=\sigma^2W, \qquad (4.7.19)$$

αν ισχύουν οι υπόλοιπες υποθέσεις, τότε η άριστη γραμμική αμερόληπτη εκτιμήτρια του b είναι η

$$\tilde{b}=(X'W^{-1}X)^{-1}X'W^{-1}y,$$

η αμερόληπτη εκτιμήτρια της διακύμανσης σ^2 είναι η

$$\tilde{\sigma}^2=\frac{\tilde{e}'W^{-1}\tilde{e}}{n-(k+1)}$$

και η αμερόληπτη εκτιμήτρια του πίνακα συνδιακυμάνσεων του $\tilde{b}$ είναι η

$$C(\tilde{b})=\tilde{\sigma}^2(X'W^{-1}X)^{-1}.$$

Θα εξετάσουμε τώρα τι είδους εκτιμήτριες θα πάρουμε αν στο υπόδειγμα (4.7.19) εφαρμόσουμε (από λάθος) τη μέθοδο Ε.Τ. αντί της μεθόδου A.G.L.S. Στην περίπτωση αυτή οι εκτιμήτριες Ε.Τ. του υποδείγματος θα είναι:

$$\hat{b}=(X'X)^{-1}X'y$$
$$=(X'X)^{-1}X'(Xb+e)$$
$$=b+(X'X)^{-1}X'e.$$

Άρα

$$E(\hat{b})=b$$ Αμεροληψία.

δηλαδή η μέθοδος Ε.Τ. όταν εφαρμοστεί στο υπόδειγμα (4.7.19) εξακολουθεί να δίνει αμερόληπτες εκτιμήτριες.

Ο αληθής πίνακας συνδιακυμάνσεων των εκτιμητριών Ε.Τ. είναι:

$$C(\hat{b})=E[(\hat{b}-b)(\hat{b}-b)']$$
$$=E[(X'X)^{-1}X'ee'X(X'X)^{-1}]$$
$$=\sigma^2(X'X)^{-1}X'WX(X'X)^{-1}, \quad \text{αφού} \quad E(ee')=\sigma^2W,$$

και αν $W=I$, τότε είναι ο γνωστός μας πίνακας $\sigma^2(X'X)^{-1}$. Αν όμως $W\neq I$, τότε έχουμε τα εξής προβλήματα:

(i) $$(X'X)^{-1}X'WX(X'X)^{-1}\neq(X'W^{-1}X)^{-1} \to \varepsilon(b)$$

συνεπώς οι εκτιμήτριες ελαχίστων τετραγώνων δεν είναι οι άριστες

(ii) ο τύπος $\sigma^2(X'X)^{-1}$ του πίνακα συνδιακυμάνσεων Ε.Τ. είναι μεροληπτική εκτιμήτρια του αληθούς πίνακα συνδιακυμάνσεων $\sigma^2(X'X)^{-1}X'WX(X'X)^{-1}$, και

(iii) η εκτιμήτρια Ε.Τ. της διακύμανσης

$$\hat{\sigma}^2=\frac{\hat{e}'\hat{e}}{n-(k+1)}$$

είναι μεροληπτική εκτιμήτρια της σ^2 (αποδεικνύεται ότι υποεκτιμά συστηματικά την σ^2).

Αποτέλεσμα των προβλημάτων αυτών είναι να μην ισχύει, εκτός από την αμεροληψία, κανένα από τα συμπεράσματα της στατιστικής επαγωγής του γραμμικού υποδείγματος στα οποία φτάσαμε χρησιμοποιόντας τη μέθοδο Ε.Τ. Μπορούμε εντούτοις να

αποδείξουμε (να αποδειχτεί ως άσκηση) ότι οι εκτιμήτριες Ε.Τ. εξακολουθούν να είναι συνεπείς εκτιμήτριες των αληθών παραμέτρων στον πληθυσμό. Είναι όμως μη αποτελεσματικές για μικρά αλλά και για μεγάλα δείγματα.

Εδώ συμπληρώνεται η γενική ανάλυση για την περίπτωση ενός γνωστού, θετικού πεπερασμένου και μη ιδιάζοντος πίνακα W. Στη συνέχεια θα εξετάσουμε χωριστά τις δύο ειδικές περιπτώσεις της ετεροσκεδαστικότητας και της αυτοσυσχέτισης.

4.8. ΕΤΕΡΟΣΚΕΔΑΣΤΙΚΟΤΗΤΑ κύριως σε διαστρωματικά στοιχεία.

Η ετεροσκεδαστικότητα των σφαλμάτων εμφανίζεται πολύ συχνά στις περιπτώσεις οικονομικών αναλύσεων με διαστρωματικά στοιχεία. Π.χ. στη διαμόρφωση της καταναλωτικής συμπεριφοράς ενός συνόλου οικογενειών, οι αποκλίσεις από τη μέση κατανάλωση που ορίζεται από την παλινδρόμηση, είναι μεγαλύτερες καθώς το εισόδημα αυξάνει διότι οι επιλογές στη διάθεση του εισοδήματος γίνονται ευρύτερες. Επίσης, οι επιχειρήσεις με μεγαλύτερο κύκλο εργασιών ή μεγαλύτερα κέρδη αναμένεται να παρουσιάζουν μεγαλύτερες διαφοροποιήσεις στη δραστηριότητά τους ως προς την παραγωγή, την πολιτική των μερισμάτων, τις επενδύσεις κλπ., από τις επιχειρήσεις με μικρότερο κύκλο εργασιών ή με μικρότερα κέρδη. Ετεροσκεδαστικότητα μπορεί όμως να παρουσιαστεί και στις διαχρονικές οικονομικές αναλύσεις και οι λόγοι είναι οι εξής: (i) διαχρονικά τόσο τα άτομα όσο και οι επιχειρήσεις αποκτούν καλλίτερη πληροφόρηση με αποτέλεσμα να διαπράττουν μικρότερα λάθη στη διεκπεραίωση των οικονομικών τους υποθέσεων (άρα οι διακυμάνσεις σ_i^2 της συμπεριφοράς είναι μικρότερες με την πάροδο του χρόνου) και (ii) καθώς οι τεχνικές της συλλογής και της επεξεργασίας των στατιστικών στοιχείων βελτιώνονται διαχρονικά οι διακυμάνσεις σ_i^2 αναμένεται να ελαττώνονται.

Ας θεωρήσουμε τη γενική περίπτωση του υποδείγματος:

$$y=Xb+e$$

$$E(e)=0$$

$$E(e_i^2)=\sigma_i^2=\sigma^2 w_{ii}, \quad i=1,2,\dots,n \qquad (4.8.1)$$

$$E(e_i e_j)=0, \qquad i,j=1,2,\dots,n, \quad i\neq j.$$

Στο υπόδειγμα αυτό έχουμε μόνο ετεροσκεδαστικότητα των σφαλμάτων και ο πίνακας W έχει τη μορφή:

$$W=\begin{bmatrix} w_{11} & & & 0 \\ & w_{22} & & \\ 0 & & \ddots & \\ & & & w_{nn} \end{bmatrix} \qquad (4.8.2)$$

όπου, χωρίς βλάβη της γενικότητας, μπορούμε να υποθέσουμε ότι $tr(W)=n$[1].

Στην περίπτωση αυτή ο μη ιδιάζων πίνακας K (τέτοιος ώστε $K'WK=I$ και $KK'=W^{-1}$) είναι ο

$$K=\begin{bmatrix} \frac{1}{\sqrt{w_{11}}} & & & 0 \\ & \frac{1}{\sqrt{w_{22}}} & & \\ 0 & & \ddots & \\ & & & \frac{1}{\sqrt{w_{nn}}} \end{bmatrix} \qquad (4.8.3)$$

και ο μετασχηματισμός

$$K'y=K'Xb+K'e$$

οδηγεί στο υπόδειγμα:

$$\frac{Y_i}{\sqrt{w_{ii}}}=\sum_{j=1}^{k}\frac{X_{ij}}{\sqrt{w_{ii}}}b_j+\frac{e_i}{\sqrt{w_{ii}}} \qquad (4.8.4)$$

---

1. Αν $tr(W)=\mu\neq n$, τότε θέτουμε $w_{ii}=\frac{\mu}{n}\left(\frac{nw_{ii}}{\mu}\right)=\frac{\mu}{n}\cdot\lambda_{ii}$ και η υπόθεση της ετεροσκεδαστικότητας γράφεται:

$$E(e_i^2)=\sigma_i^2=\sigma^2\left(\frac{\mu}{n}\lambda_{ii}\right)=\left(\sigma^2\frac{\mu}{n}\right)\lambda_{ii}=\sigma_*^2\lambda_{ii}, \quad \text{όπου} \quad \sum_{i=1}^{n}\lambda_{ii}=tr(W)=n$$

το οποίο θα εκτιμήσουμε με τη μέθοδο Ε.Τ., δηλαδή θα ελαχιστοποιήσουμε το άθροισμα

$$\sum_{i=1}^{n}\left(\frac{e_i}{\sqrt{w_{ii}}}\right)^2=\sum_{i=1}^{n}\frac{e_i^2}{w_{ii}} \qquad (4.8.5)$$

που είναι το άθροισμα των τετραγώνων των σφαλμάτων e_i σταθμισμένων με το αντίστοιχο στοιχείο του πίνακα W. Για το λόγο αυτό η μέθοδος AGLS ονομάζεται, στην περίπτωση της ετεροσκεδαστικότητας, και μέθοδος των "σταθμισμένων ελαχίστων τετραγώνων (Weighted Least Squares)".

Γενικά ο πίνακας W δεν είναι γνωστός και δεν μπορούμε να τον εκτιμήσουμε (δεν μπορούμε να εκτιμήσουμε τα n στοιχεία του W συν τα (k+1) στοιχεία του b από τις n παρατηρήσεις του δείγματος) παρά μόνο αν έχουμε ορισμένους περιορισμούς για τα στοιχεία του που θα ελαττώσουν τον αριθμό των παραμέτρων. Οι συνήθεις περιορισμοί είναι της μορφής

$$w_{ii}=f(X_{ij}) \qquad (4.8.6)$$

δηλαδή η διακύμανση των σφαλμάτων είναι συνάρτηση των τιμών κάποιας από τις ερμηνευτικές μεταβολές π.χ. της X_j, $j=1,2,\dots,k$. Μια απλή μορφή της (4.8.6) είναι η

$$w_{ii}=\sum_{r=\lambda}^{m}c_r X_j^r \qquad (4.8.7)$$

Αν (i) $\lambda=m=0$, τότε $w_{ii}=c_0=$σταθ. και τα σφάλματα είναι ομοσκεδαστικά (μέθοδος Ε.Τ.).

(ii) $\lambda=m$, τότε $w_{ii}=c_m X_j^m$ και η περίπτωση χαρακτηρίζεται ως "απλή ετεροσκεδαστικότητα". ΑΠΛΗ ΕΤΕΡΟ/ΣΚ

(iii) $\lambda\neq m$, τότε $w_{ii}=c_\lambda X_{ij}^{\lambda}+c_{\lambda+1}X_{ij}^{\lambda+1}+\dots+c_m X_{ij}^m$ και η περίπτωση χαρακτηρίζεται ως "σύνθετη ετεροσκεδαστικότητα". ΣΥΝΘΕΤΗ ΕΤΕ/ΣΚ

Οι συνηθέστερες μορφές της ετεροσκεδαστικότητας που χρησιμοποιούνται στην εφαρμοσμένη οικονομετρία είναι οι εξής:

(α) $\boxed{E(e_i^2)=k^2X_{ij}^2, \quad i=1,2,\dots,n} \qquad (4.8.8)$

Στην περίπτωση αυτή έχουμε $w_{ii}=X_{ij}^2$, $i=1,2,\dots,n$ και το προς εκτίμηση υπόδειγμα είναι το

Το j σταθερό
Άτακτη μεταβλητή.

$$\frac{Y_i}{X_{ij}}=b_0\frac{1}{X_{ij}}+b_1\frac{X_{i1}}{X_{ij}}+\dots+b_j+\dots+b_k\frac{X_{ik}}{X_{ij}}+\frac{e_i}{X_{ij}} \qquad (4.8.9)$$

$i=1,2,\dots,n$,

και είναι προφανές ότι

$$E\left(\frac{e_i}{X_{ij}}\right)=0, \quad i=1,2,\dots,n$$

$$E\left(\frac{e_i^2}{X_{ij}^2}\right)=\frac{k^2X_{ij}^2}{X_{ij}^2}=k^2=\sigma^2, \text{ σταθερή που θα εκτιμηθεί}$$

αμερόληπτα από την $\quad \tilde{\sigma}^2=\dfrac{\tilde{e}'W^{-1}\tilde{e}}{n-(k+1)}$, και

$$E\left(\frac{e_ie_k}{X_{ij}X_{kj}}\right)=0, \quad i\neq k.$$

Ακόμα πρέπει να παρατηρήσουμε ότι ο σταθερός όρος του αρχικού υποδείγματος δεν συμπίπτει με το σταθερό όρο του υποδείγματος (4.8.9) που είναι ο συντελεστής της νέας μεταβλητής $1/X_{ij}$.

(β) $\boxed{E(e_i^2)=k^2X_{ij}, \quad i=1,2,\dots,n.}$ (4.8.10)

Εδώ έχουμε $w_{ii}=X_i$, $i=1,2,\dots,n$ και το προς εκτίμηση υπόδειγμα είναι:

$$\frac{Y_i}{\sqrt{X_{ij}}}=b_0\frac{1}{\sqrt{X_{ij}}}+b_1\frac{X_{i1}}{\sqrt{X_{ij}}}+\dots+b_j\frac{X_{ij}}{\sqrt{X_{ij}}}+\dots+b_k\frac{X_{ik}}{\sqrt{X_{ij}}}+\frac{e_i}{\sqrt{X_{ij}}} \qquad (4.8.11)$$

όπου και πάλι $k^2=\sigma^2$.

(γ) $\boxed{E(e_i^2)=k^2f(X_{ij}),}$ όπου $f(X_{ij})$ είναι γνωστή συνάρτηση της X_j. (4.8.12)

Εδώ $w_{ii}=f(X_{ij})$, $i=1,2,\dots,n$ και το προς εκτίμηση υπόδειγμα είναι το

$$\frac{Y_i}{\sqrt{f(X_{ij})}}=b_0\frac{1}{\sqrt{f(X_{ij})}}+b_1\frac{X_{i1}}{\sqrt{f(X_{ij})}}+\dots+b_k\frac{X_{ik}}{\sqrt{f(X_{ij})}}+\frac{e_i}{\sqrt{f(X_{ij})}}$$

$$k^2=\sigma^2. \qquad (4.8.13)$$

(δ) $\boxed{E(e_i^2)=k^2[E(Y_i)]^2, \quad i=1,2,\dots,n.}$ (4.8.14)

Σύμφωνα με την υπόθεση αυτή η διακύμανση των σφαλμάτων είναι ανάλογη προς το τετράγωνο του μέσου των Y_i, με συντελεστή αναλογίας τη σταθερή $k^2=\sigma^2$. Αλλά

$$E(Y_i)=b_0+b_1X_{i1}+\dots+b_kX_{ik}$$

Αν λοιπόν χρησιμοποιήσουμε το μετασχηματισμό

$$\frac{Y_i}{E(Y_i)}=b_0\frac{1}{E(Y_i)}+b_1\frac{X_{i1}}{E(Y_i)}+\dots+b_k\frac{X_{ik}}{E(Y_i)}+\frac{e_i}{E(Y_i)}$$

τότε τα σφάλματα της νέας παλινδρόμησης είναι ομοσκεδαστικά και μπορούμε να χρησιμοποιήσουμε τη μέθοδο Ε.Τ. Είναι φανερό ότι ο μετασχηματισμός αυτός δεν είναι πρακτικά εφαρμόσιμος αφού ο μέσος $E(Y_i)$ εξαρτάται από τις άγνωστες παραμέτρους $b_0, b_1, \dots, b_k$. Για το λόγο αυτό μπορούμε να ακολουθήσουμε τα εξής βήματα:

(i) αγνοώντας την ύπαρξη της ετεροσκεδαστικότητας εκτιμούμε με τη μέθοδο Ε.Τ. την παλινδρόμηση $\hat{Y}_i=\hat{b}_0+\hat{b}_1X_{i1}+\dots+\hat{b}_kX_{ik}$ και

(ii) εκτελούμε το μετασχηματισμό

$$\frac{Y_i}{\hat{Y}_i}=b_0\frac{1}{\hat{Y}_i}+b_1\frac{X_{i1}}{\hat{Y}_i}+\dots+\frac{X_{ik}}{\hat{Y}_i}+\frac{e_i}{\hat{Y}_i}. \qquad (4.8.15)$$

Αν και οι $\hat{Y}_i$ δεν είναι ίσες προς τις $E(Y_i)$ είναι όμως εκτιμήτριες, δηλαδή καθώς το μέγεθος του δείγματος αυξάνει συγκλίνουν προς τις αληθείς τιμές $E(Y_i)$. Άρα ο μετασχηματισμός είναι πρακτικά εφαρμόσιμος αν το μέγεθος του δείγματος είναι αρκετά μεγάλο.

Το ερώτημα που τίθεται στη συνέχεια είναι το πως θα ελέγξουμε αν υπάρχει ή όχι ετεροσκεδαστικότητα σε κάθε συγκεκριμένη εφαρμογή και ποια είναι η μορφή της. Από την αρχή πρέπει να γνωρίζουμε ότι δεν υπάρχουν απόλυτα και άμεσα κριτήρια για την ύπαρξη της ετεροσκεδαστικότητας. Αυτό είναι φανερό αφού οι πραγματικές διακυμάνσεις σ_i^2 των σφαλμάτων είναι άγνωστες και θα παραμείνουν άγνωστες αν δεν διαθέτουμε ολόκληρους τους πληθυσμούς της Y που αντιστοιχούν στις δεδομένες τιμές των μεταβλητών $X_1, X_2, \ldots, X_k$, και αυτό δεν συμβαίνει στις οικονομετρικές αναλύσεις. Παρόλα αυτά υπήρξαν αρκετές "εμπειρικές" προσπάθειες για την επισήμανση της ετεροσκεδαστικότητας. Αναφέρουμε μερικές από τις μεθόδους που εμφανίζονται συχνότερα στα εγχειρίδια της οικονομετρίας:

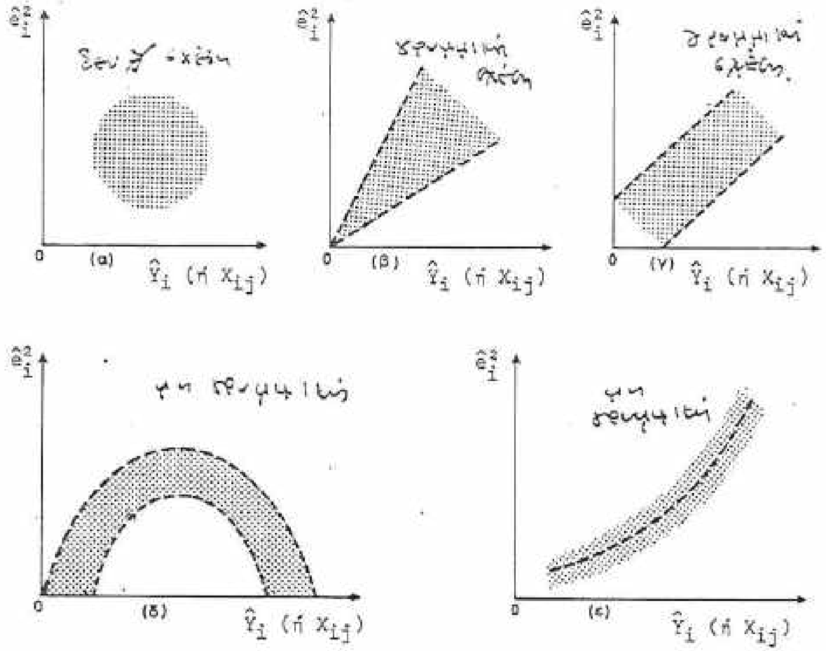
i) Η εξέταση της φύσης του προβλήματος

Πολύ συχνά η φύση του συγκεκριμένου οικονομικού φαινομένου που προσπαθούμε να ερμηνεύσουμε μπορεί να μας υποδείξει την ύπαρξη ή όχι ετεροσκεδαστικότητας. Π.χ., σύμφωνα με τα αποτελέσματα της πρωτοποριακής έρευνας των Prais-Houthaker[1] για την ανάλυση των οικογενειακών προϋπολογισμών, η διακύμανση των καταλοίπων της παλινδρόμησης της κατανάλωσης πάνω στο εισόδημα αυξάνει με το εισόδημα. Και είναι τώρα γενικά αποδεκτό ότι σε παρόμοιες έρευνες το φαινόμενο της ετεροσκεδαστικότητας είναι ο κανόνας. Γενικά, πρέπει να παρατηρήσουμε ότι η παρουσία της ετεροσκεδαστικότητας είναι συνήθης σε κάθε διαστρωματική ανάλυση που αναφέρεται σε ετερογενείς μονάδες. Π.χ. στη διαστρωματική ανάλυση των επενδύσεων πάνω στις πωλήσεις, πρέπει να αναμένουμε την ύπαρξη ετεροσκεδαστικότητας όταν το δείγμα περιέχει μικρές, μεσαίες και μεγάλες επιχειρήσεις.

1. S.J. Prais and H.S. Houthakker, *The analysis of Family Budgets*, Cambridge University Press, New York, 1955.

ii) Η εξέταση του διαγράμματος των καταλοίπων της παλινδρόμησης.

Αν δεν γνωρίζουμε τη μορφή της ετεροσκεδαστικότητας, τότε μπορούμε να εκτιμήσουμε την παλινδρόμηση κάτω από την υπόθεση της ομοσκεδαστικότητας και, στη συνέχεια, να εξετάσουμε τη συμπεριφορά των τετραγώνων των καταλοίπων $\hat{e}_i^2$. Βέβαια, τα τετράγωνα των καταλοίπων $\hat{e}_i^2$ δεν είναι τα ίδια με τα τετράγωνα e_i^2 των σφαλμάτων, μπορούμε όμως να τα χρησιμοποιήσουμε ως προσεγγίσεις των e_i^2 αν το δείγμα είναι αρκετά μεγάλο.



Σχήμα 4-6: Διαγράμματα των καταλοίπων $\hat{e}_i^2$ συναρτήσει των τιμών $\hat{Y}_i$ (ή των τιμών της X_{ij}) της εξαρτημένης μεταβλητής

Στο σχήμα 4-6 δίνονται μερικά υποθετικά διαγράμματα των $\hat{e}_i^2$ συναρτήσει των εκτιμήσεων $\hat{Y}_i$ των τιμών της εξαρτημένης μεταβλητής από την παλινδρόμηση. Στην περίπτωση (α) δεν υπάρχει συστηματική σχέση ανάμεσα στις μεταβλητές $\hat{e}_i^2$ και $\hat{Y}_i$ ενώ στις υπόλοιπες περιπτώσεις η ύπαρξη συστηματικής σχέσης είναι προφανής.

Στην περίπτωση (γ) η σχέση είναι γραμμική, ενώ στις περιπτώσεις (δ) και (ε) είναι μη γραμμική. Διαθέτοντας τις πληροφορίες αυτές μπορούμε να μετασχηματίσουμε τα αρχικά στοιχεία έτσι ώστε να απαλείψουμε την ετεροσκεδαστικότητα.

Αντί του διαγράμματος των $\hat{e}_i^2$ και $\hat{Y}_i$ μπορούμε να εξετάσουμε το διάγραμμα των $\hat{e}_i^2$ συναρτήσει των τιμών X_{ij} μιας οποιασδήποτε ερμηνευτικής μεταβλητής $(j=1,2,\ldots,k)$, αν πιστεύουμε ότι η διακύμανση των σφαλμάτων είναι συνάρτηση των τιμών της X_j.

Τις πληροφορίες από τα διαγράμματα μπορούμε να τις ποσοτικοποιήσουμε χρησιμοποιώντας τη γενική εξίσωση[1]:

$$\boxed{\sigma_i^2=\sigma^2 X_{ij}^b e^{u_i}} \qquad (4.8.16)$$

ή

$$\boxed{\ln\sigma_i^2=\ln\sigma^2+b\ln X_{ij}+u_i.} \rightarrow \text{ουσ.} \qquad (4.8.17)$$

Επειδή οι διακυμάνσεις σ_i^2 δεν είναι γνωστές, μπορούμε να τις προσεγγίσουμε με τα τετράγωνα των καταλοίπων της παλινδρόμησης και να εκτιμήσουμε την

$$\ln\hat{e}_i^2=\widehat{\ln\sigma^2}+\hat{b}\ln X_{ij}+u_i. \qquad (4.8.18)$$

Αν η εκτιμήτρια b είναι στατιστικά σημαντική, τότε δεχόμαστε την ύπαρξη ετεροσκεδαστικότητας στο δείγμα μας ενώ στην αντίθετη περίπτωση δεχόμαστε την υπόθεση της ομοσκεδαστικότητας. Αν και ο εμπειρικός αυτός έλεγχος παρουσιάζει αρκετό ενδιαφέρον εντούτοις οι Goldfeld και Quandt[2] επισημαίνουν τον κίνδυνο τα σφάλματα της εξίσωσης (4.8.18) να μην ικανοποιούν τις γνωστές υποθέσεις για την εφαρμογή της μεθόδου των Ε.Τ.

1. R.E. Park, "Estimation with Heteroskedastic Error Terms", *Econometrica*, Vol. 34 (1966), p. 888.

2. S.M. Goldfeld and R.E. Quandt, *Nonlinear Methods in Econometrics*, North-Holland, Amsterdam, 1972, pp. 93-94.

ii) Οι έλεγχοι του Glejser[1].

Οι έλεγχοι του Glejser είναι σχετικοί με τον προηγούμενο και συνεπώς παρουσιάζουν τα ίδια προβλήματα. Και εδώ υπολογίζονται με τη μέθοδο Ε.Τ. τα κατάλοιπα $\hat{e}_i$ της παλινδρόμησης

$$Y_i=\hat{b}_0+\hat{b}_1X_{i1}+\cdots+\hat{b}_kX_{ik}+\hat{e}_i$$

και χρησιμοποιούνται για την εκτίμηση διαφόρων παλινδρομήσεων με σκοπό τον προσδιορισμό της συγκεκριμένης μορφής της ετεροσκεδαστικότητας:

$$|\hat{e}_i|=b_1X_{ij}+u_i$$
$$|\hat{e}_i|=b_1\sqrt{X_{ij}}+u_i$$
$$|\hat{e}_i|=b_1\frac{1}{\sqrt{X_{ij}}}+u_i$$
$$|\hat{e}_i|=b_1\frac{1}{X_{ij}}+u_i$$
$$|\hat{e}_i|=b_0+b_1X_{ij}+u_i$$
$$|\hat{e}_i|=\sqrt{b_0+b_1X_{ij}}+u_i$$
$$|\hat{e}_i|=\sqrt{b_0+b_1X_{ij}^2}+u_i \quad \text{κ.λ.π.}$$

Μια πρόσθετη δυσκολία στην προσέγγιση του Glejser είναι ότι τα σφάλματα u_i δεν έχουν μέσο μηδέν ενώ παρουσιάζουν αυτοσυσχέτιση και ετεροσκεδαστικότητα[2]. Επίσης η εκτίμηση

1. H. Glejser, "A New Test for Heteroskedasticity" *Journal of the American Statistical Association*, Vol. 64 (1969), pp. 316-323.

2. S.M. Goldfeld and R.E. Quand, *Nonlinear Methods in Econometrics* ο.π. Κεφ. 3.

των δύο τελευταίων παλινδρομήσεων θα γίνει αριθμητικά γιατί δεν είναι γραμμικές. Ο Glejser διαπίστωσε ότι, για μεγάλα δείγματα, οι πρώτες τέσσερις παλινδρομήσεις δίνουν, γενικά, ικανοποιητικά αποτελέσματα στόν προσδιορισμό της μορφής της ετεροσκεδαστικότητας. Επομένως οι έλεγχοι του Glejser μπορούν να χρησιμοποιηθούν σε μεγάλα δείγματα ενώ σε μικρά δείγματα η χρησιμότητά τους είναι μόνο ενδεικτική για τη μορφή της ετεροσκεδαστικότητας.

iv) Ο έλεγχος των Goldfeld-Quandt[1].

Ο έλεγχος αυτός αναφέρεται στο γενικό γραμμικό υπόδειγμα

$$y = Xb + e$$

υποθέτοντας ότι τα σφάλματα είναι ανεξάρτητα και κατανέμονται κανονικά και ότι το μέγεθος του δείγματος είναι τουλάχιστον διπλάσιο από τον αριθμό των προς εκτίμηση παραμέτρων. Η υπόθεση η οποία ελέγχεται είναι η

H_0: τα σφάλματα είναι ομοσκεδαστικά

κατά της

H_1: τα σφάλματα παρουσιάζουν ετεροσκεδαστικότητα της μορφής $E(e_i^2) = \sigma^2 X_{ij}^2$, όταν X_j είναι κάποια από τις μεταβλητές $X_1, X_2, \dots, X_k$.

Τα βήματα για τη διεξαγωγή του ελέγχου είναι τα ακόλουθα:

1. Διατάσσουμε τις παρατηρήσεις του δείγματος και τις αύξουσες τιμές της μεταβλητής X_j.

2. Παραλείπουμε c το πλήθος κεντρικές παρατηρήσεις (με-

1. S.M. Goldfeld and R.E. Quandt, "Some Tests for Heteroskedasticity", *Journal of the American Statistical Association*, Vol. 60 (1965), pp. 539-47.

τά από σχετικούς πειραματισμούς οι συγγραφείς υποδεικνύουν ότι ο αριθμός c πρέπει να είναι ίσος περίπου προς το ένα τέταρτο του δείγματος).

3. Προσαρμόζουμε με τη μέθοδο Ε.Τ. χωριστές παλινδρομήσεις στα δύο δείγματα $\left(\text{μεγέθους } \frac{n-c}{2}\right)$ που απομένουν μετά την αφαίρεση των c κεντρικών παρατηρήσεων.

4. Αν $\Sigma\hat{e}_{i1}^2$ και $\Sigma\hat{e}_{i2}^2$ είναι το άθροισμα των τετραγώνων των καταλοίπων από τις δύο παλινδρομήσεις, τότε η στατιστική

$$F^* = \frac{\Sigma\hat{e}_{i2}^2 \Big/ \left[\frac{n-c}{2} - (k+1)\right]}{\Sigma\hat{e}_{i1}^2 \Big/ \left[\frac{n-c}{2} - (k+1)\right]} = \frac{\Sigma\hat{e}_{i2}^2}{\Sigma\hat{e}_{i1}^2}$$

ακολουθεί, κάτω από την υπόθεση της ομοσκεδαστικότητας, την κατανομή $F_{\frac{n-c}{2}-(k+1),\ \frac{n-c}{2}-(k+1)}$ και θα τείνει προς τη μονάδα ενώ αν τα σφάλματα είναι ετεροσκεδαστικά (με αυξανόμενη διακύμανση) τότε θα πέρνει μεγαλύτερες τιμές. Έτσι αν $F^* > F_{0,05}$ δεχόμαστε ότι υπάρχει ετεροσκεδαστικότητα της μορφής που αναφέραμε ενώ αν $F^* < F_{0,05}$ δεχόμαστε ότι τα σφάλματα είναι ομοσκεδαστικά σε επίπεδο σημαντικότητας 5%.

v) Ο συντελεστής συσχέτισης τάξεων του Spearman.

Ο έλεγχος αυτός είναι απλός, μπορεί να χρησιμοποιηθεί σε μεγάλα αλλά και σε μικρά δείγματα και βασίζεται στον υπολογισμό του συντελεστή συσχέτισης της τάξης των τιμών των καταλοίπων $\hat{e}_i$ της παλινδρόμησης:

$$Y_i = \hat{b}_0 + \hat{b}_1 X_{i1} + \hat{b} X_{i2} + \dots + \hat{b}_k X_{ik} + \hat{e}_i$$

και των τιμών της μεταβλητής X_j (αν η διακύμανση των σφαλμάτων πιστεύεται ότι μεταβάλλεται με τις τιμές της X_j). Τα συγκεκριμένα βήματα για τη διεξαγωγή του ελέγχου είναι τα ακόλουθα:

(1.) Εκτιμούμε την παλινδρόμηση με τη μέθοδο Ε.Τ. και υπολογίζουμε τα κατάλοιπα $\hat{e}_i$.

(2.) Διατάσσουμε τα κατάλοιπα $\hat{e}_i$ (αγνοώντας το πρόσημό τους) και τις τιμές της X_j κατά τάξη μεγέθους (αυξανόμενου ή φθίνοντος) και υπολογίζουμε το συντελεστή συσχέτισης των διατάξεων αυτών κατά Spearman:

$$r_s = 1 - \frac{\sigma \sum_{i=1}^{n} d_i^2}{n^3 - n}$$

Στατιστική Ι σελ. 191

όπου d_i είναι η διαφορά των τάξεων των τιμών των $\hat{e}_i$ και X_{ij} στη νέα διάταξη και n είναι το μέγεθος του δείγματος. Κάτω από την υπόθεση ότι ο συντελεστής του Spearman στον πληθυσμό είναι 0, η στατιστική

$$t^* = \frac{r_s\sqrt{n-2}}{\sqrt{1-r_s^2}}$$

ακολουθεί την κατανομή t με (n-2) βαθμούς ελευθερίας. Έτσι, αν $t^* > t_{0,05}$ των πινάκων δεχόμαστε την υπόθεση της ετεροσκεδαστικότητας. Ο έλεγχος αυτός μπορεί να γίνει για όλες τις μεταβλητές $X_1, X_2, \ldots, X_k$ έτσι ώστε να εντοπίσουμε την ή τις μεταβλητές στις οποίες οφείλεται η ετεροσκεδαστικότητα.

Γενικά μπορούμε να πούμε ότι είναι σκόπιμο να προηγείται ο έλεγχος κατά Spearman ή ο έλεγχος των Goldfeld-Quandt για να διαπιστωθεί αν υπάρχει ετεροσκεδαστικότητα και μετά να προχωρήσουμε σε παλινδρομήσεις τύπου Glejser ή Park για τον προσδιορισμό της μορφής της ετεροσκεδαστικότητας.

4.9. ΑΥΤΟΣΥΣΧΕΤΙΣΗ (κυρίως σε χρονολογικές σειρές)

Στα πλαίσια του γενικού γραμμικού υποδείγματος όταν αναφερόμαστε στον όρο "αυτοσυσχέτιση" εννοούμε ότι δεν ισχύει η βασική υπόθεση (υ.3α) της ανεξαρτησίας των σφαλμάτων· δηλαδή

$$E(e_i e_j) \neq 0, \quad i,j = 1,2,\ldots,n \ (i \neq j).$$

Το φαινόμενο της αυτοσυσχέτισης παρουσιάζεται κυρίως στην εκτίμηση διαχρονικών ποσοτικών σχέσεων, δηλαδή όταν τα στοιχεία που χρησιμοποιούμε είναι χρονολογικές σειρές, και σημαίνει ότι το σφάλμα μιας περιόδου εξαρτάται από τα σφάλματα μιας ή περισσοτέρων από τις προηγούμενες χρονικές περιόδους.

Οι μορφές που μπορεί να πάρει η αυτοσυσχέτιση είναι οι εξής:

(i)

$$e_t = \rho_1 e_{t-1} + u_t$$
$$e_t = \rho_1 e_{t-1} + \rho_2 e_{t-2} + u_t \qquad (4.9.1)$$
$$\cdots\cdots\cdots\cdots$$
$$e_t = \rho_1 e_{t-1} + \rho_2 e_{t-2} + \cdots + \rho_\theta e_{t-\theta} + u_t$$

Αυτοπαλίνδρομα Σχήματα AR(θ)

όπου οι συντελεστές $\rho_1, \rho_2, \ldots, \rho_\theta$ δεν είναι απαραίτητα όλοι διάφοροι από το μηδέν.

Οι μορφές αυτές της αυτοσυσχέτισης ονομάζονται "αυτοπαλίνδρομα σχήματα" πρώτης, δεύτερης, ..., θ τάξης, αντίστοιχα, και το σφάλμα e_t της περιόδου t εξαρτάται γραμμικά από τα σφάλματα των περιόδων t-1, t-2, ..., t-θ και από το τυχαίο σφάλμα u_t της περιόδου t. Στην αγγλική βιβλιογραφία χρησιμοποιείται ο όρος "autoregressive processes" και ο γενικός συμβολισμός AR(θ), θ=1,2,... Τα τυχαία σφάλματα u_t θεωρούμε ότι έχουν τις γνωστές ιδιότητες:

$$E(u) = 0 \quad \text{και} \quad C(u) = \sigma_u^2 I \qquad (4.9.2)$$

(ii)

$$e_t = u_t + \alpha_1 u_{t-1}$$
$$e_t = u_t + \alpha_1 u_{t-1} + \alpha_2 u_{t-2} \qquad (4.9.3)$$
$$\cdots\cdots\cdots\cdots$$
$$e_t = u_t + \alpha_1 u_{t-1} + \alpha_2 u_{t-2} + \cdots + \alpha_\ell u_{t-\ell}$$

Σχήματα κινητού μέσου. MA(ℓ)

όπου οι συντελεστές $\alpha_1, \alpha_2, \ldots, \alpha_\ell$ δεν είναι απαραίτητα όλοι διάφοροι από το μηδέν.

Οι μορφές αυτές της αυτοσυσχέτισης ονομάζονται "σχήματα κινητού μέσου" πρώτης, δεύτερης, ..., ℓ τάξης, αντίστοιχα, και το σφάλμα e_t της περιόδου t είναι ένας γραμμικός συνδυασμός των τυχαίων σφαλμάτων $u_t, u_{t-1}, \dots, u_{t-\ell}$ της ίδιας περιόδου και των προηγουμένων περιόδων. Ο αγγλικός όρος για τα σχήματα αυτά είναι "moving average processes" και χρησιμοποιείται ο γενικός συμβολισμός MA(ℓ), ℓ=1,2,...

ARMA(θ,ℓ)

(iii) $e_t=\rho_1 e_{t-1}+\dots+\rho_\theta e_{t-\theta}+u_t+\alpha_1 u_{t-1}+\dots+\alpha_\ell u_{t-\ell}$ (4.9.4)

Τα σχήματα που αντιπροσωπεύονται από τη γενική αυτή μορφή είναι μεικτά σχήματα: αυτοπαλίνδρομα και κινητού μέσου και ο αντίστοιχος διεθνής συμβολισμός τους είναι ARMA(θ,ℓ).

Οι αιτίες στις οποίες μπορεί να οφείλεται η ύπαρξη της αυτοσυσχέτισης είναι πολλές. Αναφέρουμε παρακάτω μερικές από αυτές:

ΑΙΤΙΕΣ ΑΥΤΟΣΥΣΧΕΤΙΣΗΣ

1. Η ύπαρξη διαχρονικών τάσεων στις χρονολογικές σειρές.

Όπως γνωρίζουμε, πολλές χρονολογικές σειρές, όπως η ακαθάριστη εθνική παραγωγή, η απασχόληση, η κατανάλωση, οι δείκτες τιμών κλπ., παρουσιάζουν, σε περιόδους ανάκαμψης ή ύφεσης της οικονομίας, έντονη συσχέτιση των τιμών τους στις διαδοχικές χρονικές περιόδους. Έτσι, σε περίοδο ανάκαμψης η κατανάλωση στην περίοδο t+1 είναι πάντοτε μεγαλύτερη από την κατανάλωση στην περίοδο t. Συνεπώς, στην εκτίμηση παλινδρομήσεων με χρονολογικές σειρές η παρουσία της αυτοσυσχέτισης είναι συνηθισμένο φαινόμενο.

2. Παράλειψη σημαντικών ερμηνευτικών μεταβλητών.

Αν από την αληθή παλινδρόμηση

$$Y_t=b_0+b_1X_{t1}+b_2X_{t2}+e_t$$

παραλείψουμε για κάποιο λόγο τη σημαντική ερμηνευτική μεταβλητή X_2 και εκτιμήσουμε την παλινδρόμηση

$$Y_t=\alpha_0+\alpha_1X_{t1}+\upsilon_t$$

τότε τα σφάλματα υ_t της παλινδρόμησης αυτής θα είναι:

$$\upsilon_t=b_2X_{t2}+e_t$$

και θα παρουσιάζουν αυτοσυσχέτιση στο βαθμό που οι τιμές της X_2 αυτοσυσχετίζονται και στο βαθμό που η X_2 επηρεάζει την Y.

3. Εσφαλμένος προσδιορισμός της συναρτησιακής μορφής του υποδείγματος

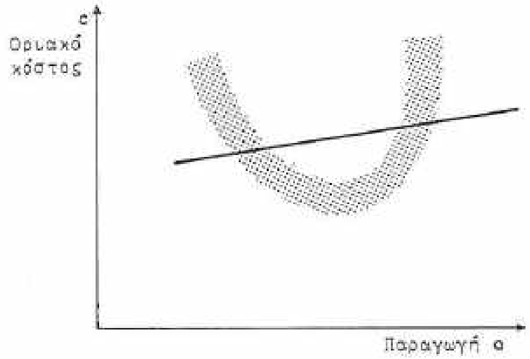
Ας υποθέσουμε ότι η αληθής παλινδρόμηση είναι μη γραμμική (οριακό κόστος συναρτήσει της παραγωγής):

$$C_i=b_0+b_1Q_i+b_2Q_i^2+e_i$$

και ότι από λάθος εκτιμούμε τη γραμμική παλινδρόμηση

$$C_i=\alpha_0+\alpha_1Q_i+\upsilon_i.$$

Στην περίπτωση αυτή οι πραγματικές τιμές του κόστους C_t που βρίσκονται πάνω από τη γραμμική παλινδρόμηση του σχήματος 4.7 θα υποεκτιμόνται συστηματικά ενώ αυτές που βρίσκονται κάτω από τη γραμμική παλινδρόμηση θα υπερεκτιμόνται επίσης συστηματικά. Άρα στα κατάλοιπα της γραμμικής παλινδρόμησης θα έχουμε αυτοσυσχέτιση και αυτό θα οφείλεται βέβαια στο ότι τα σφάλματα της γραμμικής παλινδρόμησης δεν εί-



Σχήμα 4.7: Εσφαλμένος προσδιορισμός της συναρτησιακής μορφής του υποδείγματος.

ναι τυχαία αλλά έχουν ενσωματωμένη και την επίδραση του δευτεροβάθμιου όρου Q_i^2.

4. Φαινόμενα Cobweb. (Στα της αρχικής)

Σε πολλές αγροτικές καλλιέργειες η προσφορά S στην περίοδο t προσδιορίζεται συχνά από την τιμή P στην προηγούμενη χρονική περίοδο:

$$S_t = b_0 + b_1 P_{t-1} + e_t$$

Αν η τιμή στο τέλος της περιόδου t είναι μικρότερη από την τιμή της περιόδου t-1 τότε, είναι πιθανό, οι καλλιεργητές να αποφασίσουν να ελαττώσουν την παραγωγή τους κατά την επόμενη περίοδο t+1. Σε μια τέτοια κατάσταση, είναι φανερό ότι, τα σφάλματα στις διάφορες χρονικές περιόδους θα συσχετίζονται αρνητικά: αν υπάρχει υπερπαραγωγή στην περίοδο t θα υπάρχει μειωμένη παραγωγή στην περίοδο t+1 κ.ο.κ., οδηγόντας έτσι στο γνωστό φαινόμενο Cobweb.

5. Ύπαρξη χρονικών υστερήσεων.

Στην εκτίμηση της κατανάλωσης συναρτήσει του εισοδήματος χρησιμοποιείται συχνά η εξίσωση

$$C_t = b_0 + b_1 Y_t + b_2 C_{t-1} + e_t$$

στην οποία η μεταβλητή C_{t-1} ερμηνεύει τις διαμορφωμένες καταναλωτικές συνήθειες. Αν λοιπόν παραλείψουμε από την εξίσωση την C_{t-1} τότε τα σφάλματα της ελλειπούς παλινδρόμησης δεν θα είναι τυχαία αλλά οι τιμές τους θα επηρεάζονται συστηματικά από τις τιμές της κατανάλωσης στην προηγούμενη χρονική περίοδο.

6. Επεμβάσεις στις χρονολογικές σειρές.

Συχνά στην ποσοτική ανάλυση γίνονται επεμβάσεις στα αρχικά στατιστικά στοιχεία. Π.χ. αν οι αρχικές παρατηρήσεις είναι μηνιαίες και, για διάφορους λόγους, επιθυμούμε να χρησιμοποιήσουμε τριμηνιαίες παρατηρήσεις για την εκτίμηση μιας παλινδρόμησης, τότε η συνήθης πρακτική που ακολουθείται είναι να πέρνουμε το μέσο όρο των τριών μηνιαίων παρατηρήσεων του τριμήνου. Σε τέτοιου είδους επεμβάσεις οι χρονολογικές σειρές γίνονται πιο "ομαλές" με αποτέλεσμα η συμπεριφορά τους να παρουσιάζει εντονότερη αυτοσυσχέτιση. Ακόμα πιο σοβαρά προβλήματα αυτοσυσχέτισης προκύπτουν όταν γίνονται "παρεμβολές" στις χρονολογικές σειρές. Π.χ. η απογραφή του πληθυσμού γίνεται κάθε δέκα χρόνια ενώ ο πληθυσμός στα ενδιάμεσα χρόνια υπολογίζεται -κάτω από ορισμένες υποθέσεις- με τη μέθοδο της παρεμβολής, δηλαδή με κάποιο συστηματικό τρόπο που συνδέει τον πληθυσμό ενός έτους με τον πληθυσμό του επόμενου έτους. Αυτή ακριβώς η μέθοδος εισάγει συστηματική συμπεριφορά, δηλαδή έντονη αυτοσυσχέτιση στη χρονολογική σειρά του πληθυσμού.

ΑΥΤΟΠΑΛΙΝΔΡΟΜΟ ΣΧΗΜΑ 1ης ΤΑΞΗΣ.

Στη συνέχεια θα ασχοληθούμε αποκλειστικά με την περίπτωση του αυτοπαλίνδρομου σχήματος πρώτης τάξης που είναι και το συνηθέστερο είδος αυτοσυσχέτισης που αντιμετωπίζουμε στην εμπειρική ανάλυση.

Ας θεωρήσουμε το γενικό γραμμικό υπόδειγμα:

$$Y_t = b_0 + b_1 X_{t1} + b_2 X_{t2} + \cdots + b_k X_{tk} + e_t \qquad (4.9.5)$$

όπου

$$e_t = \rho e_{t-1} + u_t, \quad |\rho| < 1 \qquad (4.9.6)$$

και

$$E(u_t) = 0, \quad \forall t,$$

$$E(u_t u_s) = \delta_{ts}\sigma_u^2, \quad \delta_{ts} = \begin{cases} 1, & t=s \\ 0, & t \neq s, \end{cases} \quad \forall t,s. \qquad (4.9.7)$$

Θα υπολογίσουμε τους μέσους, τις διακυμάνσεις και τις συνδιακυμάνσεις των σφαλμάτων e_t που ακολουθούν το αυτοπαλίνδρομο σχήμα πρώτης τάξης.

Με συνεχείς αντικαταστάσεις στην (4.9.6), εύκολα προκύπτει ότι:

$$e_t=\rho e_{t-1}+u_t$$

$$=\rho(\rho e_{t-2}+u_{t-1})+u_t=\rho^2 e_{t-2}+\rho u_{t-1}+u_t$$

$$=\rho^2(\rho e_{t-3}+u_{t-2})+\rho u_{t-1}+u_t=\rho^3 e_{t-3}+\rho^2 u_{t-2}+\rho u_{t-1}+u_t$$

. . . .

$$=\rho^s e_{t-s}+\sum_{j=0}^{s-1}\rho^j u_{t-j}\,. \qquad (4.9.9)$$

Επειδή το ρ είναι δεκαδικός

Όπως είναι γνωστό $\lim \rho^s e_{t-s}=0$ και η σειρά $\sum_{j=0}^{\infty}\rho^j u_{t-j}$ συγκλίνει προς ένα πεπερασμένο αριθμό, τότε και μόνο τότε αν, $|\rho|<1$. Αυτός είναι και ο λόγος για τον οποίο στην(4.8.6) έχουμε υποθέσει ότι $|\rho|<1$ έτσι ώστε τα σφάλματα e_t να έχουν πεπερασμένη τιμή. Αλλοιώς, αν $|\rho|>1$ τότε, το αυτοπαλίνδρομο σχήμα θα ακολουθούσε εκρηκτική (μη ελεγχόμενη) συμπεριφορά κάτι που δεν δικαιολογείται για τα κατάλοιπα μιας παλινδρόμησης με οικονομικές χρονολογικές σειρές.

Άρα μετά από άπειρες αντικαταστάσεις $(s\to\infty)$ η (4.8.9) παίρνει τη μορφή:

$$e_t=\sum_{j=0}^{\infty}\rho^j u_{t-j}\,,\quad t=1,2,\ldots,n. \qquad (4.9.10)$$

Εύκολα τώρα προκύπτουν τα εξής:

$$E(e_t)=\sum_{j=0}^{\infty}\rho^j E(u_{t-j})=0,\quad t=1,2,\ldots,n, \qquad (4.9.11)$$

$$E(e_t^2)=E\Big(\sum_{j=0}^{\infty}\rho^j u_{t-j}\Big)^2$$

$$=E\Big(\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\rho^{j+k}u_{t-j}u_{t-k}\Big)$$

$$=\sum_{j=0}^{\infty}\sum_{k=0}^{\infty}\rho^{j+k}E(u_{t-j}u_{t-k})$$

$$=\sigma_u^2\sum_{j=0}^{\infty}\rho^{2j},\ [\text{λόγω της } (4.8.8)]$$

$$=\frac{\sigma_u^2}{1-\rho^2}\quad,\Big[\text{διότι}\ \sum_{j=0}^{\infty}(\rho^2)^j=\frac{1}{1-\rho^2}\Big] \qquad (4.9.12)$$

$$t=1,2,\ldots,n\,.$$

$$E(e_te_{t-1})=E[(\rho e_{t-1}+u_t)e_{t-1}]$$

$$=\rho E(e_{t-1}^2)+\underbrace{E(e_{t-1}u_t)}_{}$$

$$=\rho\,\frac{\sigma_u^2}{1-\rho^2}\quad,\quad t=2,3,\ldots,n. \qquad (4.9.13)$$

και γενικά, αποδεικνύεται ότι:

$$E(e_te_{t-s})=\rho^s\,\frac{\sigma_u^2}{1-\rho^2}\quad,\quad t=s+1,s+2,\ldots,n. \qquad (4.9.14)$$

Σύμφωνα με τις (4.9.11) και (4.9.12) τα σφάλματα e που ακολουθούν το αυτοπαλίνδρομο σχήμα πρώτης τάξης έχουν μέσο μηδέν και σταθερή διακύμανση $\frac{\sigma_u^2}{1-\rho^2}$, ενώ σύμφωνα με την (4.9.14) δεν ισχύει η υπόθεση (υ.3β) αλλά τα σφάλματα των διαφορετικών περιόδων συσχετίζονται και η συνδιακύμανση των σφαλμάτων που διαφέρουν κατά s περιόδους (s=1,2,...) προσδιορίζεται από την (4.8.14).

Μπορούμε ακόμα να παρατηρήσουμε ότι ο συντελεστής απλής συσχέτισης των σφαλμάτων που διαφέρουν κατά μία χρονική περίοδο είναι:

$$r_{e_te_{t-1}}=\frac{E(e_te_{t-1})}{\sqrt{E(e_t^2)}\,\sqrt{E(e_{t-1}^2)}}$$

$$=\frac{E(e_te_{t-1})}{E(e_t^2)},\qquad \Big[E(e_t^2)=E(e_{t-1}^2)=\frac{\sigma_u^2}{1-\rho^2}\Big],$$

$$=\rho\,,\qquad [\text{λόγω των } (4.9.12) \text{ και } (4.9.13)]\,, \qquad (4.9.15)$$

και γενικά, ο συντελεστής συσχέτισης των σφαλμάτων που διαφέρουν κατά s περιόδους είναι:

συνάρτηση αυτοσυσχέτισης

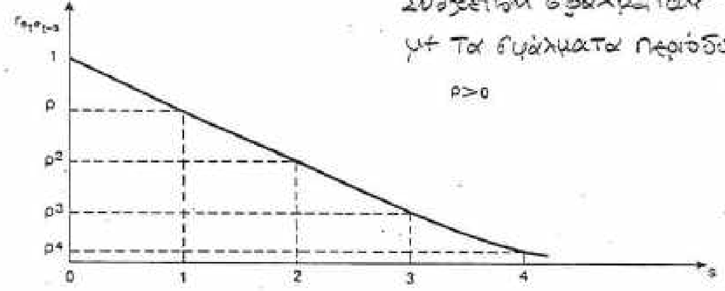
$$\boxed{r_{e_te_{t-s}}=\rho^s.} \qquad (4.9.16)$$

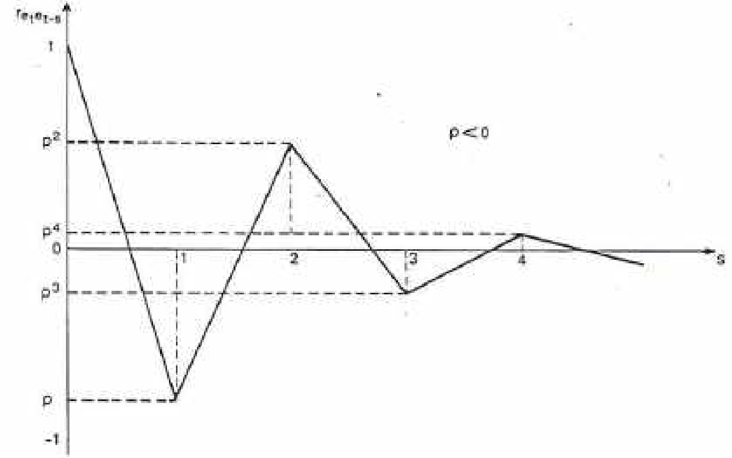
correlogram

Άρα η παράμετρος ρ είναι ο συντελεστής συσχέτισης των σφαλμάτων e_t σε διαδοχικές περιόδους στον πληθυσμό και συνεπώς η υπόθεση $|\rho|<1$ είναι απόλυτα δικαιολογημένη.

Η (4.9.16) ονομάζεται "συνάρτηση αυτοσυσχέτισης" (autocorrelation function ή correlogram) του αυτοπαλίνδρομου σχήματος πρώτης τάξης [AR(1)] και στο σχήμα 4.8 δίνεται το διάγραμμά της, δηλαδή η τιμή του συντελεστή απλής συσχέτισης $r_{e_t e_{t-s}}$ συναρτήσει του s.

Συσχέτιση σφαλμάτων περιόδου t με τα σφάλματα περιόδου t-s.





Σχήμα 4.8: Η συνάρτηση αυτοσυσχέτισης του αυτοπαλίνδρομου σχήματος πρώτης τάξης AR(1).

Από τις σχέσεις (4.9.11), (4.9.12) και (4.9.14) προκύπτει ότι, ο πίνακας συνδιακυμάνσεων των σφαλμάτων e_t, t=1, 2,...,n που ακολουθούν το αυτοπαλίνδρομο σχήμα πρώτης τάξης, είναι:

$$E(ee')=\sigma_u^2\,\frac{1}{1-\rho^2}\begin{bmatrix} 1 & \rho & \rho^2 & \dots & \rho^{n-1} \\ \rho & 1 & \rho & \dots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \dots & \rho^{n-3} \\ \vdots & & & & \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-s} & \dots & 1 \end{bmatrix}=\sigma_u^2 W \qquad (4.9.17)$$

Ο αντίστροφος του W είναι:

$$W^{-1}=\begin{bmatrix} 1 & -\rho & & & & 0 \\ -\rho & 1+\rho^2 & -\rho & & & \\ & -\rho & 1+\rho^2 & \ddots & & \\ & & \ddots & \ddots & -\rho & \\ & & & -\rho & 1+\rho^2 & -\rho \\ 0 & & & & -\rho & 1 \end{bmatrix} \qquad (4.9.18)$$

ενώ ο πίνακας

$$K'=\begin{bmatrix} \sqrt{1-\rho^2} & & & 0 \\ -\rho & 1 & & \\ & -\rho & 1 & \\ 0 & & \ddots & \ddots \\ & & -\rho & 1 \end{bmatrix} \qquad (4.9.19)$$

είναι τέτοιος ώστε

$$K'WK=I \qquad KK'=W^{-1}.$$

Αν τώρα μετασχηματίσουμε το υπόδειγμα $y=Xb+e$ χρησιμοποιόντας τον πίνακα K', θα έχουμε

$$K'y=K'Xb+K'e \qquad (4.9.20)$$

ή αναλυτικά:

$$K'y=\begin{bmatrix}\sqrt{1-\rho^2}Y_1\\ Y_2-\rho Y_1\\ Y_3-\rho Y_2\\ \vdots\\ Y_n-\rho Y_{n-1}\end{bmatrix},\quad K'X=\begin{bmatrix}\sqrt{1-\rho^2} & \sqrt{1-\rho^2}X_{11} & \dots & \sqrt{1-\rho^2}X_{1k}\\ 1-\rho & X_{2i}-\rho X_{11} & \dots & X_{2k}-\rho X_{1k}\\ \rho-\rho & X_{3i}-\rho X_{21} & \dots & X_{3k}-\rho X_{2k}\\ \vdots & \vdots & & \vdots\\ 1-\rho & X_{n1}-\rho X_{n-1} & \dots & X_{nk}-\rho X_{n-k}\end{bmatrix}$$

$$K'e=\begin{bmatrix}\sqrt{1-\rho^2}e_1\\ e_2-\rho e_1\\ e_3-\rho e_2\\ \vdots\\ e_n-\rho e_{n-1}\end{bmatrix}.$$

Είναι εύκολο να δούμε ότι οι εξισώσεις του μετασχηματισμένου υποδείγματος (4.9.20), από τις οποίες θα πάρουμε τις άριστες, αμερόληπτες, γραμμικές εκτιμήτριες των στοιχείων του διανύσματος b (AGLS), είναι:

$$\sqrt{1-\rho^2}Y_1=b_0\sqrt{1-\rho^2}+b_1\sqrt{1-\rho^2}X_{11}+\dots+b_k\sqrt{1-\rho^2}X_{1k}+\sqrt{1-\rho^2}e_1 \qquad (4.9.21)$$

για την πρώτη εξίσωση (t=1) και

$$Y_t-\rho Y_{t-1}=b_0(1-\rho)+b_1(X_{t1}-\rho X_{t-1,1})+\dots+$$
$$+b_k(X_{tk}-\rho X_{t-1,k})+e_t-\rho e_{t-1} \qquad (4.9.22)$$

για τις υπόλοιπες (t=2,3,...,n).

Συχνά, αντί του πίνακα K' χρησιμοποιείται ο πίνακας

$$K_1'=\begin{bmatrix}-\rho & 1 & & & 0\\ & -\rho & 1 & & \\ & & \ddots & \ddots & \\ & & & -\rho & 1\\ 0 & & & & -\rho \quad 1\end{bmatrix} \qquad (4.9.23)$$

που προκύπτει αν από τον πίνακα K' αφαιρέσουμε την πρώτη γραμμή. Αυτό σημαίνει ότι χρησιμοποιούμε (n-1) μετασχηματισμένες παρατηρήσεις αντί n (δεν χρησιμοποιείται η πρώτη παρατήρηση) και, όπως προκύπτει από την (4.8.22), οι (n-1) μετασχηματισμένες παρατηρήσεις είναι ίσες με τις αρχικές τιμές τους μείον ρ φορές την τιμή τους στην προηγούμενη χρονική περίοδο. Έχουμε όμως δείξει ότι οι άριστες, γραμμικές, αμερόληπτες εκτιμήτριες των παραμέτρων του αρχικού υποδείγματος προκύπτουν από το μετασχηματισμό με τον πίνακα K', ενώ οι εκτιμήτριες που προκύπτουν από τον μετασχηματισμό με τον πίνακα K_1' είναι απλώς προσεγγίσεις συχνά όχι ικανοποιητικές.

Στις εφαρμογές, η παράμετρος ρ, από την οποία εξαρτάται ο πίνακας K', δεν είναι βέβαια γνωστή και συνεπώς πρέπει να εκτιμηθεί με κάποια μέθοδο. Με τις μεθόδους εκτίμησης της παραμέτρου ρ θα ασχοληθούμε αμέσως μετά την παρουσίαση των κριτηρίων με τα οποία ελέγχουμε αν υπάρχει ή όχι αυτοσυσχέτιση πρώτης τάξης.

Η βιβλιογραφία γύρω από το θέμα των ελέγχων για την ύπαρξη αυτοσυσχέτισης είναι τεράστια ακόμα και αν περιοριστούμε στα κριτήρια για την ύπαρξη αυτοσυσχέτισης πρώτης τάξης. Στις περισσότερες εφαρμογές χρησιμοποιείται το κριτήριο των Durbin-Watson[1] το οποίο στηρίζεται στη στατιστική d των Durbin-Watson (D-W): DURBIN WATSON

$$\boxed{d=\frac{\sum_{t=2}^{n}(\hat{e}_t-\hat{e}_{t-1})^2}{\Sigma\hat{e}_t^2}} \qquad (4.9.24)$$

όπου $\hat{e}_t$, t=1,2,...,n, είναι τα κατάλοιπα από την εφαρμογή της μεθόδου Ε.Τ. στο αρχικό υπόδειγμα $y=Xb+e$. Η τιμή της στατιστικής d δίνεται από όλα σχεδόν τα προγράμματα των ηλεκτρονικών υπολογιστών.

1. J. Durbin and G.S. Watson, "Testing for Serial Correlation in Least-squares Regression" *Biometrica* Vol 37,(1950), pp. 409-428, and Vol. 38 (1951), pp. 159-178.

Οι υποθέσεις στις οποίες στηρίζεται ο έλεγχος των D-W είναι οι εξής:

ΥΠΟΘΕΣΕΙΣ D-W

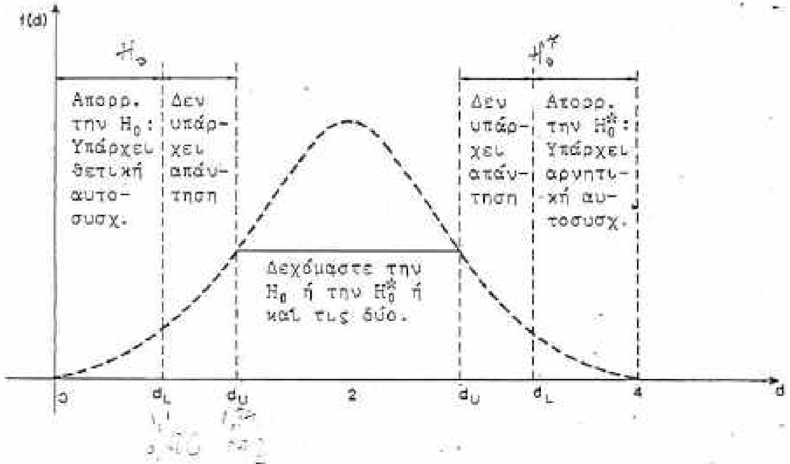
(i) Στην παλινδρόμηση υπάρχει σταθερός όρος

(ii) Οι τιμές των ερμηνευτικών μεταβλητών παραμένουν σταθερές σε επανειλημμένα δείγματα.

(iii) Η μορφή της αυτοσυσχέτισης είναι το αυτοπαλίνδρομο σχήμα πρώτης τάξης.

(iv) Δεν εισάγονται, ως ερμηνευτικές μεταβλητές, τιμές της εξαρτημένης μεταβλητής με χρονική υστέριση.

Επειδή η στατιστική d εξαρτάται από τις τιμές των ερμηνευτικών μεταβλητών $X_1, X_2, \ldots, X_k$, η ακριβής κατανομή της είναι δύσκολο να προσδιοριστεί και συνεπώς δεν ορίζονται μονοσήμαντα οι θεωρητικές τιμές της. Οι Durbin-Watson υπολόγισαν ένα κατώτερο και ένα ανώτερο όριο θεωρητικών τιμών (d_L και d_U αντίστοιχα) έτσι ώστε αν η υπολογιζόμενη τιμή της στατιστικής d βρίσκεται έξω από τα όρια αυτά να μπορούμε να δεχθούμε ή να απορρίψουμε την υπόθεση H_0 (δεν υπάρχει αυτοσυσχέτιση πρώτης τάξης). Τα όρια d_L και d_U εξαρτώνται μόνο από τον αριθμό των παρατηρήσεων (n=15 έως 100) και από τον



Σχήμα 4.9: Ο μηχανισμός του κριτηρίου των D-W για τον έλεγχο της αυτοσυσχέτισης πρώτης τάξης.

αριθμό των ερμηνευτικών μεταβλητών (k=1 έως 5) και όχι από τις τιμές των ερμηνευτικών μεταβλητών.

Ο μηχανισμός του κριτηρίου των D-W απεικονίζεται στο σχήμα 4.9 και τα βήματα για τη διεξαγωγή του είναι τα εξής:

(1) Υπολογίζουμε τις εκτιμήτριες Ε.Τ. του υποδείγματος $y = Xb + \varepsilon$ και τα κατάλοιπα $\hat{\varepsilon}$.

(2) Υπολογίζουμε την τιμή της στατιστικής d από την (4.9.24).

(3) Από τους πίνακες βρίσκουμε το κατώτερο και το ανώτερο όριο (d_L και d_U) για n παρατηρήσεις και k ερμηνευτικές μεταβλητές.

(4) Ελέγχουμε την υπόθεση

ΘΕΤΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ. (1ης τάξης)

$\boxed{H_0}$ Δεν υπάρχει θετική αυτοσυσχέτιση ($\rho > 0$) πρώτης τάξης.

Αν

$d < d_L$: απορρίπτουμε την H_0 (⇒ θετική αυτοσυσχ.)

$d > d_U$: δεν απορρίπτουμε την H_0 (δεχόμαστε H_0)

$d_L \leqslant d \leqslant d_U$: ο έλεγχος δεν δίνει απάντηση.

Αν θέλουμε να ελέγξουμε την υπόθεση

ΑΡΝΗΤΙΚΗ ΑΥΤΟΣΥΣΧΕΤΙΣΗ (1ης τάξης)

$\boxed{H_0^*}$ δεν υπάρχει αρνητική αυτοσυσχέτιση ($\rho < 0$) πρώτης τάξης

τότε, αντί των d_L και d_U χρησιμοποιούμε τα όρια $4-d_U$ και $4-d_L$ αντίστοιχα. Έτσι, αν

$d > 4-d_L$: απορρίπτουμε την H_0^*

$d < 4-d_U$: δεν απορρίπτουμε την H_0^*

$4-d_U \leqslant d \leqslant 4-d_L$: ο έλεγχος δεν δίνει απάντηση.

Ο λόγος για τον οποίο χρησιμοποιούμε τα όρια $4-d_U$ και $4-d_L$ για τον έλεγχο της αρνητικής αυτοσυσχέτισης μπορεί να δικαιολογηθεί ως εξής: η στατιστική d γράφεται

$$d=\frac{\Sigma\hat{e}_t^2+\Sigma\hat{e}_{t-1}^2-2\Sigma\hat{e}_t\hat{e}_{t-1}}{\Sigma\hat{e}_t^2}$$

$$\simeq 2\left(1-\frac{\Sigma\hat{e}_t\hat{e}_{t-1}}{\Sigma\hat{e}_t^2}\right) \qquad (4.9.25)$$

επειδή τα αθροίσματα $\Sigma\hat{e}_t^2$ και $\Sigma\hat{e}_{t-1}^2$ διαφέρουν κατά ένα μόνο όρο και μπορούμε να δεχτούμε ότι είναι περίπου ίσα κυρίως σε περίπτωση μεγάλου δείγματος. Άρα

$$\boxed{d\simeq 2(1-\hat{\rho})} \qquad (4.9.26)$$

όπου

$$\hat{\rho}=\frac{\Sigma\hat{e}_t\hat{e}_{t-1}}{\Sigma\hat{e}_t^2} \qquad (4.9.27)$$

είναι ο συντελεστής συσχέτισης πρώτης τάξης των καταλοίπων της παλινδρόμησης.

Από την (4.9.26) εύκολα προκύπτει ότι:

$\hat{\rho}=1 \Leftrightarrow d=0$ πλήρης θετική συσχέτιση

$\hat{\rho}=0 \Leftrightarrow d=2$ μηδενική συσχέτιση

$\rho=-1 \Leftrightarrow d=4.$ πλήρης αρνητική συσχέτιση

Άρα, όσο η τιμή της στατιστικής d πλησιάζει το μηδέν η θετική αυτοσυσχέτιση ενισχύεται, όσο πλησιάζει το 2, εκ των κάτω ή εκ των άνω, η αυτοσυσχέτιση ελαττώνεται ενώ, όσο πλησιάζει το 4 ενισχύεται η αρνητική αυτοσυσχέτιση.

Πολλές προσπάθειες έχουν γίνει από διάφορους ερευνητές για την αντιμετώπιση του προβλήματος της έλλειψης απάντησης από τον έλεγχο των D-W στα διαστήματα μεταξύ d_L, d_U και $4-d_U$, $4-d_L$. Π.χ. οι Theil-Nagar[1] και Hannan-Terrell[2], κάτω από την

1. H. Theil and A.L. Nagar, "Testing the Independence of Regression Disturbances" *Journal of the American Statistical Association*, Vol. 56 (1961), pp. 793-806.

2. E.J. Hannan and R.D. Terrell, "Testing for Serial Correlation after Least Squares Regresion", *Economica*, Vol. 36 (1968), pp. 133-150.

υπόθεση ότι οι οικονομικές χρονολογικές σειρές μεταβάλλονλονται με αργό ρυθμό, έδειξαν ότι μπορούμε να αγνοήσουμε το κατώτατο όριο d_L και να χρησιμοποιήσουμε μόνο την τιμή του ανωτέρου ορίου d_U (ή $4-d_U$) για τον έλεγχο της υπόθεσης H_0 (ή της H_0^*). Σχετικά με άλλες προσπάθειες βλέπε G.S. Maddala, Econometrics, σελ. 285-291 ενώ σχετικά με άλλα κριτήρια για την ύπαρξη αυτοσυσχέτισης πρώτης τάξης βλέπε J. Johnston, Econometric Methods, 2nd edition σελ. 250 και 254-258.

Προχωρούμε τώρα στην παρουσίαση των κυριωτέρων μεθόδων εκτίμησης της άγνωστης παραμέτρου ρ που χρησιμοποιούνται στις εφαρμογές:

1. Εκτίμηση της παραμέτρου ρ από τη στατιστική των Durbin-Watson: Από τη σχέση (4.9.26) εύκολα προκύπτει ότι

$$\boxed{\hat{\rho}\simeq 1-\frac{d}{2}} \quad \text{μεγάλα δείγματα.} \qquad (4.9.28)$$

η οποία μας δίνει την εκτίμηση της παραμέτρου ρ συναρτήσει της στατιστικής d των D-W. Συνεπώς, κάτω από τις υποθέσεις του κριτηρίου D-W, η στατιστική d παρέχει μια κατά προσέγγιση εκτίμηση της παραμέτρου ρ. Πρέπει όμως να διευκρινίσουμε ότι η σχέση (4.9.28) ισχύει ασυμπτωτικά, δηλαδή για μεγάλα δείγματα, ενώ για μικρά δείγματα η προσέγγιση μπορεί να μην είναι καλή. Για μικρά δείγματα οι Theil-Nagar[1] υπέδειξαν την ακόλουθη εκτίμηση της παραμέτρου ρ:

$$\boxed{\hat{\rho}=\frac{n^2\left(1-\frac{d}{2}\right)+(k+1)^2}{n^2-(k+1)^2}} \quad \text{μικρά δείγματα.} \qquad (4.9.29)$$

όπου n είναι το μέγεθος του δείγματος, k ο αριθμός των ερμηνευτικών μεταβλητών, και d η στατιστική των D-W.

Αν η τιμή της στατιστικής d είναι μηδέν ή περίπου μηδέν τότε $\hat{\rho}\simeq 1$ και η πρώτη παρατήρηση (4.9.21) δεν έχει έν-

1. H. Theil and A.L. Nagar, "Testing the Independence of Regression Distrubances" ο.π.

νοια ενώ οι μετασχηματισμένες παρατηρήσεις (4.9.22) εκφράζονται σε πρώτες διαφορές των μεταβλητών:

$$Y_t-Y_{t-1}=b_1(X_{t1}-\rho X_{t-1,1})+\cdots+b_k(X_{tk}-\rho X_{t-1,k})+e_t-e_{t-1} \quad (4.9.30)$$

ή

$$\Delta Y_t=b_1\Delta X_{t1}+b_2\Delta X_{t2}+\cdots+b_k\Delta X_{tk}+e_t^* \quad (4.9.30)$$

όπου

$$\Delta Y_t=Y_t-Y_{t-1},\quad \Delta X_t=X_t-X_{t-1} \quad \text{και} \quad e_t^*=\Delta e_t=e_t-e_{t-1}.$$

Αν $d\simeq 4$ τότε $\rho\simeq -1$ και οι μετασχηματισμένες εξισώσεις (4.9.22) παίρνουν τη μορφή:

$$Y_t+Y_{t-1}=2b_0+b_1(X_{t1}-X_{t-1,1})+\cdots+b_k(X_{tk}-X_{t-1,k})+e_t+e_{t-1} \quad (4.9.32)$$

και το υπόδειγμα (4.9.32) είναι γνωστό ως υπόδειγμα κινητού μέσου δύο περιόδων.

$\boxed{2}$ Η μέθοδος των Cochrane-Orcutt[1]: Υποθέτοντας ότι η αυτοσυσχέτιση εκφράζεται από το αυτοπαλίνδρομο σχήμα $e_t=\rho e_{t-1}+u_t$, $|\rho|<1$, και τα σφάλματα u_t ικανοποιούν τις γνωστές υποθέσεις οι Cochrane-Orcutt υποδεικνύουν την εξής μέθοδο για την εκτίμηση της παραμέτρου ρ: εφαρμόζουμε τη μέθοδο Ε.Τ. στο αρχικό υπόδειγμα αγνοώντας την ύπαρξη αυτοσυσχέτισης, υπολογίζουμε τα κατάλοιπα $\hat{e}_t$, $t=1,2,\ldots,n$ της παλινδρόμησης και εκτιμούμε την παράμετρο ρ από την παλινδρόμηση $\hat{e}_t=\rho\hat{e}_{t-1}+u_t$, παίρνοντας την εκτίμηση $\hat{\rho}=\Sigma\hat{e}_t\hat{e}_{t-1}/\Sigma\hat{e}_t^2$. Στη συνέχεια χρησιμοποιούμε την εκτίμηση $\hat{\rho}$ για να μετασχηματίσουμε τις αρχικές εξισώσεις

$-b_1\hat{\rho}$

$$Y_t-\hat{\rho}Y_{t-1}=b_0(1-\hat{\rho})+b_1(X_{t1}-\hat{\rho}X_{t-1,1})+\cdots+$$
$$+b_k(X_{tk}-\hat{\rho}X_{t-1,k})+e_t-\hat{\rho}e_{t-1}, \quad (4.9.33)$$

1. D. Cochrane and G.H. Orcutt, "Application of Least-squares Regressions to Relationships Containing Auto-correlated Error Terms", *Journal of the American Statistical Association*, Vol. 44 (1949), pp. 32-61.

εκτιμούμε την (4.9.33) με τη μέθοδο Ε.Τ., υπολογίζουμε τα νέα κατάλοιπα $\hat{\hat{e}}_t$ και από την παλινδρόμηση $\hat{\hat{e}}_t=\rho\hat{\hat{e}}_{t-1}+u_t$ παίρνουμε τη νέα εκτίμηση της παραμέτρου $\hat{\hat{\rho}}=\Sigma\hat{\hat{e}}_t\hat{\hat{e}}_{t-1}/\Sigma\hat{\hat{e}}_t^2$ την οποία χρησιμοποιούμε για να μετασχηματίσουμε ξανά το αρχικό υπόδειγμα:

$$Y_t-\hat{\hat{\rho}}Y_{t-1}=b_0(1-\hat{\hat{\rho}})+b_1(X_{t1}-\hat{\hat{\rho}}X_{t-1,1})+\cdots+b_k(X_{tk}-\hat{\hat{\rho}}X_{t-1,k})+e_t-\hat{\hat{\rho}}e_{t-1}.$$

Η επαναληπτική αυτή διαδικασία συνεχίζεται μέχρις ότου οι διαδοχικές εκτιμήσεις της παραμέτρου ρ δε διαφέρουν σημαντικά μεταξύ τους. Ως αρχική εκτίμηση $\hat{\rho}$ μπορεί να χρησιμοποιηθεί και η (4.9.29).

$\boxed{3}$ Η μέθοδος των Hildreth-Lu[1]: Επειδή στο αυτοπαλίνδρομο σχήμα πρώτης τάξης $e_t=\rho e_{t-1}+u_t$ έχουμε $-1<\rho<1$ οι Hildreth-Lu χώρισαν το διάστημα (-1,1) σε ίσα διαστήματα (π.χ. -0.9, -0.8, ..., -0.1, 0, 0.1, ..., 0.8, 0.9) και χρησιμοποίησαν τις τιμές αυτές του ρ για να μετασχηματίσουν τις αρχικές εξισώσεις. Από το σύνολο των παλινδρομήσεων επέλεξαν εκείνη που μεγιστοποιεί το R^2. Αν επιθυμούμε καλλίτερη προσέγγιση τότε μπορούμε να πειραματιστούμε με διάφορες τιμές γύρω από την τιμή $\hat{\rho}$ που επιλέξαμε στην πρώτη διαδικασία.

$\boxed{4}$ Η μέθοδος του Durbin σε δύο βήματα[2]: Η μέθοδος αυτή έχει ως αφετηρία τις μετασχηματισμένες εξισώσεις (4.9.22) οι οποίες γράφονται ως εξής

$$Y_t=b_0(1-\rho)+\rho Y_{t-1}+b_1X_{t1}-b_1\rho X_{t-1,1}+\cdots+$$
$$+b_kX_{tk}-b_k\rho X_{t-1,k}+u_t. \quad (4.9.34)$$

Αν στο υπόδειγμα (4.9.34) εφαρμόσουμε τη μέθοδο Ε.Τ.

1. G. Hildreth and J.Y. Lu, "Demand Relations with Autocorrelated Disturbances", Michigan State University, Agricultural Experiment Station, Tech. Bull. 276, November 1960.

2. J. Durbin, "Estimation of Parameters in Time-Series Regression Models", *Journal of the American Statistical Association, ser. B*, Vol. 22 (1960), pp. 139-153.

ο συντελεστής της μεταβλητής Y_{t-1} είναι συνεπής εκτιμήτρια της παραμέτρου ρ. Την εκτιμήτρια $\hat{\rho}$ που θα προκύψει από την (4.9.34) χρησιμοποιούμε, κατά τα γνωστά, για τον μετασχηματισμό του αρχικού υποδείγματος. Υπάρχουν ενδείξεις ότι η μέθοδος του Durbin είναι πιο αποτελεσματική από τις προηγούμενες (βλέπε σχετικά, J. Johnston, Econometric Methods, σελ. 264-265).

4.10. Η ΥΠΟΘΕΣΗ (υ.4): ΠΟΛΥΣΥΓΓΡΑΜΜΙΚΟΤΗΤΑ (χρονολογικές σειρές)

Η τέταρτη βασική υπόθεση του κλασικού γραμμικού υποδείγματος απαιτεί ο πίνακας X των παρατηρήσεων των ανεξάρτητων μεταβλητών $X_1, X_2, \ldots, X_k$ να είναι πλήρους βαθμού:

$$r(X)=k+1 \qquad (4.10.1)$$

δηλαδή οι (k+1) στήλες του πίνακα X να είναι γραμμικά ανεξάρτητες. Αν η υπόθεση (υ.4) ισχύει τότε ο πίνακας $X'X$ είναι μη ιδιάζων, υπάρχει ο αντίστροφός του και το σύστημα των κανονικών εξισώσεων

$$(X'X)\hat{b}=X'y \qquad (4.10.2)$$

μπορεί να ληθεί και να δώσει τις εκτιμήτριες $\hat{b}$ των Ε.Τ. Αν όμως $r(X)<k+1$, τότε ο πίνακας $X'X$ είναι ιδιάζων ($|X'X|=0$) το σύστημα των κανονικών εξισώσεων δεν έχει λύση και το διάνυσμα $\hat{b}$ των εκτιμητριών Ε.Τ. δεν μπορεί να υπολογιστεί. Η περίπτωση αυτή που είναι γνωστή ως "πλήρης πολυσυγγραμμικότητα" (perfect multicollinearity) παρατηρείται σε περιπτώσεις που ο ερευνητής διαπράττει βασικά λάθη στον προσδιορισμό των ανεξάρτητων μεταβλητών -όπως π.χ. αν μία ερμηνευτική μεταβλητή είναι ο μέσος όρος μερικών άλλων, ή αν μία ερμηνευτική μεταβλητή παραμένει σταθερή σε όλη τη διάρκεια του δείγματος οπότε οι τιμές της είναι πολλαπλάσια της μοναδιαίας στήλης που αντιστοιχεί στο σταθερό όρο, ή αν εξαντλήσουμε όλες τις κατηγορίες στην περίπτωση των ψευδομεταβλητών ενώ συγχρόνως εισάγουμε και σταθερό όρο (παγίδα των ψευδομεταβλητών) κλπ.- και η διαπίστωσή της είναι εύκολη αφού δεν μπορούμε να υπολογίσουμε το διάνυσμα $\hat{b}$. Για την αντιμετώπιση της πλήρους πολυσυγγραμμικότητας αρκεί να απαλείψουμε από την παλινδρόμηση τις μεταβλητές που δημιουργούν το πρόβλημα. Πρέπει να τονίσουμε ότι η αδυναμία της εκτίμησης του διανύσματος $\hat{b}$ δεν σημαίνει ότι δεν μπορούμε να εκτιμήσουμε κανένα από τα στοιχεία του, αλλά ότι δεν μπορούμε να εκτιμήσουμε τους συντελεστές των ερμηνευτικών μεταβλητών που παρουσιάζουν πλήρη γραμμική συσχέτιση. Ας θεωρήσουμε π.χ. το υπόδειγμα:

$$Y_t=b_0+b_1X_{t1}+b_2X_{t2}+b_3X_{t3}+e_t, \quad t=1,2,\ldots,n \qquad (4.10.3)$$

στο οποίο οι μεταβλητές X_2 και X_3 συνδέονται με την ακριβή γραμμική σχέση

$$X_{t3}=3X_{t2}. \qquad (4.10.4)$$

Αν αντικαταστήσουμε την (4.10.4) στην (4.10.3) προκύπτει το υπόδειγμα:

$$Y_t=b_0+b_1X_{t1}+(b_2+3b_3)X_{t2}+e_t$$

από το οποίο ενώ μπορούμε να υπολογίσουμε τις εκτιμήτριες $\hat{b}_0$ και $\hat{b}_1$, αντίθετα, δεν μπορούμε να υπολογίσουμε τις $\hat{b}_2$ και $\hat{b}_3$ παρά μόνο το γραμμικό συνδιασμό τους $\widehat{b_2+3b_3}$. Αυτό σημαίνει ότι δεν μπορούμε να εκτιμήσουμε τη χωριστή επίδραση που ασκεί κάθε μία από τις μεταβλητές X_2 και X_3 πάνω στην Y, αλλά μόνο τη συνδιασμένη επίδρασή τους.

Όμως στην ποσοτική οικονομική ανάλυση, πιο συχνή δεν είναι η περίπτωση της πλήρους πολυσυγγραμμικότητας αλλά η περίπτωση της "πολυσυγγραμμικότητας" (multicollinearity) κατά την οποία ο πίνακας $X'X$ δεν είναι ιδιάζων αλλά "περίπου" ιδιάζων, με την έννοια ότι η τιμή της ορίζουσας $|X'X|$ είναι περίπου ίση με το μηδέν. ΠΟΛΥΣΥΓΓΡΑΜΜΙΚΟΤΗΤΑ. Το ότι $|X'X|\approx 0$ σημαίνει ότι οι παρατηρήσεις πάνω στις ανεξάρτητες μεταβλητές $X_1, X_2, \ldots, X_k$ δεν συνδέονται με ακριβείς γραμμικές σχέσεις αλλά με περίπου γραμμικές σχέσεις. Το φαινόμενο αυτό είναι χαρακτηριστικό όλων σχεδόν των οικονομετρικών αναλύσεων που χρησιμοποιούν χρονολογικές σειρές και ιδιαίτερα μακρο-οικονομικές χρονολογικές σειρές οι οποίες, κυρίως σε περιόδους ανάκαμψης, κινούνται προς την ίδια κατεύθυνση και παρουσιάζουν κοινές διαχρονικές τάσεις. Αν λοιπόν σε μια παλινδρόμηση χρησιμοποιήσουμε ως ερμηνευτικές μεταβλητές δύο ή περισσότερες τέτοιες μεταβλητές, είναι σχεδόν βέβαιο ότι θα αντιμετωπίσουμε το πρόβλημα της πολυσυγγραμμικότητας.

Αν $|X'X|\approx 0$ τότε, τα διαγώνια στοιχεία του πίνακα $(X'X)^{-1}$ θα είναι μεγάλα και επειδή $C(\hat{b})=\sigma^2(X'X)^{-1}$, οι διακυμάνσεις (και τα τυπικά σφάλματα) των εκτιμητριών $\hat{b}$ Ε.Τ. θα είναι μεγάλες με αποτέλεσμα οι λόγοι t -που είναι αντιστρόφως ανάλογοι προς τα τυπικά σφάλματα- να παίρνουν μικρές τιμές. Συνέπεια των μικρών τιμών των λόγων είναι οι εκτιμήτριες $b_1, b_2, \dots, b_k$ να μην είναι στατιστικά σημαντικές. Από την άλλη μεριά ενδέχεται ο συντελεστής πολλαπλού προσδιορισμού R^2 της παλινδρόμησης να είναι υψηλός και το κριτήριο F να δείχνει ότι η συνολική ερμηνευτική ικανότητα της παλινδρόμησης είναι στατιστικά σημαντική. Τα συμπτώματα αυτά (υψηλή ερμηνευτική ικανότητα της παλινδρόμησης και μη σημαντικοί συντελεστές $\hat{b}_1, \hat{b}_2, \dots, \hat{b}_k$) αποτελούν σοβαρές ενδείξεις για την ύπαρξη πολυσυγγραμμικότητας. Ακόμα, μπορούμε να προσθέσουμε ότι, στην περίπτωση που υπάρχει πολυσυγγραμμικότητα μεταξύ των ερμηνευτικών μεταβλητών, οι εκτιμήτριες $\hat{b}_1, \hat{b}_2, \dots, \hat{b}_k$ (ή τουλάχιστον μερικές από αυτές) είναι πολύ ευαίσθητες στην εισαγωγή ή στην παράλειψη ερμηνευτικών μεταβλητών καθώς και στην προσθήκη ή αφαίρεση παρατηρήσεων από το δείγμα.

ΕΝΔΕΙΞΕΙΣ ΓΙΑ ΤΗΝ ΥΠΑΡΞΗ ΠΟΛΥΣΥΓΓΡΑΜΜΙΚΟΤΗΤΑΣ

Όπως αναφέραμε πιο πάνω χαρακτηριστική περίπτωση πολυσυγγραμμικότητας έχουμε όταν ο συντελεστής R^2 είναι πολύ υψηλός ενώ συγχρόνως κανένας από τους συντελεστές $\hat{b}_1, \hat{b}_2, \dots, \hat{b}_k$ δεν είναι στατιστικά σημαντικός. Αυτό όμως αποτελεί ακραία περίπτωση.

Αν στο υπόδειγμά μας έχουμε μόνο δύο ερμηνευτικές μεταβλητές, η διαπίστωση της πολυσυγγραμμικότητας μπορεί να γίνει εύκολα από την τιμή του συντελεστή συσχέτισης r_{X_1,X_2}.

Αν η τιμή του είναι υψηλή τότε οι ερμηνευτικές μεταβλητές X_1 και X_2 παρουσιάζουν υψηλή γραμμική συσχέτιση και συνεπώς υπάρχει πολυσυγγραμμικότητα. Αν όμως στο υπόδειγμά μας έχουμε περισσότερες από δύο ερμηνευτικές μεταβλητές τότε η εξέταση των συντελεστών συσχέτισης ή ο υπολογισμός της ορίζουσας $|R_X|$ του πίνακα R_X των συντελεστών απλής συσχέτισης δεν αρκούν για τη διαπίστωση της πολυσυγγραμμικότητας γιατί, είναι δυνατόν, οι συντελεστές συσχέτισης των ερμηνευτικών μεταβλητών να είναι χαμηλοί και όμως το πρόβλημα της πολυσυγγραμμικότητας να είναι έντονο. Στην περίπτωση περισσότερων ερμηνευτικών μεταβλητών πρέπει να εξετάσουμε και τους συντελεστές μερικής συσχέτισης -π.χ. τους $r_{YX_1 \cdot X_2X_3}$, $r_{YX_2 \cdot X_1X_3}$ και $r_{YX_3 \cdot X_1X_2}$ αν έχουμε τρεις ερμηνευτικές μεταβλητές. Αν ο R^2 είναι υψηλός και οι συντελεστές μερικής συσχέτισης είναι χαμηλοί τότε είναι πολύ πιθανό να έχουμε πολυσυγγραμμικότητα, ενώ αν και ο R^2 και οι συντελεστές μερικής συσχέτισης είναι υψηλοί τότε η διαπίστωση της πολυσυγγραμμικότητας δεν είναι άμεσα δυνατή. Στην περίπτωση αυτή εκτιμούμε τις παλινδρομήσεις κάθε ερμηνευτικής μεταβλητής X_i, $i=1,2,\dots,k$, πάνω στις υπόλοιπες ερμηνευτικές μεταβλητές και υπολογίζουμε τους αντίστοιχους συντελεστές πολλαπλού προσδιορισμού R_i^2 των παλινδρομήσεων αυτών. Αν κάποιος συντελεστής R_i^2 είναι υψηλός αυτό σημαίνει ότι η μεταβλητή X_i συσχετίζεται υψηλά με τις υπόλοιπες και συνεπώς μπορούμε να την απαλείψουμε από την παλινδρόμηση αρκεί να μην δημιουργείται σοβαρό πρόβλημα λανθασμένου προσδιορισμού του υποδείγματος.

ΕΛΕΓΧΟΙ

Η διαπίστωση της πολυσυγγραμμικότητας αποτελεί το ένα σκέλος του προβλήματος. Το δεύτερο σκέλος είναι η αντιμετώπισή της.

Στο σημείο αυτό, πρέπει να τονίσουμε, ότι η πολυσυγγραμμικότητα είναι ένα "χαρακτηριστικό του δείγματος" με την έννοια ότι το συγκεκριμένο δείγμα δεν παρέχει την "ποσότητα πληροφόρησης" που απαιτείται για την αποτελεσματική εκτίμηση των παραμέτρων του υποδείγματος που έχουμε προσδιορίσει. Αν μάλιστα λάβουμε υπόψη ότι τα στατιστικά στοιχεία που χρησιμοποιούνται για την εκτίμηση των οικονομετρικών υποδειγμά-

ΧΑΡΑΚΤΗΡΙΣΤΙΚΟ ΤΟΥ ΔΕΙΓΜΑΤΟΣ

των δεν προέρχονται από "κατευθυνόμενα πειράματα" αλλά εκφράζουν τη μοναδικότητα του οικονομικού περελθόντος, τότε γίνεται αμέσως αντιληπτό ότι δεν έχουμε καμιά ελπίδα να καλύψουμε το χάσμα μεταξύ των πληροφοριακών απαιτήσεων ενός σωστά προσδιορισμένου υποδείγματος και της προσφοράς πληροφοριών από ένα δείγμα στο οποίο οι ερμηνευτικές μεταβλητές δεν παρουσιάζουν "αρκετή" ανεξάρτητη μεταβλητικότητα. Το πρόβλημα συνεπώς δεν έχει λύση αν δεν υπάρξει κάποιος "συμβιβασμός", λιγότερο ή περισσότερο επώδυνος, μεταξύ των απαιτήσεων του υποδείγματος και της προσφοράς πληροφοριών από το δείγμα. Οι μέθοδοι που αναφέρονται στα διάφορα εγχειρίδια για την "αντιμετώπιση" της πολυσυγγραμμικότητας αποτελούν ακριβώς προσπάθειες είτε για τη μείωση των προς εκτίμηση παραμέτρων με την απαλοιφή των "προβληματικών" μεταβλητών είτε για την ενίσχυση της πληροφοριακής ικανότητας του δείγματος με την αναζήτηση συμπληρωματικών πληροφοριών που θα καταστήσουν δυνατή την αποτελεσματικότερη εκτίμηση των παραμέτρων του υποδείγματος. Βέβαια, υπάρχει και η άποψη να αποδεχτούμε την πραγματική κατάσταση που εκφράζει η πολυσυγγραμμικότητα και να αφήσουμε τα πράγματα όπως είναι. Άλλωστε το ότι δεν μπορούμε να υπολογίσουμε αξιόπιστα τις χωριστές επιδράσεις των ερμηνευτικών μεταβλητών πάνω στην εξαρτημένη μεταβλητή είναι ένα πραγματικό γεγονός και δεν υπάρχει συνεπώς κανένας λόγος να προχωρήσουμε σε επιζήμιους συμβιβασμούς για να επιτύχουμε εκτιμήσεις πιο ακριβείς από τις πραγματικές τιμές των παραμέτρων. Ποιον από τους τρεις δρόμους θα ακολουθήσουμε εξαρτάται βέβαια από τη φύση του οικονομικού φαινομένου που ερευνούμε καθώς και από το σκοπό για τον οποίο διεξάγεται η έρευνα. Αν π.χ. ο σκοπός της έρευνας είναι η διατύπωση προβλέψεων τότε η παρουσία της πολυσυγγραμμικότητας ίσως να μην είναι τόσο σοβαρό πρόβλημα. Συνήθως επιτυγχάνονται καλές προβλέψεις από σωστά προσδιορισμένα υποδείγματα με υψηλή ερμηνευτική ικανότητα (υψηλή τιμή του R^2) παρά την ύπαρξη της πολυσυγγραμμικότητας, δεδομένου ότι η ίδια σχέση μεταξύ των ερμηνευτικών μεταβλητών αναμένεται να ισχύει και κατά την περίοδο της πρόβλεψης. Αν-

προβλέψεις

τίθετα, αν ο σκοπός της έρευνας είναι η διαρθρωτική ανάλυση και ο έλεγχος υποθέσεων τότε απαιτούνται αξιόπιστες εκτιμήτριες των παραμέτρων του υποδείγματος και η πολυσυγγραμμικότητα είναι πράγματι πολύ σοβαρό πρόβλημα που πρέπει να το αντιμετωπίσουμε.

Έλεγχος Υποθέσεων

Παρουσιάζουμε στη συνέχεια τις κυριώτερες "λύσεις" που έχουν προταθεί για την αντιμετώπιση της πολυσυγγραμμικότητας μαζί με κάποια κριτική των διαφόρων μεθόδων:

1. Απαλοιφή των προβληματικών ερμηνευτικών μεταβλητών.

Η απλούστερη λύση στην περίπτωση σοβαρής πολυσυγγραμμικότητας είναι να απαλείψουμε από την εξίσωση την προβληματική ή τις προβληματικές μεταβλητές. Αλλά, παραλείποντας μία ή περισσότερες μεταβλητές από ένα σωστά προσδιορισμένο υπόδειγμα, διαπράττουμε σφάλμα προσδιορισμού το οποίο, όπως αναλύσαμε σε προηγούμενο κεφάλαιο, έχει ως συνέπεια οι εκτιμήτριες που θα πάρουμε από το νέο υπόδειγμα να είναι μεροληπτικές. Π.χ. αν το σωστό υπόδειγμα για την κατανάλωση είναι → σελ. 144

$$C_t=b_0+b_1Y_t+b_2W_t+e_t \qquad (4.10.5)$$

όπου C_t είναι η κατανάλωση, Y_t το διαθέσιμο εισόδημα και W_t η περιουσία στην περίοδο t και, λόγω της υψηλής συσχέτισης των Y_t και W_t, παραλείψουμε την W_t και εκτιμήσουμε το υπόδειγμα

$$C_t=\alpha_0+\alpha_1Y_t+u_t \qquad (4.10.6)$$

τότε, έχουμε δείξει ότι,

$$E(\hat{\alpha}_1)=b_1+\hat{d}b_2 \qquad (4.10.7)$$

→ σελ. 143

όπου $\hat{d}$ είναι η κλίση της παλινδρόμησης

$$W_t=c+dY_t+\upsilon_t. \qquad (4.10.8)$$

Γίνεται λοιπόν φανερό από την (4.10.7) ότι η $\hat{\alpha}_1$ θα είναι μεροληπτική εκτιμήτρια της αληθούς οριακής ροπής προς κατανάλωση b_1 στον πληθυσμό αφού η $\hat{d}$ είναι προφανώς διάφορη από το μηδέν (οι μεταβλητές W_t και Y_t έχουν υψηλή γραμμική συσχέτιση). Ακόμα αν $\hat{d}b_2>0$ (πότε συμβαίνει αυτό;) η $\hat{a}_1$ θα υπερεκτιμά την b_1 ενώ αν $\hat{d}b_2<0$ η $\hat{a}_1$ θα υποεκτιμά την b_1.

Παραλείποντας επομένως μία ή περισσότερες ερμηνευτικές μεταβλητές για να αντιμετωπίσουμε το πρόβλημα της πολυσυγγραμμικότητας μπορεί να οδηγηθούμε σε παραπλανητικά αποτελέσματα, περισσότερο ή λιγότερο σοβαρά, ανάλογα με την περίπτωση.

2. Χρησιμοποίηση "a priori" πληροφοριών.

Στο προηγούμενο παράδειγμα της κατανάλωσης ας υποθέσουμε ότι το εισόδημα Y και η περιουσία W παρουσιάζουν υψηλή γραμμική συσχέτιση αλλά, από σχετικές οικονομετρικές έρευνες, έχει διαπιστωθεί ότι $b_2\simeq 0,15b_1$ (να ερμηνεύσετε οικονομικά τη σχέση αυτή). Στην περίπτωση αυτή η (1) γράφεται:

$$C_t=b_0+b_1Y_t+0,15b_1W_t+e_t$$

$$=b_0+b_1X_t+e_t, \qquad (4.10.9)$$

όπου $X_t=Y_t+0,15W_t$, και από την εκτίμηση $\hat{b}_1$ προκύπτει ότι $\hat{b}_2=0,15\hat{b}_1$.

Σχετική είναι και η περίπτωση κατά την οποία χρησιμοποιούμε ένα συνδυασμό διαχρονικών και διαστρωματικών στοιχείων. Αν π.χ. θέλουμε να εκτιμήσουμε την εξίσωση:

$$\log Q_t=b_0+b_1\log P_t+b_2\log Y_t+e_t \qquad (4.10.10)$$

όπου Q_t και P_t είναι, αντίστοιχα, η ζητούμενη ποσότητα και η τιμή ενός διαρκούς αγαθού και Y_t το διαθέσιμο εισόδημα στην περίοδο t, τότε, είναι πολύ πιθανό οι μεταβλητές P_t και Y_t να παρουσιάζουν υψηλή γραμμική συσχέτιση (δικαιολογείται το φαινόμενο αυτό στην περίπτωση χρονολογικών σειρών;) με αποτέλεσμα οι εκτιμήτριες $\hat{b}_1$ και $\hat{b}_2$ της ελαστικότητας της ζήτησης ως προς την τιμή και το εισόδημα να είναι ασταθείς και αναξιόπιστες. Σε τέτοιου είδους οικονομετρικές αναλύσεις η

συνήθως πρακτική είναι να εκτιμήσουμε την ελαστικότητα b_2 της ζήτησης ως προς το εισόδημα από διαστρωματικά στοιχεία (οι τιμές στην περίπτωση αυτή δε διαφέρουν σημαντικά) και στη συνέχεια να εκτιμήσουμε την ελαστικότητα b_1 της ζήτησης ως προς την τιμή P_t από την εξίσωση:

$$\log Q_t-\hat{b}_2\log Y_t=b_0+b_1\log P_t+u_t. \qquad (4.10.11)$$

Βέβαια, δεν πρέπει να μας διαφεύγει το ότι, η $\hat{b}_2$ είναι μια εκτιμήτρια και όχι η πραγματική τιμή της ελαστικότητας b_2 και αυτό πρέπει να λαμβάνεται υπόψη στον υπολογισμό των διακυμάνσεων των $\hat{b}_0$ και $\hat{b}_1$[1]. Ακόμη, πρέπει να έχουμε υπόψη ότι, η $\hat{b}_2$ εκτιμά την ελαστικότητα της ζήτησης ως προς το εισόδημα σε κάποια δεδομένη χρονική στιγμή (ενώ το περιεχόμενό της στην εξίσωση (4.10.10) είναι διαφορετικό) και η τιμή της ενδέχεται να μην παραμένει σταθερή κατά την περίοδο που καλύπτεται από το δείγμα.

3. Μετασχηματισμός των μεταβλητών.

Σε πολλές εφαρμογές χρησιμοποιούνται οι πρώτες διαφορές ή λόγοι των αρχικών μεταβλητών για την αντιμετώπιση της πολυσυγγραμμικότητας. Π.χ. αν από την εξίσωση:

ΠΡΩΤΕΣ ΔΙΑΦΟΡΕΣ

$$C_t=b_0+b_1Y_t+b_2W_t+e_t \qquad (4.10.12)$$

αφαιρέσουμε την τιμή της στην περίοδο (t-1):

$$C_{t-1}=b_0+b_1Y_{t-1}+b_2W_{t-1}+e_{t-1} \qquad (4.10.13)$$

θα προκύψει η εξίσωση:

$$C_t-C_{t-1}=b_1(Y_t-Y_{t-1})+b_2(W_t-W_{t-1})+u_t \qquad (4.10.14)$$

όπου $u_t=e_t-e_{t-1}$.

---

1. Για το πως θα γίνει αυτό βλέπε G.S. Maddala, "The Likelihood Approach to Pooling Cross-Setion and Time-Series Data", *Econometrica*, Vol. 39 (1971), pp. 939-953.

ΠΡΩΤΕΣ ΔΙΑΦΟΡΕΣ

Είναι γενική διαπίστωση ότι η πολυσυγγραμμικότητα στην (4.10.14) θα είναι λιγότερο σοβαρή από ότι στην (4.10.12) γιατί έστω και αν οι Y_t και W_t συσχετίζονται υψηλά δεν υπάρχουν "a priori" λόγοι να πιστεύουμε ότι θα συσχετίζονται υψηλά και οι πρώτες διαφορές τους χωρίς να αποκλείεται βέβαια και το αντίθετο. Η χρησιμοποίηση όμως των πρώτων διαφορών έχει αρκετό "κόστος": Πρώτον, δεν μπορούμε να εκτιμήσουμε την παράμετρο b_0. Δεύτερον, τα σφάλματα $u_t=e_t-e_{t-1}$ στην εξίσωση (4.10.14) θα αυτοσυσχετίζονται έστω και αν τα σφάλματα e_t της αρχικής εξίσωσης έχουν τυχαία συμπεριφορά και η αυτοσυσχέτιση αυτή δεν αντιμετωπίζεται διότι η αντιμετώπισή της σημαίνει την επιστροφή στην αρχική εξίσωση (4.10.12). Τρίτον, η χρησιμοποίηση των πρώτων διαφορών ελαττώνει τον αριθμό των παρατηρήσεων κατά μία και αυτό, σε μικρά δείγματα, ίσως δεν είναι επιθυμητό. Τέλος, οι πρώτες διαφορές ίσως δεν είναι κατάλληλες στην περίπτωση διαστρωματικών αναλύσεων όπου δεν υπάρχει συνήθως δεδομένη φυσική ή λογική διάταξη των παρατηρήσεων.

ΛΟΓΟΙ

Συχνά, αντί των πρώτων διαφορών, χρησιμοποιούνται λόγοι. Π.χ. αντί της εξίσωσης (4.10.12) μπορούμε να χρησιμοποιήσουμε την

$$\frac{Q_t}{P_t}=b_1+b_0\frac{1}{P_t}+b_2\frac{Y_t}{P_t}+\frac{e_t}{P_t} \qquad (4.10.15)$$

στην οποία όμως τα σφάλματα θα είναι ετεροσκεδαστικά έστω και αν τα e_t είναι ομοσκεδαστικά και η ετεροσκεδαστικότητα αυτή δεν αντιμετωπίζεται διότι η αντιμετώπισή της σημαίνει επιστροφή στην εξίσωση (4.10.12).

4. Αύξηση του μεγέθους του δείγματος.

Όπως ήδη τονίσαμε η πολυσυγγραμμικότητα αποτελεί πρόβλημα των στοιχείων του δείγματος και εκφράζει ακριβώς τη μη ικανοποιητική ανταπόκριση του σ υ γ κ ε κ ρ ι μ έ ν ο υ δ ε ί γ μ α τ ο ς στις πληροφοριακές απαιτήσεις του υποδείγματος. Σε ένα καινούργιο λοιπόν δείγμα ή σε ένα δείγμα που θα προκύψει από την αύξηση του μεγέθους (αν αυτό είναι δυνατό) του αρχικού δείγματος, ενδέχεται η πολυσυγγραμμικότητα να είναι λιγότερο σοβαρή αν βέβαια τα καινούργια στοιχεία είναι διαφορετικής φύσης από τα αρχικά.

5. Η μέθοδος των ορθογώνιων συνιστωσών (principal components).

Η μέθοδος αυτή, που είναι καθαρά στατιστική, έχει ως στόχο την αντικατάσταση των σχεδόν συγγραμμικών μεταβλητών $X_1, X_2, \ldots, X_k$ με ένα νέο σύνολο μεταβλητών $Z_1, Z_2, \ldots, Z_k$ οι οποίες να είναι ανά δύο γραμμικά ασυσχέτιστες (ορθογώνιες) και τέτοιες ώστε η Z_1 να έχει τη μεγαλύτερη δυνατή διακύμανση, η Z_2 τη μεγαλύτερη δυνατή διακύμανση από τις υπόλοιπες κ.ο.κ. Ας θεωρήσουμε π.χ. τη μεταβλητή (όλες οι μεταβλητές εκφράζονται σε αποκλίσεις από τους αντίστοιχους μέσους):

$$z_{1t}=\alpha_{11}x_{t1}+\alpha_{12}x_{t2}+\cdots+\alpha_{1k}x_{tk}, \quad t=1,2,\ldots,n \qquad (4.10.16)$$

αποκλίσεις από τους μέσους

ή

$$z_1=X\alpha_1 \qquad (4.10.17)$$

όπου

$$X=\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1k} \\ x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix},$$

με τον περιορισμό (γιατί αλλιώς η z_1 μπορεί να πάρει τιμές απεριόριστα μεγάλες):

$$\alpha_{11}^2+\alpha_{12}^2+\cdots+\alpha_{1k}^2=1, \quad \text{περιορισμός} \qquad (4.10.18)$$

και ας ζητήσουμε τις τιμές των $\alpha_{11}, \alpha_{12}, \ldots, \alpha_{1k}$ που μεγιστοποιούν τη διακύμανση της z_1 κάτω από τον περιορισμό (4.10.18). ΣΤΟΧΟΣ

Αποδεικνύεται[1] ότι το ζητούμενο διάνυσμα $a_1'=[\alpha_{11},\alpha_{12},\ldots,\alpha_{1k}]$ προσδιορίζεται από την εξίσωση

$$(X'X)a_1=\lambda_1 a_1 \qquad (4.10.19)$$

και συγκεκριμένα είναι το χαρακτηριστικό διάνυσμα του πίνακα $(X'X)$ που αντιστοιχεί στη μεγαλύτερη χαρακτηριστική του ρίζα λ_1 (ο πίνακας $X'X$ είναι θετικός πεπερασμένος και συνεπώς όλες οι χαρακτηριστικές του ρίζες είναι θετικές).

Ας αναζητήσουμε τώρα μία δεύτερη μεταβλητή

$$z_{2t}=\alpha_{21}x_{t1}+\alpha_{22}x_{t2}+\cdots+\alpha_{2k}x_{tk},\quad t=1,2,\ldots,n \qquad (4.10.20)$$

με τον περιορισμό

$$\alpha_{21}^2+\alpha_{22}^2+\cdots+\alpha_{2k}^2=1 \qquad (4.10.21)$$

και ας ζητήσουμε να μεγιστοποιήσουμε τη διακύμανση της z_2 κάτω από τον περιορισμό (4.10.21) και τον περιορισμό

$$a_1'a_2=0. \qquad (4.10.22)$$

Ο περιορισμός αυτός τίθεται έτσι ώστε οι μεταβλητές z_1 και z_2 να είναι γραμμικά ασυσχέτιστες (ορθογώνιες). Αποδεικνύεται[1] ότι και το διάνυσμα a_2 προσδιορίζεται από την εξίσωση

$$(X'X)a_2=\lambda_2 a_2 \qquad (4.10.23)$$

όπου λ_2 είναι τώρα η δεύτερη σε μέγεθος χαρακτηριστική ρίζα του $X'X$.

Ακολουθώντας την ίδια διαδικασία μπορούμε να βρούμε k νέες μεταβλητές $z_1,z_2,\ldots,z_k$ που να είναι ανά δύο ορθογώνιες και τέτοιες ώστε η πρώτη να έχει τη μεγαλύτερη διακύμανση, η δεύτερη την αμέσως μικρότερη κ.ο.κ. Οι μεταβλητές $z_1,z_2,\ldots,z_k$ ονομάζονται "οι k ορθογώνιες συνιστώσες" του πίνακα

1. Βλέπε σχετικά J. Johnston "Econometric Methods" 2nd Edition, McGraw-Hill, 1972, σελ. 322-331.

X και προσδιορίζονται από τα αντίστοιχα διανύσματα $a_1,a_2,\ldots,a_k$ και από τη σχέση

$$z_i=Xa_i,\quad i=1,2,\ldots,k. \qquad (4.10.24)$$

Αν θέσουμε

$$A=[a_1,a_2,\ldots,a_k] \qquad (4.10.25)$$

και

$$Z=[z_1,z_2,\ldots,z_k] \qquad (4.10.26)$$

τότε ο πίνακας X των ερμηνευτικών μεταβλητών αντικαθίσταται από τον πίνακα:

$$Z=XA \qquad (4.10.27)$$

για τον οποίο ισχύει:

$$Z'Z=A'X'XA=\begin{bmatrix}\lambda_1 & & 0\\ & \lambda_2 & \\ & & \ddots \\ 0 & & & \lambda_k\end{bmatrix}. \qquad (4.10.28)$$

Χαρακτηριστικές ρίζες.

Από την (4.10.28) προκύπτει ότι οι "k ορθογώνιες συνιστώσες" του πίνακα X είναι πράγματι γραμμικά ασυσχέτιστες και οι διακυμάνσεις τους προσδιορίζονται από την

$$z_i'z_i=\lambda_i,\quad i=1,2,\ldots,k. \qquad (4.10.29)$$

Αν $r(X)=\ell<k$ τότε $(k-\ell)$ χαρακτηριστικές ρίζες του πίνακα $X'X$ θα είναι ίσες με μηδέν και θα έχουμε ℓ μόνο ορθογώνιες συνιστώσες. Στην περίπτωση που ο πίνακας X είναι πλήρους βαθμού, αλλά μερικές από τις χαρακτηριστικές ρίζες του $X'X$ είναι κοντά στο μηδέν, ένας μικρός αριθμός ορθογώνιων συνιστωσών θα εξαντλεί σχεδόν ολόκληρη τη διακύμανση των $X_1,X_2,\ldots,X_k$ και η υπόδειξη είναι να χρησιμοποιήσουμε, αντί των $X_1,X_2,\ldots,X_k$, τις σημαντικότερες από τις $Z_1,Z_2,\ldots,Z_k$ (όχι όλες διότι τότε το αποτέλεσμα θα είναι το ίδιο).

Υπάρχουν βέβαια αρκετά προβλήματα στη χρησιμοποίηση της μεθόδου αυτής. Πρώτον, η ορθογώνια συνιστώσα Z παρά το ότι ερμηνεύει το μεγαλύτερο μέρος της συνολικής διακύμανσης των $X_1, X_2, \ldots, X_k$, δεν είναι και αναγκαστικά αυτή που συσχετίζεται υψηλότερα με την εξαρτημένη μεταβλητή Y, δηλαδή δεν υπάρχει συστηματική σχέση ανάμεσα στη σειρά των ορθογώνιων συνιστωσών και στο βαθμό συσχέτισής τους με την Y. Δεύτερον, οι $Z_1, Z_2, \ldots, Z_k$, κατά κανόνα, δεν είναι εύκολο να ερμηνευτούν οικονομικά. Ποια οικονομική ερμηνεία θα δώσουμε π.χ. στο γραμμικό συνδυασμό $\frac{2}{\sqrt{13}}Y - \frac{3}{\sqrt{13}}P$ αν Y είναι το διαθέσιμο εισόδημα και P είναι η τιμή στον προσδιορισμό της ζήτησης ενός αγαθού; Τρίτον, ο υπολογισμός των ορθογωνίων συνιστωσών στηρίζεται στην υπόθεση ότι όλες οι μεταβλητές $X_1, X_2, \ldots, X_k$ έχουν μετρηθεί με την ίδια μονάδα. Στην περίπτωση που αυτό δε συμβαίνει ο πίνακας X πρέπει να αντικατασταθεί με τον πίνακα R_X των συντελεστών απλής συσχέτισης $r_{X_iX_j}$ αλλά τότε το πρόβλημα της οικονομικής ερμηνείας των παραμέτρων γίνεται ακόμα δυσκολότερο.

Παρά τα μειονεκτήματα αυτά, η μέθοδος των ορθογώνιων συνιστωσών είναι χρήσιμη κυρίως στο προκαταρκτικό στάδιο της έρευνας. Π.χ. αν η διακύμανση της Z_1 ερμηνεύει το 80% της διακύμανσης των $X_1, X_2, \ldots, X_k$ και η διακύμανση της Z_2 το 19% από το υπόλοιπο 20%, αυτό σημαίνει ότι υπάρχουν μόνο δύο σημαντικές πηγές ανεξάρτητης διακύμανσης και επομένως η εισαγωγή περισσότερων από δύο ερμηνευτικών μεταβλητών θα δημιουργήσει προβλήματα πολυσυγγραμμικότητας. Αν μπορούμε βέβαια να ερμηνεύσουμε και οικονομικά τις Z_1 και Z_2 τότε αυτό αποτελεί πράγματι ορθή αντιμετώπιση του προβλήματος.

Η χρησιμότητα λοιπόν της μεθόδου των ορθογώνιων συνιστωσών για την αντιμετώπιση της πολυσυγγραμμικότητας είναι γενικά πολύ περιορισμένη.

(6) Η μέθοδος των Hoerl-Kennard (Ridge Regressions).

Η κεντρική ιδέα της μεθόδου αυτής είναι η εξής: επειδή στην περίπτωση σοβαρής πολυσυγγραμμικότητας ο πίνακας $X'X$

είναι "σχεδόν" ιδιάζων μπορούμε να πολλαπλασιάσουμε τα διαγώνια στοιχεία του με την ποσότητα 1+d, όπου d είναι ένας μικρός αριθμός. Η υπόδειξη των Hoerl-Kennard είναι να αρχίσουμε με μία πολύ μικρή τιμή του d (π.χ. d=0,01) και να την αυξάνουμε συνεχώς μέχρις ότου οι εκτιμήτριες Ε.Τ. γίνουν "σταθερές" ή δεν παρουσιάζουν σημαντικές μεταβολές.

Η μέθοδος αυτή παρουσιάζει γενικά πολλά προβλήματα[1] στην εφαρμογή της και δεν έχει χρησιμοποιηθεί αρκετά στις οικονομετρικές αναλύσεις.

4.11. Η ΥΠΟΘΕΣΗ (υ.5): ΣΤΟΧΑΣΤΙΚΕΣ ΕΡΜΗΝΕΥΤΙΚΕΣ ΜΕΤΑΒΛΗΤΕΣ. ΤΕΧΝΗΤΕΣ ΜΕΤΑΒΛΗΤΕΣ

Όπως διαπιστώσαμε στην παράγραφο 4.1 του κεφαλαίου 4, η υπόθεση (υ.5) υπεισέρχεται σε όλο το φάσμα της στατιστικής επαγωγής του κλασικού γραμμικού υποδείγματος, από τον υπολογισμό των μέσων και του πίνακα διακυμάνσεων-συνδιακυμάνσεων των εκτιμητριών Ε.Τ. μέχρι τους ελέγχους υποθέσεων, τον προσδιορισμό διαστημάτων εμπιστοσύνης και τη διατύπωση προβλέψεων.

Όμως, η υπόθεση ότι οι ερμηνευτικές μεταβλητές $X_1, X_2, \ldots, X_k$ δεν είναι τυχαίες (στοχαστικές) αλλά ότι οι τιμές τους παραμένουν σταθερές σε επανειλημμένα δείγματα, σπάνια ισχύει στις οικονομετρικές αναλύσεις δεδομένου ότι τα στατιστικά στοιχεία που χρησιμοποιούμε δεν προέρχονται από κατευθυνόμενα πειράματα, στα οποία έχουμε τη δυνατότητα να προσδιορίζουμε κατά βούληση -και επομένως να διατηρούμε σταθερές- τις τιμές των μεταβλητών που ελέγχουμε δηλαδή των ερμηνευτικών μεταβλητών. Χαρακτηριστική περίπτωση, στην οποία δεν ισχύει η βασική υπόθεση (υ.5), αποτελεί το γραμμικό υπόδειγμα

1. Για τα προβλήματα της μεθόδου των Hoerl-Kennard και τις ιδιότητες των εκτιμητριών που προκύπτουν από την εφαρμογή της βλέπε H.D. Vinod, "A Survey of Ridge Regressions and Related Techniques for Improvements over Ordinary Least Squares", *Review of Economics and Statistics*, Vol. 60 (1978), pp. 121-131.

$$C_t=b_0+b_1Y_t+b_2C_{t-1}+e_t$$ χαρακτηριστική περίπτωση

όπου ως ερμηνευτική μεταβλητή χρησιμοποιείται η τιμή C_{t-1} της εξαρτημένης μεταβλητής με χρονική υστέρηση μιας περιόδου η οποία βέβαια είναι τυχαία μεταβλητή αφού η C_t είναι τυχαία μεταβλητή ως συνάρτηση των τυχαίων σφαλμάτων e_t.

Αν λοιπόν οι $X_1, X_2, \ldots, X_k$ είναι τυχαίες μεταβλητές τότε οι εκτιμήτριες Ε.Τ. του κλασσικού γραμμικού υποδείγματος παύουν να είναι γενικά αμερόληπτες και δεν ικανοποιούν το θεώρημα των Gauss-Markov. Επιπλέον ο προσδιορισμός της κατανομής των εκτιμητριών αυτών για πεπερασμένα δείγματα δεν είναι πια εύκολος ούτε γενικά δυνατός, με αποτέλεσμα να μην μπορούμε να προβούμε σε στατιστική επαγωγή ή διατύπωση προβλέψεων. Αφού λοιπόν δεν μπορούμε, κατά κανόνα, να προσδιορίσουμε τις ιδιότητες των εκτιμητριών Ε.Τ. για πεπερασμένα δείγματα, περιοριζόμαστε στην αναζήτηση των ασυμπτωτικών ιδιοτήτων των εκτιμητριών αυτών δηλαδή, των ιδιοτήτων τους όταν το μέγεθος του δείγματος τείνει στο άπειρο. Η αναζήτηση των ασυμπτωτικών ιδιοτήτων καθώς και της ασυμπτωτικής κατανομής των εκτιμητριών Ε.Τ. γίνεται για τους εξής λόγους: Πρώτον, αν οι εκτιμήτριες δεν έχουν επιθυμητές ιδιότητες (π.χ. δεν είναι συνεπείς) ούτε σε μεγάλα δείγματα τότε πρέπει να τις απορρίψουμε. Δεύτερον, αν προσδιορίσουμε την ασυμπτωτική κατανομή τους και το μέγεθος του δείγματος είναι αρκετά μεγάλο (στις εφαρμογές της οικονομετρίας η προσέγγιση της ασυμπτωτικής κατανομής είναι συνήθως πολύ καλή για $n>30$ και σε πολλές περιπτώσεις για $n>20$) τότε μπορούμε να χρησιμοποιήσουμε την ασυμπτωτική κατανομή των εκτιμητριών αυτών για τη διεξαγωγή της στατιστικής επαγωγής για τις αληθείς τιμές των παραμέτρων στους πληθυσμούς.

Ας θεωρήσουμε το γραμμικό υπόδειγμα:

$$y=Xb+e \qquad (4.11.1)$$

με τις ακόλουθες υποθέσεις

(i) $$E(e)=0 \qquad (4.11.2)$$

(ii) $$E(ee')=\sigma^2 I \qquad (4.11.3)$$

(iii) $$\operatorname{plim}\left(\frac{1}{n}e'e\right)=\sigma^2. \qquad (4.11.4)$$

Η υπόθεση αυτή εξασφαλίζει ότι το κατά πιθανότητα όριο της διακύμανσης των σφαλμάτων είναι ίσο με σ^2.

(iv) Οι μεταβλητές $X_1, X_2, \ldots, X_k$ είναι τυχαίες και τέτοιες ώστε

$$\operatorname{plim}\left(\frac{1}{n}X'X\right)=M \qquad (4.11.5)$$

όπου M είναι (θετικός) πεπερασμένος πίνακας μη ιδιάζων.

Η υπόθεση αυτή εξασφαλίζει αφενός ότι οι τυχαίες μεταβλητές $X_1, X_2, \ldots, X_k$ έχουν, στο όριο, πεπερασμένες ροπές δεύτερης τάξης ως προς την αρχή των αξόνων, άρα και ως προς τους μέσους τους, και αφετέρου ότι ο πίνακας X είναι πλήρους βαθμού για κάθε μέγεθος δείγματος.

(v) Οι τυχαίες μεταβλητές $X_1, X_2, \ldots, X_k$ είναι τουλάχιστον "ταυτόχρονα ασυσχέτιστες" με τα σφάλματα, δηλαδή

ταυτόχρονα ασυσχέτιστες

$$E(X_{ij}e_i)=0, \quad i=1,2,\ldots,n, \quad j=1,2,\ldots,k. \qquad (4.11.6)$$

Η υπόθεση αυτή εξασφαλίζει [βλέπε σχετικά θεώρημα Mann-Wald, Παράρτημα Β, θεώρημα (Β.3.32)] ότι:

$$\operatorname{plim}\left(\frac{1}{n}X'e\right)=0. \qquad (4.11.7)$$

οι X είναι ορθογώνιοι (ασυσχέτιστες) με τα σφάλματα.

Επειδή οι σχετικές υποθέσεις ότι οι τυχαίες μεταβλητές $X_1, X_2, \ldots, X_k$ είναι "ασυσχέτιστες ή ανεξάρτητες" από τα σφάλματα είναι ισχυρότερες από την (4.11.6), γίνεται φανερό ότι η (4.11.7) θα ισχύει και στις περιπτώσεις αυτές και, συνεπώς, όσα θα αναφέρουμε στη συνέχεια, θα ισχύουν κατά μείζονα λόγο, όταν οι μεταβλητές $X_1, X_2, \ldots, X_k$ είναι ασυσχέτιστες ή ανεξάρτητες από τα σφάλματα.

(vi) Οι κατανομές των $X_1, X_2, \ldots, X_k$ δεν εξαρτώνται από

plim: κατά πιθανότητα όριο.

τις προς εκτίμηση παραμέτρους b και σ^2. Η υπόθεση αυτή επιτρέπει την εφαρμογή της μεθόδου Ε.Τ. για την εκτίμηση των παραμέτρων του υποδείγματος. (4.11.8)

Ας αναζητήσουμε τώρα τις ιδιότητες των εκτιμητριών ελαχίστων τετραγώνων του υποδείγματος (4.11.1) κάτω από τις νέες υποθέσεις (4.11.2) έως (4.11.8), θεωρώντας ότι οι δύο υπόλοιπες υποθέσεις (υ.6) και (υ.7) του κλασικού γραμμικού υποδείγματος εξακολουθούν να ισχύουν.

Για τις εκτιμήτριες Ε.Τ. του υποδείγματος (4.11.1) γνωρίζουμε ότι:

$$\hat{b}=b+(X'X)^{-1}X'e \qquad (4.11.9)$$

$E(\hat{b})=b$

και γίνεται αμέσως φανερό ότι η ιδιότητα της αμεροληψίας θα ισχύει μόνο αν

$$\boxed{E(e/X)=0} \quad \text{Υπόθεση 1} \qquad (4.11.10)$$

δηλαδή αν η υπό συνθήκη μαθηματική ελπίδα του e, δοθέντος του πίνακα X, είναι 0 για κάθε μέγεθος δείγματος. Πράγματι, από την (4.11.9) έχουμε:

$$\begin{aligned} E(\hat{b}) &= b+E[(X'X)^{-1}X']\cdot E(e/X) \\ &= b, \quad \text{λόγω της (4.11.10).} \end{aligned} \qquad (4.11.11)$$

Αν επιπλέον ισχύει

$$\boxed{E(ee'/X)=\sigma^2 I} \quad \text{Υπόθεση 2} \qquad (4.11.12)$$

τότε

$$\begin{aligned} C(\hat{b}) &= E(\hat{b}-b)(\hat{b}-b)' \\ &= E[(X'X)^{-1}X'ee'X(X'X)^{-1}] \\ &= E\{E[(X'X)^{-1}X'ee'X(X'X)^{-1}]/X\} \\ &= E[(X'X)^{-1}X'(\sigma^2 I)X(X'X)^{-1}] \\ &= \sigma^2 E(X'X)^{-1} \end{aligned} \qquad (4.11.13)$$

όπου βέβαια έχουμε χρησιμοποιήσει το γεγονός ότι η μαθημα-

βάσιμη υπό συνθήκη πιθανότητα διότι ο X είναι στοχαστικός.
Πριν δεν ήταν.

τική ελπίδα μιας τυχαίας μεταβλητής μπορεί να γραφτεί ως η μαθηματική ελπίδα των υπό συνθήκη μαθηματικών της ελπίδων.

Όμοια, για το άθροισμα των τετραγώνων των καταλοίπων της παλινδρόμησης προκείπτει ότι:

$$\begin{aligned} E(\hat{e}'\hat{e}) &= E[E(\hat{e}'\hat{e})/X]=E[\sigma^2[n-(k+1)]] \\ &= \sigma^2[n-(k+1)], \end{aligned} \qquad (4.11.14)$$

έτσι ώστε για την εκτιμήτρια της διακύμανσης:

$$\hat{\sigma}^2=\frac{\hat{e}'\hat{e}}{n-(k+1)} \qquad (4.11.15)$$

να ισχύει:

$$\begin{aligned} E(\hat{\sigma}^2) &= E[E(\hat{\sigma}^2/X)] \\ &= E\left[E\left[\frac{\hat{e}'\hat{e}}{n-(k+1)}\,\middle|\,X\right]\right] \\ &= \sigma^2 \end{aligned} \qquad (4.11.16)$$

και για την εκτιμήτρια του πίνακα συνδιακυμάνσεων του $\hat{b}$:

$$\begin{aligned} E[\widehat{C(\hat{b})}] &= E[E[\widehat{C(\hat{b})}]/X] \\ &= E[E[\hat{\sigma}^2(X'X)^{-1}]/X] \\ &= \sigma^2 E(X'X)^{-1}. \end{aligned} \qquad (4.11.17)$$

Αν οι τυχαίες μεταβλητές $X_1, X_2, \ldots, X_k$ είναι ανεξάρτητες από τα σφάλματα τότε οι (4.11.10) και (4.11.12) ισχύουν κατά μείζονα λόγο, άρα ισχύουν και όλα τα συμπεράσματα που αποδείξαμε πιο πάνω.

Συμπερασματικά λοιπόν μπορούμε να πούμε ότι:

ΣΥΜΠΕΡΑΣΜΑ ΓΙΑ ΤΗΝ ΑΜΕΡΟΛΗΨΙΑ

Αν στο κλασικό γραμμικό υπόδειγμα οι μεταβλητές $X_1, X_2, \ldots, X_k$ είναι τυχαίες αλλά ικανοποιούν τις (4.11.10) και (4.11.12) ή, κατά μείζονα λόγο, είναι ανεξάρτητες από τα σφάλματα, τότε, οι εκτιμήτριες ελαχίστων τετραγώνων $\hat{b}$, $\hat{\sigma}^2$ και $C(\hat{b})$ εξακολουθούν να είναι αμερόληπτες. (4.11.18)

Σημειώνουμε βέβαια ότι αντί του πίνακα $(X'X)^{-1}$ υπεισέρχεται τώρα ο πίνακας $E(X'X)^{-1}$. Σημειώνουμε ακόμη ότι οι εκτιμήτριες Ε.Τ. $\hat{b}=(X'X)^{-1}X'y$ δεν είναι πλέον γραμμικές συναρτήσεις των τιμών της Y αλλά μη γραμμικές συναρτήσεις όλων των τυχαίων μεταβλητών $X_1, X_2, \ldots, X_k$ και Y με συνέπεια οι εκτιμήτριες ελαχίστων τετραγώνων να μην ικανοποιούν το θεώρημα των Gauss-Markov δηλαδή να μην είναι οι "άριστες γραμμικές αμερόληπτες εκτιμήτριες (BLUE) του διανύσματος b. Πέρα όμως από το γεγονός αυτό, όταν οι μεταβλητές $X_1, X_2, \ldots, X_k$ είναι ανεξάρτητες από τα σφάλματα και οι κατανομές τους δεν εξαρτώνται από τις παραμέτρους b και σ^2, τότε οι εκτιμήτριες Ε.Τ. ταυτίζονται με τις εκτιμήτριες Μέγιστης Πιθανοφάνειας (να αποδειχτεί ως άσκηση) με αποτέλεσμα αυτές να είναι ασυμπτωτικά αποτελεσματικές. Άρα όταν το 1 μέγεθος του δείγματος είναι αρκετά μεγάλο τότε όλα τα συμπεράσματα της στατιστικής επαγωγής του κλασικού γραμμικού υποδείγματος ισχύουν αρκεί 2 οι τυχαίες μεταβλητές $X_1, X_2, \ldots, X_k$ να είναι ανεξάρτητες από τα σφάλματα και 3 οι κατανομές τους να μην εξαρτώνται από τις παραμέτρους b και σ^2. Ακόμη μπορούμε, στην περίπτωση αυτή, να δείξουμε ότι οι εκτιμήτριες $\hat{b}$, $\hat{\sigma}^2$ και $C(\hat{b})$ είναι συνεπείς εκτιμήτριες των b, σ^2 και $C(\hat{b})$ αντίστοιχα.

μη γραμμικές

συνεπείς

Ας εξετάσουμε τώρα τις ιδιότητες των εκτιμητριών Ε.Τ. όταν οι $X_1, X_2, \ldots, X_k$ είναι τυχαίες και ικανοποιούν την ασθενέστερη υπόθεση (4.11.6) δηλαδή είναι ταυτόχρονα ασυσχέτιστες με τα σφάλματα. Στην περίπτωση αυτή οι μεταβλητές $X_1, X_2, \ldots, X_k$ συσχετίζονται γενικά με τα σφάλματα και συνεπώς οι εκτιμήτριες Ε.Τ. παύουν να είναι αμερόληπτες.

ταυτόχρονα συσχετίζεται.

Ας εξετάσουμε τη συνέπεια των εκτιμητριών Ε.Τ. Από την (4.11.9) προκύπτει ότι:

$$\operatorname{plim}\hat{b} = b + \operatorname{plim}\left(\frac{1}{n}X'X\right)^{-1}\left(\frac{1}{n}X'e\right)$$

σύγκλιση κατά πιθανότητα

$$= b + \operatorname{plim}\left(\frac{1}{n}X'X\right)^{-1}\operatorname{plim}\left(\frac{1}{n}X'e\right), \quad \text{[θεώρημα (Β.3.28)]}$$

$$= b + M^{-1}\cdot 0 \quad \text{[λόγω των (4.11.5) και (4.11.7)]}$$

$$= b \qquad (4.11.19)$$

όροι

i) μεγάλο δείγμα

ii) οι τυχαίες μετ. $X_1, \ldots, X_k$ ανεξ. από σφάλματα

iii) ... σ^2

} ασυμπτωτικά αποτελεσματικές

άρα οι εκτιμήτριες Ε.Τ. είναι συνεπείς εκτιμήτριες των αληθών τιμών τους στον πληθυσμό.

Η κατανομή των εκτιμητριών $\hat{b}$, για πεπερασμένα δείγματα, είναι αδύνατον να προσδιοριστεί αν δε γνωρίζουμε τις επιμέρους κατανομές των $X_1, X_2, \ldots, X_k$. Αλλά, και στην περίπτωση που οι κατανομές αυτές είναι γνωστές, η κατανομή του $\hat{b}$ είναι αφενός δύσκολο να προσδιοριστεί και αφετέρου είναι διαφορετική για κάθε μέγεθος δείγματος. Απομένει λοιπόν να αναζητήσουμε τουλάχιστον την ασυμπτωτική κατανομή των εκτιμητριών $\hat{b}$.

Ας θεωρήσουμε την ακολουθία των τυχαίων διανυσμάτων:

$$\sqrt{n}(\hat{b}-b) = \left(\frac{1}{n}X'X\right)^{-1}\left(\frac{X'e}{\sqrt{n}}\right). \qquad (4.11.20)$$

Από τις υποθέσεις του υποδείγματος και το θεώρημα των Mann-Wald προκύπτει ότι:

$$\frac{X'e}{\sqrt{n}} \underset{a}{\sim}^{1} N(0, \sigma^2 M)$$

ασυμπτωτική κατανομή.

και από το θεώρημα του Grammer (Β.3.31) ότι:

$$\left(\frac{1}{n}X'X\right)^{-1}\left(\frac{X'e}{\sqrt{n}}\right) \underset{a}{\sim} N(0, \sigma^2 M^{-1}MM^{-1}) = N(0, \sigma^2 M^{-1}).$$

Άρα

$$\sqrt{n}(\hat{b}-b) \underset{a}{\sim} N(0, \sigma^2 M^{-1}) \qquad (4.11.21)$$

και επομένως η ασυμπτωτική κατανομή των εκτιμητριών Ε.Τ. είναι:

$$\hat{b} \underset{a}{\sim} N\left(b, \frac{1}{n}\sigma^2 M^{-1}\right) \qquad (4.11.22)$$

και είναι φανερό ότι οι εκτιμήτριες $\hat{b}$ είναι ασυμπτωτικά αμερόληπτες και η διακύμανσή τους τείνει στο μηδέν.

1. Υπενθυμίζουμε ότι το σύμβολο $\underset{a}{\sim}$ δηλώνει την ασυμπτωτική κατανομή.

Όταν το μέγεθος του δείγματος είναι αρκετά μεγάλο, τότε για τους σκοπούς της στατιστικής επαγωγής μπορούμε να χρησιμοποιήσουμε την ασυμπτωτική κατανομή (4.11.22) η οποία, για μέγεθος δείγματος n, είναι η

$$\hat{b} \sim N\left[b, \frac{1}{n}\sigma^2\left(\frac{1}{n}X'X\right)^{-1}\right] = N[b, \sigma^2(X'X)^{-1}]. \qquad (4.11.23)$$

ΕΦΑΡΜΟΓΗ

Ως εφαρμογή των όσων αναφέραμε πιο πάνω, ας θεωρήσουμε το υπόδειγμα

$$C_t = b_0 + b_1 Y_t + b_2 C_{t-1} + e_t, \quad t=1,2,\ldots,n$$

όπου C_t είναι κατανάλωση και Y_t είναι το διαθέσιμο εισόδημα στην περίοδο t. Είναι φανερό ότι η C_{t-1} είναι τυχαία μεταβλητή αφού είναι η τιμή της τυχαίας μεταβλητής C_t στην προηγούμενη χρονική περίοδο. Επειδή η C_{t-1} εξαρτάται από το σφάλμα e_{t-1} και η C_t εξαρτάται από την C_{t-1}, έπεται ότι η C_t εξαρτάται από το σφάλμα e_{t-1} καθώς επίσης και από όλα τα προηγούμενα σφάλματα e_{t-2}, e_{t-3} κλπ. Η συσχέτιση βέβαια αυτή θα είναι μικρή αν η κύρια ερμηνευτική μεταβλητή είναι το διαθέσιμο εισόδημα. Παρόλα αυτά, αν τα σφάλματα είναι ασυσχέτιστα μεταξύ τους, δεν έχουμε κανένα λόγο να πιστεύουμε ότι η C_{t-1} συσχετίζεται με το σφάλμα e_t παρά το ότι συσχετίζεται με τα σφάλματα e_{t-1}, e_{t-2}, κλπ. Άρα οι τυχαίες μεταβλητές C_{t-1} και e_t είναι "ταυτόχρονα ασυσχέτιστες". Αν δεχτούμε το ίδιο και για το διαθέσιμο εισόδημα Y_t τότε πληρούνται οι υποθέσεις που αναφέραμε πιο πάνω και οι εκτιμήτριες $\hat{b}_0$, $\hat{b}_1$ και $\hat{b}_2$ θα είναι συνεπείς και η ασυμπτωτική τους κατανομή είναι η (4.11.22).

ΟΛΟ.

ΜΕΘΟΔΟΣ ΤΕΧΝΗΤΩΝ ΜΕΤΑΒΛΗΤΩΝ

Αν τώρα μια ή περισσότερες από τις ερμηνευτικές μεταβλητές $X_1, X_2, \ldots, X_k$ συσχετίζονται με τα σφάλματα, τότε

$$\text{plim}\left(\frac{1}{n}X'e\right) \neq 0 \qquad (4.11.24)$$

και οι εκτιμήτριες Ε.Τ. του κλασικού γραμμικού υποδείγματος είναι όχι μόνο μη αμερόληπτες αλλά και ασυνεπείς. Στην περίπτωση αυτή για την εκτίμηση του υποδείγματος (4.11.1) μπορούμε να χρησιμοποιήσουμε τη μέθοδο των "τεχνητών μεταβλητών":

Ας θεωρήσουμε και πάλι το γραμμικό υπόδειγμα $y=Xb+e$ για το οποίο ισχύουν όλες οι υποθέσεις (4.11.1) έως (4.11.8), εκτός από τις (4.11.6) και (4.11.7) οι οποίες αντικαθίστανται από την (4.11.24). Ας υποθέσουμε ότι, εκτός από τις παρατηρήσεις για τις μεταβλητές Y και $X_1, X_2, \ldots, X_k$, διαθέτουμε παρατηρήσεις και για ένα σύνολο μεταβλητών $Z_1, Z_2, \ldots, Z_k$ οι οποίες είναι τουλάχιστον ταυτόχρονα ασυσχέτιστες με τα σφάλματα έτσι ώστε

$$\boxed{\text{plim}\left(\frac{1}{n}Z'e\right) = 0,} \qquad (4.11.25)$$

όπου Z είναι ο πίνακας των παρατηρήσεων πάνω στις μεταβλητές $Z_1, Z_2, \ldots, Z_k$ (στον πίνακα Z περιλαμβάνεται και η μοναδιαία στήλη για το σταθερό όρο). Επιπλέον υποθέτουμε ότι οι μεταβλητές $Z_1, Z_2, \ldots, Z_k$ συσχετίζονται όσο το δυνατόν υψηλότερα με τις μεταβλητές $X_1, X_2, \ldots, X_k$, έτσι, ώστε

$$\text{plim}\left(\frac{1}{n}Z'X\right) = M_{ZX} \qquad (4.11.26)$$

και ότι

$$\text{plim}\left(\frac{1}{n}Z'Z\right) = M_{ZZ} \qquad (4.11.27)$$

όπου M_{ZX} και M_{ZZ} είναι πίνακες πεπερασμένοι και πλήρους βαθμού (μη ιδιάζοντες).

Ας θεωρήσουμε τώρα τις αποκαλούμενες "εκτιμήτριες τεχνητών μεταβλητών"[1]

$$\boxed{\tilde{b} = (Z'X)^{-1} Z'y.} \qquad (4.11.28)$$

1. Ο όρος "τεχνητές μεταβλητές" είναι μετάφραση του αγγλικού όρου "Instrumental Variables". Η πιστότερη απόδοση του αγγλικού όρου θα ήταν "βοηθητικές μεταβλητές" αλλά υιοθετούμε και εδώ την ονομασία "τεχνητές μεταβλητές" διότι έχει χρησιμοποιηθεί σε αρκετά ελληνικά εγχειρίδια.

Οι εκτιμήτριες (4.11.28) προκύπτουν, αν αντί των κανονικών εξισώσεων

$$X'X\hat{b}=X'y$$

χρησιμοποιήσουμε τις εξισώσεις

$$Z'X\tilde{b}=Z'y$$

που προκύπτουν αν πολλαπλασιάσουμε το υπόδειγμα (4.11.1) από αριστερά με τον πίνακα Z' αντί να πολλαπλασιάσουμε με τον πίνακα X' για να πάρουμε τις εκτιμήτριες Ε.Τ. Και στις δύο περιπτώσεις παραλείπουμε τους όρους $X'e$ και $Z'e$ γιατί οι μεν μεταβλητές $X_1,X_2,\dots,X_k$ παραμένουν σταθερές σε επανειλημμένα δείγματα (άρα είναι ασυσχέτιστες με τα σφάλματα), ενώ, σύμφωνα με την (4.11.25), οι μεταβλητές $Z_1,Z_2,\dots,Z_k$ είναι, στο όριο, επίσης ασυσχέτιστες με τα σφάλματα. Τονίζουμε ότι, αν μερικές από τις μεταβλητές $X_1,X_2,\dots,X_k$ είναι ασυσχέτιστες με τα σφάλματα, αυτές πρέπει να τις διατηρήσουμε και να αναζητήσουμε "τεχνητές μεταβλητές" μόνο για εκείνες από τις $X_1,X_2,\dots,X_k$ που συσχετίζονται με τα σφάλματα.

Εύκολα προκύπτει ότι

$$\tilde{b}=b+(Z'X)^{-1}Z'e$$

και

$$\text{plim}\tilde{b}=b+\text{plim}\left(\frac{1}{n}Z'X\right)^{-1}\text{plim}\left(\frac{1}{n}Z'e\right)$$

$$=b+M_{zx}0$$

$$=b \qquad (4.11.29)$$

δηλαδή οι εκτιμήτριες τεχνητών μεταβλητών είναι συνεπείς.

Ακόμη,

$$\sqrt{n}(\tilde{b}-b)=\left(\frac{1}{n}Z'X\right)^{-1}\frac{Z'e}{\sqrt{n}}$$

και, κάτω από τις υποθέσεις που έχουμε κάνει,

$$\frac{Z'e}{\sqrt{n}}\underset{\alpha}{\sim}N(0,\sigma^2M_{zz}) \quad \text{(θεώρημα Mann-Wald)}$$

και

$$\sqrt{n}(\tilde{b}-b)\underset{\alpha}{\sim}N(0,\sigma^2M_{zx}^{-1}M_{zz}M_{xz}^{-1}) \quad \text{(θεώρημα Grammer).} \qquad (4.11.30)$$

Άρα οι εκτιμήτριες τεχνητών μεταβλητών είναι συνεπείς και ακολουθούν ασυμπτωτικά την κανονική κατανομή (4.11.30). Αν λοιπόν το μέγεθος του δείγματος είναι σχετικά μεγάλο τότε μπορούμε, για τις ανάγκες της στατιστικής επαγωγής, να δεχτούμε ότι:

$$\tilde{b}\sim N\left[b,\sigma^2\frac{1}{n}\left[\left(\frac{Z'X}{n}\right)^{-1}\left(\frac{Z'Z}{n}\right)\left(\frac{X'Z}{n}\right)^{-1}\right]\right] \qquad (4.11.31)$$

$$\sim N[b,\sigma^2(Z'X)^{-1}(Z'Z)(X'Z)^{-1}]$$

$$\sim N[b,\sigma^2[X'Z(Z'Z)^{-1}Z'X]^{-1}]$$

ή

$$\tilde{b}\sim N[b,\sigma^2(X'QX)^{-1}] \qquad (4.11.32)$$

όπου

$$\boxed{Q=Z(Z'Z)^{-1}Z'} \quad \text{Τελεστής τεχνητών μεταβλητών.} \qquad (4.11.33)$$

είναι ο "τελεστής" των τεχνητών μεταβλητών.

Στις εφαρμογές, η πραγματική δυσκολία είναι να βρούμε τις τεχνητές μεταβλητές $Z_1,Z_2,\dots,Z_k$. Τα σφάλματα e δεν είναι παρατηρήσιμα και συνεπώς είναι δύσκολο να βεβαιωθούμε ότι είναι πράγματι ασυσχέτιστα με τις τεχνητές μεταβλητές. Επιπλέον, οι τεχνητές μεταβλητές $Z_1,Z_2,\dots,Z_k$ πρέπει να συσχετίζονται όσο το δυνατόν υψηλότερα με τις αντίστοιχες μεταβλητές $X_1,X_2,\dots,X_k$ γιατί, διαφορετικά, οι εκτιμήτριες $\tilde{b}$ των τεχνητών μεταβλητών θα είναι μεν συνεπείς αλλά όχι αποτελεσματικές όπως εύκολα προκύπτει από την (4.11.31). Παρόλα αυτά, η μέθοδος των τεχνητών μεταβλητών χρησιμοποιείται, όπως θα δούμε παρακάτω, με επιτυχία σε αρκετές περιπτώσεις.

4.12. Η ΥΠΟΘΕΣΗ (υ.6): ΣΦΑΛΜΑΤΑ ΣΤΗ ΜΕΤΡΗΣΗ ΤΩΝ ΜΕΤΑΒΛΗΤΩΝ

Κατά την εκτίμηση των παραμέτρων του κλασικού γραμμικού υποδείγματος έχουμε υποθέσει ότι η εξαρτημένη μεταβλητή Y και οι ερμηνευτικές μεταβλητές $X_1, X_2, \ldots, X_k$, έχουν μετρηθεί χωρίς σφάλματα. Πρέπει λοιπόν να εξετάσουμε τις συνέπειες για τις εκτιμήτριες του κλασικού γραμμικού υποδείγματος όταν μία ή περισσότερες μεταβλητές έχουν μετρηθεί με σφάλμα.

Σφάλμα στην Y (όχι σοβαρό)

Ας εξετάσουμε πρώτα την περίπτωση που μόνο η εξαρτημένη μεταβλητή Y έχει μετρηθεί με σφάλμα. Έστω λοιπόν ότι η αληθής σχέση που συνδέει την Y με τις $X_1, X_2, \ldots, X_k$ είναι η

$$y_\alpha = Xb + e \tag{4.12.1}$$

και ότι αντί του διανύσματος y_α των πραγματικών τιμών της εξαρτημένης μεταβλητής παρατηρούμε το διάνυσμα

$$y = y_\alpha + u \tag{4.12.2}$$

όπου u είναι το διάνυσμα των σφαλμάτων μέτρησης των αντίστοιχων παρατηρήσεων της Y. Αντικαθιστόντας την (4.12.2) στην (4.12.1) έχουμε

$$\boxed{y = Xb + (e + u)} \tag{4.12.3}$$

και είναι φανερό ότι αν τα σφάλματα u είναι ασυσχέτιστα με τα σφάλματα e και $E(u) = 0$ και $E(uu') = \sigma_u^2 I$, τότε το υπόδειγμα (4.12.3) μπορεί να εκτιμηθεί με τη μέθοδο Ε.Τ. κατά τα γνωστά θέτοντας $v = e + u$. Αν οι μεταβλητές $X_1, X_2, \ldots, X_k$ είναι τυχαίες τότε πρέπει τα σφάλματα μέτρησης u να είναι τουλάχιστον ταυτόχρονα ασυσχέτιστα με τις $X_1, X_2, \ldots, X_k$. Συμπεραίνουμε λοιπόν ότι αν τα σφάλματα μέτρησης της Y δεν είναι συστηματικά τότε μπορούμε να τα ενσωματώσουμε στα σφάλματα e του υποδείγματος και να εκτιμήσουμε τις παραμέτρους του υποδείγματος με τη μέθοδο Ε.Τ. κατά τα γνωστά.

Σφάλμα στις X (σοβαρό)

Ας εξετάσουμε τώρα την περίπτωση που οι ερμηνευτικές μεταβλητές $X_1, X_2, \ldots, X_k$ έχουν μετρηθεί με σφάλματα.

Υποθέτουμε ότι η αληθής σχέση που συνδέει την εξαρτημένη μεταβλητή με τις ανεξάρτητες μεταβλητές είναι η

$$y = X_\alpha b + e \tag{4.12.4}$$

αλλά, αντί του πίνακα X_α των αληθών τιμών των $X_1, X_2, \ldots, X_k$, παρατηρούμε τον πίνακα

$$X = X_\alpha + V \tag{4.12.5}$$

όπου V είναι ο πίνακας των σφαλμάτων μέτρησης των παρατηρήσεων πάνω στις ερμηνευτικές μεταβλητές. Αν στην (4.12.4) αντικαταστήσουμε την (4.12.5) τότε έχουμε προς εκτίμηση τις παραμέτρους του υποδείγματος:

$$\boxed{y = Xb + (e - Vb).} \tag{4.12.6}$$

Οι εκτιμήτριες Ε.Τ. του υποδείγματος (4.12.6) είναι:

$$\begin{aligned} \hat{b} &= (X'X)^{-1}X'y \\ &= b + (X'X)^{-1}X'(e - Vb) \end{aligned} \tag{4.12.7}$$

και οι εκτιμήτριες αυτές θα είναι συνεπείς μόνο αν

$$\text{plim}\left[\frac{1}{n}X'(e - Vb)\right] = 0. \tag{4.12.8}$$

Αλλά,

$$\begin{aligned} \text{plim}\left[\frac{1}{n}X'(e - Vb)\right] &= \text{plim}\left(\frac{1}{n}X'e\right) - \text{plim}\left(\frac{1}{n}X'V\right)b \\ &= \text{plim}\left(\frac{1}{n}X'e\right) - \text{plim}\left[\frac{1}{n}(X_\alpha' + V')V\right]b \\ &= \text{plim}\left(\frac{1}{n}X'e\right) - \text{plim}\left(\frac{1}{n}X_\alpha'V\right)b - \text{plim}\left(\frac{1}{n}V'V\right)b \end{aligned} \tag{4.12.9}$$

και γίνεται αμέσως φανερό ότι, έστω και αν τα σφάλματα μέτρησης είναι ασυσχέτιστα με τις αντίστοιχες τιμές των $X_1, X_2, \ldots, X_k$, για τον τρίτο όρο της (4.12.9) ισχύει:

$$\boxed{\operatorname{plim}\left(\frac{1}{n}V'v\right)b\neq 0.} \qquad (4.12.10)$$

Άρα η (4.12.8) δεν ικανοποιείται και συνεπώς οι εκτιμήτριες Ε.Τ. του υποδείγματος (4.12.7) είναι ασυνεπείς και υποεκτιμούν συστηματικά τις αληθείς τιμές των παραμέτρων τους πληθυσμούς. (μεροληπτικές).

Για την αντιμετώπιση του σοβαρού προβλήματος των σφαλμάτων μέτρησης των ερμηνευτικών μεταβλητών υπάρχουν δύο βασικές μέθοδοι: ① η μέθοδος των τεχνητών μεταβλητών και ② η μέθοδος της μέγιστης πιθανοφάνειας.

DURBIN

Η κυριότερη μέθοδος τεχνητών μεταβλητών είναι η μέθοδος του Durbin. Σύμφωνα με τη μέθοδο αυτή οι μεταβλητές X_1, $X_2,\dots,X_k$ εκφράζονται σε αποκλίσεις από τους μέσους και οι αποκλίσεις κάθε μεταβλητής από τον αντίστοιχο μέσο διατάσσονται κατά τάξη μεγέθους. Στη συνέχεια ως τεχνητή μεταβλητή για κάθε μία από τις $X_1, X_2,\dots,X_k$ ορίζεται η μεταβλητή που εκφράζει την τάξη των παρατηρήσεων κάθε μιας από τις $X_1, X_2,\dots,X_k$. Ο Durbin έδειξε[1] ότι για μεγάλα δείγματα η σχετική αποτελεσματικότητα των εκτιμητριών αυτών των τεχνητών μεταβλητών ως προς τις εκτιμήτριες Ε.Τ. είναι περίπου 90% ενώ για δείγματα μεγέθους 20 είναι περίπου 86%.

Η μέθοδος της μέγιστης πιθανοφάνειας είναι πιο σύνθετη και θα αποφύγουμε να την παρουσιάσουμε εδώ. Οι ενδιαφερόμενοι μπορούν να βρουν την αναλυτική παρουσίαση της μεθόδου της μέγιστης πιθανοφάνειας στο εγχειρίδιο του J.Johnston Econometric methods, 2nd edition, σελ. 286-291.

4.13. Η ΥΠΟΘΕΣΗ (υ.7): $e \sim N(0,\sigma^2 I)$

Όπως έχουμε διαπιστώσει η υπόθεση αυτή είναι απαραίτητη για τους σκοπούς της στατιστικής επαγωγής σε μικρά δείγματα ενώ, αν το μέγεθος του δείγματος είναι μεγάλο, τότε, η οριακή κατανομή των εκτιμητριών Ε.Τ. είναι η κανονική κατανομή έστω και αν η υπόθεση (υ.7) δεν ισχύει και μπορούμε να χρησιμοποιήσουμε την οριακή αυτή κατανομή για τους σκοπούς της στατιστικής επαγωγής. Στις εφαρμογές έχει διαπιστωθεί ότι η προσέγγιση της οριακής κατανομής είναι πάρα πολύ καλή για δείγματα μεγέθους $n > 30$ ενώ είναι αρκετά καλή ακόμα και για δείγματα μεγέθους $n > 20$ ή και μικρότερα.

1. J. Durbin, "Errors in Variables" *Review of International Statistical Institute*, Vol. 22 (1954), pp. 23-32.

κεντρικό οριακό θεώρημα

Για τον έλεγχο της κανονικότητας των σφαλμάτων έχουν προταθεί αρκετές μέθοδοι[1]. Ο απλούστερος έλεγχος βασίζεται στις γνωστές στατιστικές

ασυμμετρία κύρτωση.

$$b_1 = \frac{\mu_3^2}{\mu_2^3} \quad \text{και} \quad b_2 = \frac{\mu_4}{\mu_2^2}$$

όπου μ_2, μ_3 και μ_4 είναι αντίστοιχα οι ροπές δεύτερης, τρίτης και τέταρτης τάξης γύρω από το μέσο. Όπως είναι γνωστό η στατιστική b_1 μετρά την ασυμμετρία (skewness) ενώ η στατιστική b_2 μετρά την κύρτωση (kurtosis) της κατανομής. Για την κανονική κατανομή γνωρίζουμε ότι $b_1=0$ και $b_2=3$. Ο έλεγχος αυτός προϋποθέτει την ανεξαρτησία των σφαλμάτων και για τη διεξαγωγή του θα χρησιμοποιήσουμε τα κατάλοιπα $\hat{e}$ της παλινδρόμησης αφού τα πραγματικά σφάλματα είναι άγνωστα.

Αν διαπιστωθεί ότι τα σφάλματα δεν ακολουθούν την κανονική κατανομή και το δείγμα δεν είναι αρκετά μεγάλο τότε οι λύσεις του προβλήματος είναι οι εξής:

i) Να χρησιμοποιήσουμε κάποια άλλη κατανομή -γάμμα, λογαριθμοκανονική κ.λ.π. για τους σκοπούς της στατιστικής επαγωγής.

ii) Να χρησιμοποιήσουμε κάποια άλλη μέθοδο εκτίμησης των παραμέτρων αντί της μεθόδου των Ε.Τ.

1. S.S. Shapiro, M.B. Wilk and H.J. Chen, "A Comparative Study of Various Test for Normality, *Journal of the American Statistical Association*, Vol. 63 (1968), pp. 1343-1372.
C.J. Huang and B.N. Bolch, "On the Testing of Regression Disturbances for Normality, *Journal of the American Statistical Association*, Vol. 69 (1974), pp. 330-335.

iii) Να μετασχηματίσουμε τα στοιχεία έτσι ώστε τα νέα σφάλματα να ακολουθούν την κανονική κατανομή.

Περισσότερες λεπτομέρειες και σχετική βιβλιογραφία μπορούν να βρουν οι ενδιαφερόμενοι στο εγχειρίδιο των Judge et al[1] και Maddala[2].

1. G. Judge et al., *The Theory and Practice of Econometrics*, Wiley, New York 1980, pp. 297-315.

2. G.S. Maddala, *Econometrics*, McGraw-Hill, New York 1977, pp. 305-319.

Κεφάλαιο 5

Υποδείγματα με χρονικές υστερήσεις

5.1. ΕΙΣΑΓΩΓΗ

Τα υποδείγματα που θα εξετάσουμε στο κεφάλαιο αυτό είναι εκείνα που περιέχουν ως ερμηνευτικές μεταβλητές τιμές με χρονική υστέρηση των ανεξάρτητων μεταβλητών ή και της εξαρτημένης μεταβλητής. Τα υποδείγματα αυτά είναι κατεξοχήν "εμπειρικά υποδείγματα" και αυτό οφείλεται στο ότι η οικονομική θεωρία είναι, κατά κανόνα, στατική και σπάνια αναφέρεται στις χρονικές υστερήσεις που εκφράζουν τη δυναμική του οικονομικού συστήματος. Έτσι, ο ερευνητής πρέπει να πειραματιστεί με διάφορες υποθέσεις για τη μορφή και τη διάρκεια των χρονικών υστερήσεων και να επιλέξει εκείνο το υπόδειγμα που ερμηνεύει καλλίτερα τη συγκεκριμένη πραγματικότητα.

Είναι γνωστό ότι οι χρονικές υστερήσεις υπεισέρχονται σε κάθε εκδήλωση της οικονομικής συμπεριφοράς διότι οι οικονομικές διαδικασίες είναι δυναμικές με βαθμιαίες προσαρμογές που υποβοηθούνται ή παρεμποδίζονται από το ψυχολογικό κλίμα, την τεχνολογική εξέλιξη και το θεσμικό πλαίσιο. Π.χ. η αντίδραση της παραγωγής σε κάποια αυξημένη ζήτηση δεν είναι άμεση: χρειάζεται χρόνος για την παραγγελία και την εγκατάσταση νέων μηχανημάτων, για την εκπαίδευση νέου προσωπικού καθώς και για την εξεύρεση των απαραίτητων χρηματοδο-

1. Η αγγλική ορολογία για τα υποδείγματα με χρονικές υστερήσεις είναι "distributed lag models".

τικών πόρων, χωρίς να παραλείπουμε το σημαντικό ρόλο του ψυχολογικού κλίματος και την οικονομική και κοινωνική πολιτική της εκάστοτε κυβέρνησης. Επίσης η αντίδραση της καταναλωτικής συμπεριφοράς σε απότομες μεταβολές των τιμών ή των εισοδημάτων δεν είναι άμεση: αν κάποιος κερδίσει ένα μεγάλο ποσό (λαχείο, προπό κλπ.) θα χρειαστεί κάποιο χρόνο για να αλλάξει τον τρόπο της ζωής του και να "μάθει" να διαχειρίζεται την περιουσία του.

Γίνεται λοιπόν φανερό ότι οι χρονικές υστερήσεις παίζουν τον κεντρικό ρόλο στην εξέλιξη των οικονομικών μεγεθών. Αυτός είναι και ο λόγος που διακρίνουμε τη "βραχυχρόνια" από τη "μακροχρόνια" συμπεριφορά και αυτός είναι ο λόγος που οι βραχυχρόνιες ελαστικότητες ως προς το εισόδημα είναι μικρότερες (απόλυτα) από τις μακροχρόνιες ελαστικότητες και η βραχυχρόνια οριακή ροπή προς κατανάλωση είναι γενικά μικρότερη από τη μακροχρόνια οριακή ροπή προς κατανάλωση.

Θα εξετάσουμε πρώτα τα (1) υποδείγματα με χρονικές υστερήσεις στις ανεξάρτητες μεταβλητές και στη συνέχεια τα (2) υποδείγματα που έχουν ως ερμηνευτικές μεταβλητές και τιμές της εξαρτημένης μεταβλητής με κάποια ή κάποιες χρονικές υστερήσεις.

5.2. ΘΕΩΡΗΤΙΚΑ ΥΠΟΔΕΙΓΜΑΤΑ ΜΕ ΧΡΟΝΙΚΕΣ ΥΣΤΕΡΗΣΕΙΣ ΤΩΝ ΑΝΕΞΑΡΤΗΤΩΝ Η ΚΑΙ ΤΗΣ ΕΞΑΡΤΗΜΕΝΗΣ ΜΕΤΑΒΛΗΤΗΣ

Εξετάζουμε την περίπτωση που έχουμε μία μόνο ερμηνευτική μεταβλητή. Η ανάλυση που ακολουθεί μπορεί να επεκταθεί εύκολα στην περίπτωση περισσότερων ανεξάρτητων μεταβλητών. Η γενική μορφή ενός υποδείγματος με χρονικές υστερήσεις στην ανεξάρτητη μεταβλητή X είναι η εξής:

$$\boxed{Y_t=\alpha+b_0X_t+b_1X_{t-1}+\cdots+b_kX_{t-k}+\cdots+e_t} \qquad (5.2.1)$$

και γίνεται φανερό ότι, στη διαμόρφωση της τιμής της Y κατά την περίοδο t, δεν επιδρά μόνο η τιμή της X στην ίδια περίοδο αλλά και τιμές της X σε προηγούμενες χρονικές περιόδους. Μερικοί από τους συντελεστές $b_0, b_1, \ldots, b_k, \ldots$ μπορεί να εί-

ναι ίσοι με μηδέν και αυτό σημαίνει ότι οι αντίστοιχες χρονικές υστερήσεις της X δεν επιδρούν στη διαμόρφωση της Y_t.

Ο αριθμός των χρονικών υστερήσεων μπορεί να είναι είτε πεπερασμένος είτε άπειρος και για να αποφύγουμε την περίπτωση μη ελεγχόμενης συμπεριφοράς της Y υποθέτουμε γενικά ότι

$$\sum_{i=0}^{\infty} b_i < +\infty . \qquad (5.2.2)$$

Ως μέση χρονική υστέρηση (mean lag ή average lag) ορίζεται ο σταθμικός μέσος όλων των χρονικών υστερήσεων με σταθμιστές τα σχετικά μεγέθη των αντίστοιχων συντελεστών b_i:

ΜΕΣΗ ΧΡΟΝΙΚΗ ΥΣΤΕΡΗΣΗ

$$\boxed{\text{Μέση χρονική υστέρηση}=\frac{\sum_{i=0}^{k} ib_i}{\sum_{i=0}^{k} b_i}=\sum_{i=0}^{k} i\,\frac{b_i}{\sum_{i=0}^{k} b_i}} \qquad (5.2.3)$$

και εκφράζει το μέσο χρονικό διάστημα που απαιτείται ώστε η επίδραση πάνω στην Y από μία μεταβολή της X (π.χ. κατά μία μονάδα) να εξαντληθεί πλήρως.

Αν ο αριθμός k των χρονικών υστερήσεων είναι πεπερασμένος τότε η εκτίμηση των παραμέτρων του υποδείγματος (5.1.1) με τη μέθοδο Ε.Τ. παρουσιάζει τις εξής πρακτικές δυσκολίες:

ΔΥΣΚΟΛΙΕΣ

(i) Χάνουμε k βαθμούς ελευθερίας αφού, αν διαθέτουμε n παρατηρήσεις, για την εκτίμηση του υποδείγματος

$$Y_t=\alpha+b_0X_t+b_1X_{t-1}+\cdots+b_kX_{t-k}+e_t \qquad (5.2.4)$$

θα χρησιμοποιηθούν μόνο οι n-k παρατηρήσεις. Αν ο αριθμός k είναι μεγάλος αυτό θα μειώσει σημαντικά τους βαθμούς ελευθερίας του δείγματος.

(ii) Συνήθως υπάρχει υψηλή πολυσυγγραμμικότητα μεταξύ των $X_t, X_{t-1}, \ldots, X_{t-k}$ με τις γνωστές συνέπειες για την αποτελεσματικότητα και τη σταθερότητα των εκτιμητριών Ε.Τ. ΠΟΛΥΣΥΓΓΡΑΜΜΙΚΟΤΗΤΑ

Για την αντιμετώπιση των δυσκολιών αυτών έχουν υποδειχθεί διάφοροι τρόποι που όλοι γενικά έχουν ως στόχο να μειώσουν τον αριθμό των προς εκτίμηση παραμέτρων. Αυτό επιτυγχάνεται με την επιβολή περιορισμών πάνω στους συντελεστές $b_0, b_1, \ldots, b_k$ ή, ισοδύναμα, με την "κατασκευή" νέων μεταβλητών

ΑΝΤΙΜΕΤΩΠΙΣΗ ΔΥΣΚΟΛΙΩΝ

Τα υποδείγματα αυτά αντιμετωπίζουν το πρόβλημα της πολυσυγγραμμικότητας, όχι όμως των β.ε.

Z_t από γραμμικούς συνδυασμούς των $X_t, X_{t-1}, \dots, X_{t-k}$. Μερικά από τα σημαντικότερα σχήματα που έχουν προταθεί είναι τα ακόλουθα:

ΑΝΤΙΜΕΤΩΠΙΣΗ ΠΟΛΥΣΥΓΓΡΑΜΜΙΚΟΤΗΤΑΣ $X_t, X_{t-1}, \dots, X_{t-k}$

1. Το αριθμητικό (ή τριγωνικό) σχήμα χρονικών υστερήσεων

Στο υπόδειγμα αυτό γίνεται, για τους συντελεστές b, η υπόθεση ότι ελαττώνονται "αριθμητικά", δηλαδή είναι διαδοχικοί όροι φθίνουσας αριθμητικής προόδου με πρώτο όρο τον $(k+1)b$, λόγο $-b$ και πλήθος $(k+1)$, (σχήμα 5.1.α):

$$b_i=(k+1-i)b, \quad 0\leqslant i\leqslant k \qquad (5.2.5)$$
$$=0 \quad , \quad i>k,$$

και αυτό σημαίνει ότι μεταβολές στην τιμή της X ασκούν τη μεγαλύτερη επίδρασή τους πάνω στην Y στην ίδια χρονική περίοδο και στη συνέχεια η επίδραση αυτή ελαττώνεται "αριθμητικά" μέχρι να μηδενιστεί.

Αν αντικαταστήσουμε την (5.2.5) στην (5.2.4) τότε προκύπτει το υπόδειγμα

$$Y_t=\alpha+b\left[\sum_{i=0}^{k}(k+1-i)X_{t-i}\right]+e_t \qquad (5.2.6)$$

ή

$$\boxed{Y_t=\alpha+bZ_t+e_t} \qquad (5.2.7)$$

όπου Z_t είναι ένας γραμμικός συνδυασμός των $X_t, X_{t-1}, \dots, X_{t-k}$. Μπορούμε ακόμα να τυποποιήσουμε τους συντελεστές b_i διαιρόντας τους με το άθροισμα $\frac{k(k+1)}{2}$ των όρων της αριθμητικής προόδου $0,1,2,\dots,k$. Π.χ. η μορφή του "τυποποιημένου" υποδείγματος (5.2.7) για $k=2,3,4$ είναι:

$$k=2 : \quad Y_t=\alpha+b^*\left(\frac{2X_t+X_{t-1}}{3}\right)+e_t, \quad b^*=3b,$$

1. Με τον όρο "τυποποιημένο" εννοούμε ότι το άθροισμα $\sum_{i=0}^{k}(k+1-i)$ είναι ίσο με τη μονάδα.

$$k=3 : \quad Y_t=\alpha+b^*\left(\frac{3X_t+2X_{t-1}+X_{t-2}}{6}\right)+e_t, \quad b^*=6b,$$

$$k=4 : \quad Y_t=\alpha+b^*\left(\frac{4X_t+3X_{t-1}+2X_{t-2}+X_{t-3}}{10}\right)+e_t, \quad b^*=10b.$$

2. Το ορθογώνιο σχήμα χρονικών υστερήσεων.

Στο υπόδειγμα αυτό όλες οι χρονικές υστερήσεις έχουν την ίδια επίδραση πάνω στην Y (σχήμα 5.1.β)

$$b_i=b, \quad i=0,1,2,\dots,k \qquad (5.2.8)$$
$$=0, \quad i>k.$$

Αν αντικαταστήσουμε την (5.2.8) στην (5.2.4) τότε προκύπτει το υπόδειγμα:

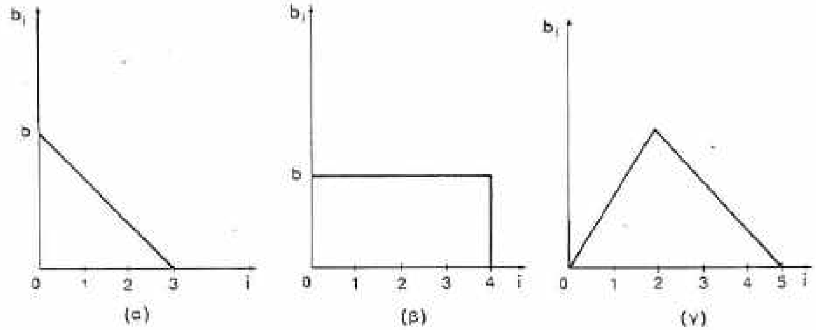
$$Y_t=\alpha+b(X_t+X_{t-1}+\dots+X_{t-k})+e_t \qquad (5.2.9)$$

ή

$$Y_t=\alpha+bZ_t+e_t \qquad (5.2.10)$$

όπου $Z_t=X_t+X_{t-1}+\dots+X_{t-k}$. Για $k=2$ π.χ. το υπόδειγμα (5.2.10) γράφεται

$$Y_t=\alpha+b(X_t+X_{t-1}+X_{t-2})+e_t$$



Σχήμα 5.1: Το "τριγωνικό", το "ορθογώνιο" και το "ανάστροφο V" σχήμα χρονικής υστέρησης.

3. Το σχήμα του "ανάστροφου V".

Στο υπόδειγμα αυτό οι συντελεστές b_i αρχικά αυξάνουν μέχρι να πάρουν κάποια μέγιστη τιμή και στη συνέχεια φθίνουν μέχρι να μηδενιστού (σχήμα 5.1.γ). Τα δύο σκέλη του ανάστροφου V μπορεί να είναι ίσα ή άνισα. Π.χ.

$$Y_t=\alpha+b\left(\frac{1}{8}X_t+\frac{1}{4}X_{t-1}+\frac{1}{2}X_{t-2}+\frac{1}{5}X_{t-3}+\frac{1}{7}X_{t-4}+\frac{1}{9}X_{t-5}\right)+e_t \tag{5.2.11}$$

ή

$$Y_t=\alpha+bZ_t+e_t \tag{5.2.12}$$

όπου

$$Z_t=\frac{1}{8}X_t+\frac{1}{4}X_{t-1}+\frac{1}{2}X_{t-2}+\frac{1}{5}X_{t-3}+\frac{1}{7}X_{t-4}+\frac{1}{9}X_{t-5}.$$

4. Το πολυωνυμικό σχήμα Almon.

Η βασική υπόθεση του υποδείγματος αυτού είναι ότι οι συντελεστές b_i είναι πλυώνυμα του i. Αν π.χ. υποθέσουμε ότι

$$b_i=\alpha_0+\alpha_1 i+\alpha_2 i^2,\quad i=0,1,2,\dots,k, \tag{5.2.13}$$

και αντικαταστήσουμε στο βασικό υπόδειγμα (5.2.4),τότε προκύπτει το υπόδειγμα του Almon στο οποίο οι συντελεστές b_i είναι συναρτήσεις δευτέρου βαθμού του i:

$$\begin{aligned}Y_t&=\alpha+\sum_{i=0}^{k}(\alpha_0+\alpha_1 i+\alpha_2 i^2)X_{t-i}+e_t\\&=\alpha+\alpha_0\sum_{i=0}^{k}X_{t-i}+\alpha_1\sum_{i=0}^{k}iX_{t-i}+\alpha_2\sum_{i=0}^{k}i^2X_{t-i}+e_t\end{aligned}$$

ή

$$\boxed{Y_t=\alpha+\alpha_0 Z_{0t}+\alpha_1 Z_{1t}+\alpha_2 Z_{2t}+e_t} \tag{5.2.14}$$

όπου

$$Z_{0t}=\sum_{i=0}^{k}X_{t-i},\quad Z_{1t}=\sum_{i=0}^{k}iX_{t-i},\quad Z_{2t}=\sum_{i=0}^{k}i^2X_{t-i}. \tag{5.2.15}$$

Η εκτίμηση του υποδείγματος (5.2.14) θα δώσει τις εκτιμήσεις των α_0, α_1 και α_2 και στη συνέχεια από τη σχέση (5.2.13) θα υπολογίσουμε τις εκτιμήσεις των συντελεστών

$$\hat{b}_i=\hat{\alpha}_0+\hat{\alpha}_1 i+\hat{\alpha}_2 i^2. \tag{5.2.16}$$

Γενικά, μπορούμε να υποθέσουμε ότι

$$b_i=\alpha_0+\alpha_1 i+\cdots+\alpha_\lambda i^\lambda$$

και να ελέγξουμε τη στατιστική σημαντικότητα των $\alpha_0,\alpha_1,\dots,\alpha_\lambda$ από το υπόδειγμα

$$Y_t=\alpha+\alpha_0 Z_{0t}+\alpha_1 Z_{1t}+\cdots+\alpha_\lambda Z_{\lambda t}+e_t \tag{5.2.17}$$

όπου

$$Z_{0t}=\sum_{i=0}^{k}X_{t-i},\quad Z_{1t}=\sum_{i=0}^{k}iX_{t-i},\dots,Z_{\lambda t}=\sum_{i=0}^{k}i^\lambda X_{t-i}. \tag{5.2.18}$$

Πρέπει να τονίσουμε ότι όλα τα υποδείγματα που αναφέραμε εμπερίεχουν αυθαιρεσία και ως προς τον αριθμό των χρονικών υστερήσεων και ως προς το σχήμα που ακολουθούν. Ακόμη πρέπει να τονίσουμε ότι ενώ αντιμετωπίζουν το πρόβλημα της πολυσυγγραμμικότητας, δεν αντιμετωπίζουν το πρόβλημα της μείωσης των βαθμών ελευθερίας.

Ας υποθέσουμε τώρα ότι ο αριθμός των χρονικών υστερήσεων της ανεξάρτητης μεταβλητής είναι άπειρος. Τότε το υπόδειγμα (5.2.1) γράφεται:

$$Y_t=\alpha+\sum_{i=0}^{\infty}b_iX_{t-i}+e_t \tag{5.2.19}$$

και γίνεται αμέσως φανερό ότι δεν μπορούμε να εκτιμήσουμε τις άπειρες παραμέτρους του υποδείγματος αυτού, από κάποιο πεπερασμένο δείγμα, χωρίς κάποιους περιορισμούς που να ελαττώνουν τον αριθμό των προς εκτίμηση συντελεστών b_i.Η συνήθως πρακτική είναι να γράψουμε το υπόδειγμα (5.2.19) ως εξής:

$$Y_t=\alpha+b\sum_{i=0}^{\infty}w_iX_{t-i}+e_t \tag{5.2.20}$$

και να ορίσουμε τις ποσότητες w_i ως πιθανότητες που ακολουθούν κάποια συγκεκριμένη κατανομή και συνεπώς θα ισχύει:

$$\sum_{i=0}^{\infty}w_i=1. \tag{5.2.21}$$

Ο συντελεστής b είναι το άθροισμα των συντελεστών όλων των χρονικών υστερήσεων και ερμηνεύεται ως η "μακροχρόνια" επίδραση της X πάνω στην Y. Ακόμα, επειδή οι συντελεστές w_i είναι πιθανότητες που ακολουθούν κάποια κατανομή, μπορούμε να υπολογίσουμε το μέσο και τη διακύμανση της κατανομής των χρονικών υστερήσεων.

Πριν αναφέρουμε τις κυριότερες κατανομές που έχουν προταθεί, θα ορίσουμε τον "τελεστή L" της χρονικής υστέρησης:

$$\boxed{LX_t=X_{t-1}.} \qquad (5.2.22)$$

Εύκολα προκύπτει ότι:

$$\boxed{L^2X_t=L(LX_t)=LX_{t-1}=X_{t-2}}$$

και γενικά

$$\boxed{L^kX_t=X_{t-k}.} \qquad (5.2.23)$$

Ακόμα, ο τελεστής L διατηρεί τη γραμμικότητα, δηλαδή,

$$L(\alpha X_t+bY_t)=\alpha X_{t-1}+bY_{t-1}. \qquad (5.2.24)$$

Αναφέρουμε στη συνέχεια τις κυριότερες κατανομές που έχουν προταθεί για τους συντελεστές w_i:

5. Το γεωμετρικό υπόδειγμα (υπόδειγμα Koyck[1]).

Στο υπόδειγμα αυτό οι συντελεστές w_i ακολουθούν τη γεωμετρική κατανομή: $P(X)=(1-\lambda)\lambda^i$

πιθανότητα $\leftarrow w_i=(1-\lambda)\lambda^i, \quad 0<\lambda<1,$ (5.2.25)

όπου

$$\sum_{i=0}^{\infty} w_i=(1-\lambda)\sum_{i=0}^{\infty}\lambda^i=(1-\lambda)\frac{1}{(1-\lambda)}=1. \qquad (5.2.26)$$

1. L.M. Koyck, *Distributed Lags and Investment Analysis*, North - Holland, Amsterdam, 1954.

X ~ Γεωμετρική κατανομή

$E(X)=\frac{\lambda}{1-\lambda}$, $V(X)=\frac{\lambda}{(1-\lambda)^2}$.

(1-λ): επιτυχία

λ: αποτυχία.

διωνυμική δοκιμή, διακριτή κατανομή

Από τη θεωρία της γεωμετρικής κατανομής γνωρίζουμε ότι η μέση τιμή και η διακύμανσή της είναι $\frac{\lambda}{1-\lambda}$ και $\frac{\lambda}{(1-\lambda)^2}$ αντίστοιχα. Άρα, οσο μεγαλύτερη είναι η τιμή της παραμέτρου λ τόσο υψηλότερη είναι η μέση χρονική υστέρηση (τόσο βραδύτερα φθίνει η κατανομή των χρονικών υστερήσεων) και η διακύμανση της κατανομής των χρονικών υστερήσεων.

Αν αντικαταστήσουμε την (5.2.25) στην (5.2.20) έχουμε το υπόδειγμα:

$$Y_t=\alpha+b(1-\lambda)\sum_{i=0}^{\infty}\lambda^iX_{t-i}+e_t \qquad (5.2.27)$$

ή, αν χρησιμοποιήσουμε τον τελεστή L, $\quad \Sigma(\lambda L)^i=\frac{1}{1-\lambda L}$

$$Y_t=\alpha+b(1-\lambda)\sum_{i=0}^{\infty}(\lambda L)^iX_t+e_t$$

$$=\alpha+\frac{b(1-\lambda)}{1-\lambda L}X_t+e_t. \qquad (5.2.28)$$

Πολλαπλασιάζοντας και τα δύο μέλη της (5.2.28) επί (1-λL) έχουμε

$$(1-\lambda L)Y_t=(1-\lambda L)\alpha+b(1-\lambda)X_t+(1-\lambda L)e_t$$

ή

$$Y_t-\lambda Y_{t-1}=\alpha(1-\lambda)+b(1-\lambda)X_t+e_t-\lambda e_{t-1},$$

από την οποία προκύπτει για εκτίμηση η εξίσωση:

$$Y_t=\alpha(1-\lambda)+b(1-\lambda)X_t+\lambda Y_{t-1}+u_t, \qquad (5.2.29)$$

όπου

$$u_t=e_t-\lambda e_{t-1}.$$

Συγκρίνοντας το υπόδειγμα ~~(5.2.28)~~ με το υπόδειγμα (5.2.29) παρατηρούμε ότι στο τελευταίο αντιμετωπίζονται με επιτυχία τα προβλήματα της πολυσυγγραμμικότητας και των βαθμών ελευθερίας αφού έχουμε να εκτιμήσουμε μόνο τις τρεις παραμέτρους α, b και λ. Πρέπει όμως να επισημάνουμε ότι αυτό δεν έχει γίνει χωρίς κόστος. Στο υπόδειγμα (5.2.29) αντιμετωπίζουμε δύο νέα σοβαρά προβλήματα:

(i) Την παρουσία της Y_{t-1} στο δεύτερο μέλος. Η Y_{t-1} είναι τυχαία μεταβλητή και είναι φανερό ότι συσχετίζεται με το σφάλμα u_t με αποτέλεσμα οι εκτιμήτριες Ε.Τ. να είναι μεροληπτικές και ασυνεπείς.

(ii) Την αυτοσυσχέτιση των σφαλμάτων u_t. Τα σφάλματα u_t ακολουθούν το σχήμα του κινητού μέσου πρώτης τάξης με αποτέλεσμα οι εκτιμήτριες Ε.Τ. να μην είναι αποτελεσματικές.

Ο συνδιασμός λοιπόν της παρουσίας της Y_{t-1} ως ερμηνευτικής μεταβλητής και της αυτοσυσχέτισης των σφαλμάτων u_t καθιστά τη μέθοδο Ε.Τ. τελείως ακατάλληλη για την εκτίμηση των παραμέτρων του υποδείγματος (5.2.29).

Βέβαια, αν το σφάλμα e_t δεν εισαχθεί στην αρχική εξίσωση (5.2.27) αλλά το προσθέσουμε στην (5.2.29) μετά το μετασχηματισμό του Koyck, τότε προκύπτει το υπόδειγμα

$$Y_t=\alpha(1-\lambda)+b(1-\lambda)X_t+\lambda Y_{t-1}+e_t \qquad (5.2.30)$$

στο οποίο η Y_{t-1} και το σφάλμα e_t είναι τώρα "ταυτόχρονα ασυσχέτιστες" τυχαίες μεταβλητές και γνωρίζουμε ότι, στην περίπτωση αυτή, οι εκτιμήτριες Ε.Τ. είναι συνεπείς. Το ίδιο θα ισχύει αν υποθέσουμε ότι τα σφάλματα e_t της (5.2.27) ακολουθούν το αυτοπαλίνδρομο σχήμα κινητού μέσου με παράμετρο λ, οπότε, τα σφάλματα της (5.2.29) θα έχουν τις γνωστές ιδιότητες και η μέθοδος Ε.Τ. θα δώσει συνεπείς εκτιμήτριες των παραμέτρων λ, α και b.

Η μέθοδος των τεχνητών μεταβλητών[1] αποτελεί μία πρώτη λύση για την εκτίμηση των παραμέτρων του υποδείγματος (5.2.29). Επειδή η X_t δεν είναι τυχαία μεταβλητή θα χρησιμοποιήσουμε τεχνητή μεταβλητή μόνο για την Y_{t-1} και ως τέτοια μπορούμε

1. Βλέπε σχετικά N. Liviatan, "Consistent Estimation of Distributed Lags", *International Economic Review*, Vol 4 (1963) pp. 44-42 και E. J. Hannan, "The Estimation of Relationships Involving Distributed Lags" *Econometrica*, Vol. 33 (1965) pp. 206-224. Βελτιώσεις της μεθόδου των τεχνητών μεταβλητών έχουν προταθεί από τους T. Amemiya and W.A. Fuller, "A Comparative Study of Alternative Estimators in a Distributed Lag Model", *Econometrica*, Vol. 35 (1967) pp. 509-529.

να χρησιμοποιήσουμε την X_{t-1} που συσχετίζεται υψηλά με την Y_{t-1} και δε συσχετίζεται με τα σφάλματα u_t. Οι εκτιμήτριες $\bar{\alpha}$, $\bar{b}$ και $\bar{\lambda}$ των τεχνητών μεταβλητών θα είναι συνεπείς αλλά δεν θα είναι αποτελεσματικές αφού δεν αντιμετωπίσαμε την αυτοσυσχέτιση των σφαλμάτων. Μπορούμε όμως να προχωρήσουμε ένα ακόμα βήμα και να χρησιμοποιήσουμε τη συνεπή εκτιμήτρια $\bar{\lambda}$ των τεχνητών μεταβλητών για να υπολογίσουμε τον πίνακα W των διακυμάνσεων-συνδιακυμάνσεων των σφαλμάτων u και να εκτιμήσουμε ξανά το υπόδειγμα (5.2.29) με τη γενικευμένη μέθοδο Ε.Τ. του Aitkens. Έτσι βελτιώνουμε την αποτελεσματικότητα των εκτιμητριών με την προϋπόθεση βέβαια ότι δεν υπάρχει υψηλή πολυσυγγραμμικότητα μεταξύ των X_t και X_{t-1}.

Άλλες μέθοδοι για την εκτίμηση των παραμέτρων του υποδείγματος (5.2.29) έχουν υποδειχθεί από τους Klein[1], Dhrymes[2], Zellner-Geisel[3] και άλλους, αλλά δε θα τις αναπτύξουμε εδώ. Οι ενδιαφερόμενοι μπορούν να συμβουλευθούν τα σχετικά άρθρα καθώς και το εγχειρίδιο των Judge et al[4].

Ας θεωρήσουμε το ακόλουθο συγκεκριμένο παράδειγμα: από στοιχεία για την Ελληνική Οικονομία (1958-1973) εκτιμήθηκε η ακόλουθη συνάρτηση για την ιδιωτική κατανάλωση:

$$C_t=2239+0.465Y_t+0.347C_{t-1} \qquad (5.2.31)$$

όπου C_t είναι η ιδιωτική κατανάλωση και Y_t το διαθέσιμο εισόδημα στην περίοδο t. Συγκρίνοντας την (5.2.31) με την (5.2.29), εύκολα διαπιστώνουμε ότι:

1. L.R. Klein, "The estimation of Distributed Lags", *Econometrica*, Vol. 26 (1958), pp. 553-565.

2. P.J. Dhrymes, *Distributed Lags: Problems of Estimation and Formulation*, Edinburgh: Oliver and Boyd, 1971 chapters 4-7.

3. A. Zellner and M.S. Geisel, "Analysis of Distributed Lag Models with Applications to Consumption Function Estimation", *Econometrica*, Vol. 38 (1971), pp. 865-888.

4. G. Judge, et al., *The Theory and Practice of Econometrics*, New York 1980, Part 6, Leads and Lags.

$$\lambda=0,347$$

$$b(1-\lambda)=0,465 \qquad (5.2.32)$$

$$b=\frac{0,465}{1-0,347}=0,712\ , \qquad (5.2.33)$$

και είναι φανερό ότι η (5.2.32) προσδιορίζει τη "βραχυχρόνια" οριακή ροπή προς κατανάλωση (0,465) ενώ η (5.2.33) προσδιορίζει τη "μακροχρόνια" οριακή ροπή προς κατανάλωση (0.712). Εξάλλου, η πραγματική σχέση ανάμεσα στην ιδιωτική κατανάλωση και το διαθέσιμο εισόδημα είναι η

$$C_t=\text{σταθ}+0,465Y_t+(0,465)\cdot(0,347)Y_{t-1}+(0,465)(0,347)^2Y_{t-2}+\cdots$$

και η μακροχρόνια οριακή ροπή προς κατανάλωση είναι το άθροισμα των συντελεστών όλων των χρονικών υστερήσεων:

$$b=0,465(1+0,347+0,347^2+\cdots)=\frac{0,465}{1-0,347}=0,712.$$

6. Το υπόδειγμα της κατανομής Pascal (υπόδειγμα Solow[1]).

Όπως προκύπτει από την ονομασία, στο υπόδειγμα αυτό οι συντελεστές w_i ακολουθούν την κατανομή του Pascal:

$$w_i=(1-\lambda)^r\binom{i+r-1}{i}\lambda^i,\quad 0<\lambda<1,\ r=1,2,\ldots \qquad (5.2.34)$$

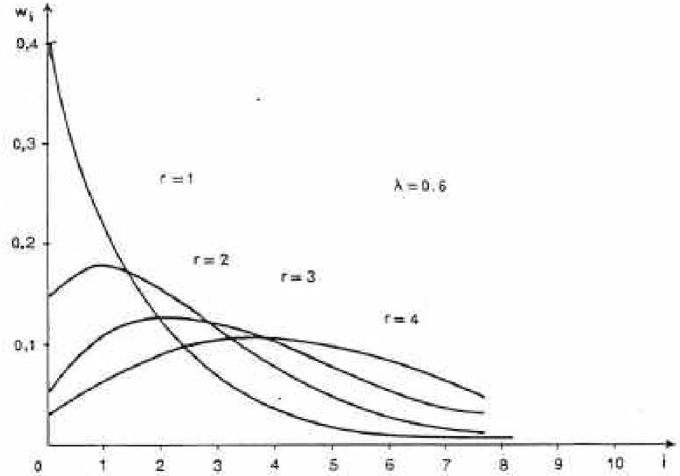
με παραμέτρους λ και r.

Στο σχήμα 5.2 φαίνεται η κατανομή των χρονικών υστερήσεων για διάφορες τιμές της παραμέτρου r και για $\lambda=0,6$. Για $r=1$ η κατανομή (5.2.34) παίρνει τη μορφή

$$w_i=(1-\lambda)\lambda^i$$

και η κατανομή Pascal είναι η γεωμετρική κατανομή που εξετάσαμε πριν.

1. R.M. Solow, "On a Family of Lag Distributions" *Econometrica*, 28 (1960), pp. 393-406.



Σχήμα 5.2: Οι μορφές της κατανομής Pascal για λ=0,6 και r=1,2,3,4

Για r=2 το γενικό υπόδειγμα της κατανομής του Pascal παίρνει τη μορφή:

$$Y_t=\alpha+b(1-\lambda)^2(X_t+2\lambda X_{t-1}+3\lambda^2X_{t-2}+4\lambda^3X_{t-3}+\cdots)+e_t$$

$$=\alpha+b(1-\lambda)^2[1+2(\lambda L)+3(\lambda L)^2+4(\lambda L)^3+\cdots]X_t+e_t$$

$$=\alpha+b(1-\lambda)^2\left[\sum_{i=0}^{\infty}(i+1)(\lambda L)^i\right]X_t+e_t$$

$$=\alpha+\frac{b(1-\lambda)^2}{(1-\lambda L)^2}X_t+e_t.$$

Συνεπώς

$$(1-\lambda L)^2Y_t=(1-\lambda L)^2\alpha+b(1-\lambda)^2X_t+(1-\lambda L)^2e_t$$

ή

$$Y_t=\alpha(1-\lambda)^2+b(1-\lambda)^2X_t+2\lambda Y_{t-1}-\lambda^2Y_{t-2}+u_t \qquad (5.2.35)$$

όπου $\quad u_t=e_t-2\lambda e_{t-1}+\lambda^2e_{t-2}.$

Το υπόδειγμα αυτό, για τους γνωστούς λόγους, δεν μπορεί να εκτιμηθεί με τη μέθοδο Ε.Τ. Εκτός από τα προβλήματα εκτίμησης που αναφέραμε στο υπόδειγμα Koyck, εδώ έχουμε και ένα ακόμα πρόβλημα: από τους συντελεστές των Y_{t-1} και Y_{t-2} θα προκύψουν δύο τιμές για την παράμετρο λ και, εκ των προτέρων, δεν έχουμε κανένα λόγο να πιστεύουμε ότι αυτές θα είναι ίσες. Το πρόβλημα αυτό είναι γνωστό ως πρόβλημα της "υπερταυτοποίησης της παραμέτρου λ" (overidentification problem).

Η απλούστερη μέθοδος εκτίμησης του υποδείγματος (5.2.35) είναι να το γράψουμε με τη μορφή

$$Y_t^*=\alpha(1-\lambda)^2+(1-\lambda)^2X_t+u_t \qquad (5.2.36)$$

όπου

$$Y_t^*=Y_t-2\lambda Y_{t-1}+\lambda^2Y_{t-2}.$$

και στη συνέχεια να το εκτιμήσουμε για διάφορες τιμές της παραμέτρου λ και να επιλέξουμε την παλινδρόμηση με το υψηλότερο R^2. Οι εκτιμήτριες των α και b θα προκύψουν από τη διαίρεση του σταθερού όρου και της κλίσης της (5.2.36) με $(1-\hat{\lambda})^2$. Παραμένει βέβαια το πρόβλημα της επιλογής της παραμέτρου r. Εντούτοις μπορούμε να εκτιμήσουμε το γενικό υπόδειγμα του Pascal

$$Y_t=\alpha+\frac{b(1-\lambda)^r}{(1-\lambda)^r}X_t+e_t \qquad (5.2.37)$$

για διάφορες τιμές του r και να επιλέξουμε την παλινδρόμηση με το υψηλότερο R^2.

Τα υποδείγματα των Koyck και Solow που εξετάσαμε είναι χαρακτηριστικά παραδείγματα υποδειγμάτων με άπειρες χρονικές υστερήσεις της ανεξάρτητης μεταβλητής που μετασχηματίζονται σε ισοδύναμα υποδείγματα με ερμηνευτικές μεταβλητές τιμές της εξαρτημένης μεταβλητής με κάποια ή κάποιες χρονικές υστερήσεις. Άλλα θεωρητικά υποδείγματα όπως το υπόδειγμα της κατανομής Γάμμα, το γεωμετρικό πολυωνυμικό υπόδειγμα, το εκθετικό υπόδειγμα, κλπ., δεν εξετάζονται εδώ. Ο ενδιαφερόμενος αναγνώστης θα βρει λεπτομερή ανάλυση των υποδειγμάτων αυτών στο εγχειρίδιο των G. Judge et al.[1].

1. G.Judge, W, Griffiths, R-C. Hill, T-C Lee, *The Theory and Practice of Econometrics*, Wiley, New York 1980, Part 6: Leads and Lags.

Πρέπει εδώ να παρατηρήσουμε ότι οι υποθέσεις (5.2.25) και (5.2.34) για την κατανομή των συντελεστών w_i είναι καθαρά μαθηματικές και περιέχουν αυθαιρεσία διότι η επιλογή τους δεν γίνεται με βάση συγκεκριμένες οικονομικές υποθέσεις. Τα υποδείγματα αυτά εντούτοις παρουσιάζουν σημαντικό ενδιαφέρον διότι, όπως θα διαπιστώσουμε παρακάτω, σε ανάλογα υποδείγματα, με ανάλογα προβλήματα εκτίμησης καταλήγουμε ξεκινόντας από καθαρά οικονομικές υποθέσεις. Τα σημαντικότερα από τα υποδείγματα αυτά εξετάζουμε παρακάτω.

5.3. ΕΜΠΕΙΡΙΚΑ ΥΠΟΔΕΙΓΜΑΤΑ ΜΕ ΧΡΟΝΙΚΕΣ ΥΣΤΕΡΗΣΕΙΣ

Από τα κυριώτερα υποδείγματα στο χώρο αυτό είναι το υπόδειγμα της 1 "Μερικής Αναπροσαρμογής", το υπόδειγμα των 2 "Αναθεωρούμενων Προσδοκιών" και το υπόδειγμα των 3 "Ρητών χρονικών υστερήσεων" τα οποία παρουσιάζουμε στη συνέχεια.

1 Το υπόδειγμα Μερικής Αναπροσαρμογής (υπόδειγμα Nerlove[1]).

Ας υποθέσουμε ότι, το "επιθυμητό" ύψος Q_t^* της παραγωγής μιας επιχείρησης στην περίοδο t, προσδιορίζεται από την εξίσωση:

$$Q_t^*=\alpha+bS_{t-1}+e_t \qquad (5.3.1)$$

όπου S_{t-1} είναι οι πωλήσεις στην προηγούμενη χρονική περίοδο. $Q_t^*-Q_{t-1}$ είναι η "επιθυμητή" αναπροσαρμογή ενώ Q_t-Q_{t-1} είναι η "πραγματοποιούμενη" αναπροσαρμογή της παραγωγής. Η υπόθεση της μερικής αναπροσαρμογής εκφράζεται από την εξίσωση:

$$\boxed{Q_t-Q_{t-1}=c(Q_t^*-Q_{t-1}),\quad 0<c<1} \qquad (5.3.2)$$

σύμφωνα με την οποία η πραγματοποιούμενη αναπροσαρμογή είναι ένα κλάσμα μόνο της επιθυμητής αναπροσαρμογής. Πώς μπο-

1. M. Nerlove, "Estimates of the Elasticities of Supply of Selected Agricultural Commodities", *Journal of Farm Economics*, Vol.38 (1956), pp. 496-509.

ρούμε όμως να αιτιολογήσουμε μια οικονομική απόφαση όπως η (5.3.2); Είναι γνωστό ότι το κόστος οποιασδήποτε αναπροσαρμογής της παραγωγής στο ύψος Q_t, που δεν είναι το επιθυμητό, συνεπάγεται αφενός το κόστος της αναπροσαρμογής και αφετέρου το κόστος λειτουργίας εκτός ισορροπίας. Αν υποθέσουμε ότι το κόστος C είναι συνάρτηση δευτέρου βαθμού, τότε μπορούμε να το εκφράσουμε με την εξίσωση

$$C_t=k(Q_t-Q_t^*)^2+\lambda(Q_t-Q_{t-1})^2, \quad k,\lambda>0, \qquad (5.3.3)$$

όπου $k(Q_t-Q_t^*)^2$ είναι το κόστος λειτουργίας εκτός ισορροπίας και $\lambda(Q_t-Q_{t-1})^2$ είναι το κόστος αναπροσαρμογής της παραγωγής από Q_{t-1} σε Q_t. Στόχος της επιχείρησης είναι να επιλέξει το ύψος της παραγωγής Q_t έτσι ώστε να ελαχιστοποιήσει το συνολικό κόστος. Εύκολα διαπιστώνουμε ότι η συνθήκη πρώτης τάξης

$$\frac{\partial C_t}{\partial Q_t}=2k(Q_t-Q_t^*)+2\lambda(Q_t-Q_{t-1})=0$$

για την ελαχιστοποίηση του κόστους, οδηγεί στη συνθήκη

$$\boxed{Q_t-Q_{t-1}=\frac{k}{k+\lambda}(Q_t^*-Q_{t-1})}$$

η οποία δε διαφέρει από την (5.3.2) και από την οποία γίνεται φανερό ότι το πραγματοποιούμενο ύψος της παραγωγής θα είναι ίσο με το επιθυμητό μόνο αν λ=0 δηλαδή μόνο αν το κόστος αναπροσαρμογής είναι μηδέν.

Αν, τώρα, αντικαταστήσουμε την (5.3.1) στην (5.3.2) και λύσουμε ως προς Q_t τότε προκύπτει η εξίσωση

ΤΕΛΙΚΗ ΜΟΡΦΗ

$$\boxed{Q_t=\alpha c+bcS_{t-1}+(1-c)Q_{t-1}+u_t} \qquad (5.3.4)$$

όπου

$$u_t=ce_t,$$

και είναι φανερό ότι η εξίσωση (5.3.4) δε διαφέρει, ως προς τη μορφή, από την εξίσωση (5.2.29) που προέκυψε από το υπόδειγμα του Koyck, εκτός από το ότι τα σφάλματα $u_t=ce_t$ είναι τώρα ασυσχέτιστα. Αν δε λάβουμε υπόψη ότι οι S_{t-1} και Q_{t-1} είναι ταυτόχρονα ασυσχέτιστες με τα σφάλματα u_t τότε η μέθοδος Ε.Τ. θα δώσει συνεπείς εκτιμήτριες των παραμέτρων α, b και c.

2. **Το υπόδειγμα των Αναθεωρούμενων Προσδοκιών (υπόδειγμα Cagan[1]).**

Μια άλλη υπόθεση με την οποία μπορούμε να αντικαταστήσουμε την (5.3.1) είναι η

$$\boxed{Q_t=\alpha+bS_t^*+e_t} \qquad (5.3.5)$$

όπου Q_t είναι η πραγματοποιούμενη παραγωγή και S_t^* οι "προσδοκόμενες" πωλήσεις στην περίοδο t. Επειδή η μεταβλητή S_t^* δεν είναι παρατηρήσιμη πρέπει να διατυπώσουμε την εξίσωση σύμφωνα με την οποία διαμορφώνονται οι προσδοκίες για το ύψος των αναμενόμενων πωλήσεων. Μια τέτοια εξίσωση εκφράζεται από την υπόθεση των "αναθεωρούμενων προσδοκιών":

$$S_t^*-S_{t-1}^*=c(S_{t-1}-S_{t-1}^*) \qquad (5.3.6)$$

σύμφωνα με την οποία η αναθεώρηση των προσδοκιών $S_t^*-S_{t-1}^*$ γίνεται με βάση το πιο πρόσφατο σφάλμα $S_{t-1}-S_{t-1}^*$, όπου S_{t-1} είναι οι πωλήσεις που πραγματοποιήθηκαν στην περίοδο t-1. Η (5.3.6) γράφεται ως εξής:

$$S_t^*-(1-c)S_{t-1}^*=cS_{t-1}. \qquad (5.3.7)$$

Από την (5.3.5) εύκολα προκύπτει ότι

$$Q_t-(1-c)Q_{t-1}=\alpha c+b[S_t^*-(1-c)S_{t-1}^*]+e_t-(1-c)e_{t-1} \qquad (5.3.8)$$

και αν λάβουμε υπόψην την (5.3.7) η (5.3.8) γράφεται:

1. P. Cagan, "The Monetary Dynamics of Hyper Inflations", in M. Friedman (ed.) *Studies of the Quantity Theory of Money*, University of Chicago Press, Chicago, 1956.

$$Q_{t-1}=\alpha+bS_{t-1}^*+e_{t-1} \Rightarrow (1-c)Q_{t-1}=\alpha(1-c)+b(1-c)S_{t-1}^*+(1-c)e_{t-1}$$

$$Q_t-(1-c)Q_{t-1}=$$

ΤΕΛΙΚΗ ΜΟΡΦΗ.

$$\boxed{Q_t=\alpha c+bcS_{t-1}+(1-c)Q_{t-1}+u_t} \tag{5.3.9}$$

όπου

$$u_t=e_t-(1-c)e_{t-1}$$

→ Βλέπε σελ. 235-237

και η εξίσωση αυτή δε διαφέρει από την (5.2.29) ούτε ως προς τη μορφή, ούτε ως προς τα προβλήματα εκτίμησης των παραμέτρων της. (όχι ολς).

ΣΥΝΔΙΑΣΜΟΣ

Μπορούμε ακόμα να συνδιάσουμε τα δύο προηγούμενα υποδείγματα υποθέτοντας ότι το επιθυμητό ύψος Q_t^* της παραγωγής συνδέεται με το ύψος S_t^* των προσδοκόμενων πωλήσεων με την εξίσωση:

$$Q_t^*=\alpha+bS_t^*+e_t \tag{5.3.10}$$

ότι η αναπροσαρμογή της παραγωγής είναι "μερική" σύμφωνα με την υπόθεση

$$Q_t-Q_{t-1}=c(Q_t^*-Q_{t-1}), \quad 0<c<1 \tag{5.3.11}$$

και ότι οι προσδοκίες για τις αναμενόμενες πωλήσεις διαμορφώνονται από την εξίσωση:

$$S_t^*-S_{t-1}^*=d(S_{t-1}-S_{t-1}^*), \quad 0<d<1 \tag{5.3.12}$$

ή, ισοδύναμα, από την εξίσωση

$$S_t^*-(1-d)S_{t-1}^*=dS_{t-1}. \tag{5.3.13}$$

Αν αντικαταστήσουμε την (5.3.10) στην (5.3.11) προκύπτει η εξίσωση:

$$Q_t=\alpha c+bcS_t^*+(1-c)Q_{t-1}+ce_t \tag{5.3.14}$$

από την οποία εύκολα προκύπτει ότι

$$Q_t-(1-d)Q_{t-1}=\alpha cd+bc[S_t^*-(1-d)S_{t-1}^*]+(1-c)[Q_{t-1}-(1-d)Q_{t-2}]+$$
$$+c[e_t-(1-d)e_{t-1}] \tag{5.3.15}$$

και ο συνδιασμός των (5.3.15) και (5.3.13) δίνει την τελική εξίσωση:

$$Q_t-(1-d)Q_{t-1}=\alpha cd+bcdS_{t-1}+(1-c)[Q_{t-1}-(1-d)Q_{t-2}]+u_t \tag{5.3.16}$$

ή,

$$Q_t=\alpha cd+bcdS_{t-1}+(1-c+1-d)Q_{t-1}-(1-c)(1-d)Q_{t-2}+u_t \tag{5.3.17}$$

όπου

$$u_t=c[e_t-(1-d)e_{t-1}]. \tag{5.3.18}$$

Τα προβλήματα εκτίμησης των παραμέτρων της εξίσωσης (5.3.17) είναι γνωστά. Στα προβλήματα αυτά πρέπει να προσθέσουμε και το ότι οι συντελεστές c και d υπερταυτοποιούνται δηλαδή δεν προσδιορίζονται μονοσήμαντα. Η εκτίμηση των παραμέτρων του υποδείγματος γίνεται, στην περίπτωση αυτή, από την (5.3.16) η οποία γράφεται

$$Q_t^+=\alpha cd+bcdS_{t-1}+(1-c)Q_{t-1}^++u_t \tag{5.3.19}$$

όπου

$$Q_t^+=Q_t-(1-d)Q_{t-1}$$

$$Q_t^+=Q_{t-1}-(1-d)Q_{t-2}.$$

Η (5.3.19) εκτιμάται για διάφορες τιμές της παραμέτρου d και επιλέγεται η τιμή της d που δίνει το υψηλότερο R^2. Οι υπόλοιποι συντελεστές c, α και b προσδιορίζονται στη συνέχεια από την (5.3.19). Για την αντιμετώπιση της αυτοσυσχέτισης των σφαλμάτων μπορούμε να προχωρήσουμε ένα βήμα ακόμα υπολογίζοντας τον πίνακα W των διακυμάνσεων-συνδιακυμάνσεων των σφαλμάτων (5.3.18) και εφαρμόζοντας τη γενικευμένη μέθοδο Ε.Τ. του Aitkens στην (5.3.19).

3. Το υπόδειγμα των ρητών χρονικών υστερήσεων (υπόδειγμα Jorgenson[1]).

Το γενικό υπόδειγμα των ρητών χρονικών υστερήσεων εί-

---

1. D.W. Jorgenson, "Rational Distributed Lag Functions", *Econometrica*, Vol. 34 (1966), pp. 135-149.

είναι της μορφής

$$Y_t = \frac{A(L)}{B(L)} X_t + e_t$$

όπου A(L) και B(L) είναι ακέραια πολυώνυμα του τελεστή L. Π.χ. αν

$$A(L) = \alpha_0 + \alpha_1 L + \alpha_2 L^2$$

$$B(L) = 1 - b_1 L - b_2 L^2$$

τότε το υπόδειγμα (5.3.20) παίρνει τη μορφή:

$$Y_t = b_1 Y_{t-1} + b_2 Y_{t-2} + \alpha_0 X_t + \alpha_1 X_{t-1} + \alpha_2 X_{t-2} + u_t \qquad (5.3.21)$$

όπου

$$u_t = e_t - b_1 e_{t-1} - b_2 e_{t-2}\,. \qquad (5.3.22)$$

Ο Jorgenson εκτίμησε το υπόδειγμα (5.3.21) με τη μέθοδο Ε.Τ. χωρίς να λάβει υπόψη την αυτοσυσχέτιση των σφαλμάτων. Η εκτίμηση των παραμέτρων του υποδείγματος των ρητών χρονικών υστερήσεων είναι πολύπλοκη και δε θα την αναφέρουμε εδώ. Οι ενδιαφερόμενοι μπορούν να συμβουλευτούν το βασικό άρθρο του Jorgenson καθώς και τα σχετικά άρθρα των Dhrymes[1], Dhrymes-Klein-Steiglitz[2] και Maddala-Rao[3].

5.4. ΕΛΕΓΧΟΣ ΑΥΤΟΣΥΣΧΕΤΙΣΗΣ ΠΡΩΤΗΣ ΤΑΞΗΣ ΣΤΑ ΥΠΟΔΕΙΓΜΑΤΑ ΜΕ ΧΡΟΝΙΚΕΣ ΥΣΤΕΡΗΣΕΙΣ ΤΗΣ ΕΞΑΡΤΗΜΕΝΗΣ ΜΕΤΑΒΛΗΤΗΣ

Πριν κλείσουμε το κεφάλαιο αυτό θα αναφερθούμε στον έλεγχο της ύπαρξης αυτοσυσχέτισης στα υποδείγματα με χρονικές υστερήσεις. Όταν δεν υπάρχουν τιμές της εξαρτημένης μεταβλητής με χρονικές υστερήσεις ως ερμηνευτικές μεταβλητές τότε δεν υπάρχει σοβαρός κίνδυνος στη χρησιμοποίηση του γνωστού κριτηρίου των Durbin-Watson για τον έλεγχο της αυτοσυσχέτισης πρώτης τάξης. Αντίθετα, το κριτήριο των Durbin - Watson, δεν ισχύει στην περίπτωση που έχουμε ως ερμηνευτικές μεταβλητές τιμές της εξαρτημένης μεταβλητής με χρονικές υστερήσεις. Για τον έλεγχο της αυτοσυσχέτισης πρώτης τάξης στα υποδείγματα αυτά ο Durbin[1] πρότεινε ένα διαφορετικό κριτήριο. Ας θεωρήσουμε π.χ. το υπόδειγμα

$$y_t = \hat{b}_1 y_{t-1} + \hat{b}_2 y_{t-2} + \hat{\alpha}_1 x_{1t} + \hat{\alpha}_2 x_{2t} + \cdots + \hat{\alpha}_k x_{kt} + e_t$$

όπου $\hat{b}_i$ και $\hat{\alpha}_i$ είναι οι εκτιμήτριες Ε.Τ. Η στατιστική που προτείνεται από τον Durbin για τον έλεγχο της αυτοσυσχέτισης πρώτης τάξης είναι η

Δεν ισχύει ο έλεγχος D.W.

$$h = \left(1 - \frac{1}{2} d^*\right) \sqrt{\frac{n}{1 - n\widehat{V(b_1)}}} = \hat{\rho} \sqrt{\frac{n}{1 - n\widehat{V(b_1)}}}$$

όπου d^* είναι η γνωστή στατιστική των D-W, n είναι το μέγεθος του δείγματος και $\widehat{V(b_1)}$ είναι η εκτίμηση της διακύμανσης της παραμέτρου $\hat{b}_1$. Η στατιστική h ακολουθεί την τυποποιημένη κανονική κατανομή N(0.1) και για τον έλεγχο της υπόθεσης ρ=0 θα χρησιμοποιηθούν οι πίνακες της κατανομής αυτής. Πρέπει να τονίσουμε ότι στην περίπτωση που ισχύει $n\widehat{V(b_1)} > 1$ -αυτό μπορεί να συμβεί όταν έχουμε λίγους βαθμούς ελευθερίας- η στατιστική h δεν ορίζεται. Για την περίπτωση αυτή ο Durbin πρότεινε ένα ασυμπτωτικά ισοδύναμο έλεγχο σύμφωνα με αυτόν εκτιμούμε την παλινδρόμηση των $\hat{e}_t$ πάνω στα $\hat{e}_{t-1}$ και στον πίνακα X όλων των ερμηνευτικών μεταβλητών (συμπεριλαμβανομένων και όλων των χρονικών υστερήσεων της Y) και ελέγχουμε τη στατιστική σημαντικότητα του συντελεστή των $\hat{e}_{t-1}$

---

1. P.J. Dhrymes, "Distributed Lags: Problems of Estimation and Formulation" Edinburg: Oliver and Boyd, 1971 Chapter 9.

2. P.J. Dhrymes, L.R. Klein and K. Steiglitz, "Estimation of Distributed Lags", *International Economic Review 11 (1970)*, pp. 235-250.

3. G.S. Maddala and A.S. Rao, "Maximum Likelihood Estimation of Solow's and Jorgenson's Distributed Lag Models", *The Review of Economics and Statistics*, February 1971, pp. 80-88.

---

1. J. Durbin, "Testing for Serial Correlation in Least Squares Regression When Some of the Regressors Are Lagged Dependent Variables", *Econometrica* 37 (1970), pp. 410-421.

χρησιμοποιόντας τη γνωστή μεθοδολογία των Ε.Τ. Επειδή ο έλεγχος αυτός του Durbin ισχύει ασυμπτωτικά, ορισμένοι ερευνητές προσπάθησαν να ελέγξουν την απόδοση του ελέγχου σε μικρά δείγματα σε σχέση με τη στατιστική h. Παρά το ότι τα αποτελέσματα των ερευνών αυτών δεν κατέληξαν σε γενικά αποδεκτά συμπεράσματα μπορούμε γενικά να δεχτούμε ότι ο ασυμπτωτικός έλεγχος του Durbin δίνει καλλίτερα αποτελέσματα και από τη στατιστική h και από τη στατιστική D-W ακόμα και σε μικρά δείγματα.

Κεφάλαιο 6

Συστήματα γραμμικών υποδειγμάτων

6.1. ΕΙΣΑΓΩΓΗ

Μέχρι τώρα έχουμε ασχοληθεί με τα προβλήματα εκτίμησης των παραμέτρων σε υποδείγματα που εκφράζονται, βασικά, με τη γραμμική εξίσωση:

$$Y_t=b_0+b_1X_{1t}+b_2X_{2t}+\cdots+b_kX_{kt}+e_t \qquad (6.1.1)$$

υποθέτοντας ότι οι μεταβολές των $X_1,X_2,\ldots,X_k$ είναι το "αίτιο" ενώ οι μεταβολές της Y είναι το "αποτέλεσμα", δηλαδή ότι η σχέση (6.1.1) είναι "μονόδρομη" και λειτουργεί μόνο από την κατεύθυνση των $X_1,X_2,\ldots,X_k$ προς την Y.

Είναι εύκολο όμως να διαπιστώσουμε ότι, σε πολλές περιπτώσεις, η υπόθεση αυτή δεν είναι σωστή, δηλαδή όχι μόνο η Y εξαρτάται από τις $X_1,X_2,\ldots,X_k$ αλλά και μερικές από τις $X_1,X_2,\ldots,X_k$ εξαρτώνται από την Y και συνεπώς υπάρχει "αμφίδρομη" σχέση μεταξύ της Y και μερικών από τις $X_1,X_2,\ldots,X_k$.

Σε μια τέτοια περίπτωση, είναι φανερό ότι, δε μπορούμε πια να μιλάμε για ανεξάρτητες μεταβλητές και εξαρτημένη μεταβλητή, αλλά για "αλληλοεξαρτημένες" μεταβλητές και η μελέτη της αλληλοεξάρτησης αυτής μπορεί να γίνει μόνο στα πλαίσια ενός συστήματος γραμμικών ή μη γραμμικών εξισώσεων.

Το εύλογο ερώτημα που τίθεται στην περίπτωση αυτή είναι το εξής: τι είδους εκτιμήτριες θα πάρουμε αν εκτιμήσουμε την εξίσωση (6.1.1), π.χ. με τη μέθοδο ελαχίστων τετραγώνων, αγνοώντας τις υπόλοιπες εξισώσεις του συστήματος, αματικότητα, τότε, δεν θα έχει πράγματι σημασία η περίοδος που αρχίζουμε την ιστορική προσομείωση. Ένας άλλος έλεγχος "ευαισθησίας" είναι η προσομείωση του υποδείγματος αφού επιφέρουμε μικρές μεταβολές στις τιμές των επιμέρους συντελεστών, διότι, μικρές μεταβολές των συντελεστών -τουλάχιστον μέχρι το 1/2 του τυπικού σφάλματος κάθε συντελεστή- δεν πρέπει να επιφέρει δραστικές μεταβολές στα αποτελέσματα της προσομείωσης. Τέλος, ένα τρίτο κριτήριο ευαισθησίας είναι η προσομείωση του υποδείγματος αφού επιφέρουμε μεταβολές στην ιστορική συμπεριφορά των εξωγενών μεταβλητών. Και πάλι, μικρές μεταβολές στην ιστορική συμπεριφορά των εξωγενών μεταβλητών δεν πρέπει να επιφέρει δραστική μεταβολή στα αποτελέσματα της ιστορικής προσομείωσης.

Όπως διαπιστώσαμε υπάρχουν αρκετά κριτήρια που μπορούν να χρησιμοποιηθούν για την αξιολόγηση της συμπεριφοράς ενός υποδείγματος, χωρίς αυτό να σημαίνει ότι δεν υπάρχουν προβλήματα στη χρησιμοποίηση όλων αυτών των κριτηρίων. Τι μπορούμε π.χ. να πούμε όταν έχουμε μία πολύ καλή "εκ των υστέρων" προσαρμογή (μικρό RMSE) αλλά η προσομείωση δεν επισημαίνει τα σημεία καμπής των ενδογενών μεταβλητών ή ακόμη όταν η ιστορική προσαρμογή είναι καλή αλλά πολύ ευαίσθητη ως προς την αρχική περίοδο που θα επιλέξουμε; Δυστυχώς, δεν υπάρχει γενικό κριτήριο για τη συνολική αξιολόγηση ενός υποδείγματος, αλλά μόνο ειδικά κριτήρια για την αξιολόγηση των επιμέρους ιδιοτήτων του. Έτσι, η κατασκευή ενός υποδείγματος είναι, ακριβώς, η "τέχνη" που επιτρέπει τον κατά περίπτωση συγκερασμό των πλεονεκτημάτων και των μειονεκτημάτων ενός υποδείγματος για την καλύτερη εξυπηρέτηση των στόχων για τους οποίους έχει κατασκευασθεί[1].

1. Μερικά πολύ χρήσιμα παιδαγωγικά παραδείγματα προσομειώσεων μπορεί ο αναγνώστης να βρει στο βιβλίο του R.S. Pindyck and D.L. Rubinfeld, *Econometric Models and Economic Forecasts*, ο.π., Κεφ. 12.

Παράρτημα Α

Στοιχεία από τον λογισμό των διανυσμάτων και των πινάκων

Α.1. ΔΙΑΝΥΣΜΑΤΑ

Ένα πραγματικό διάνυσμα x, διαστάσεων n×1, συμβολίζεται ως

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Οι πραγματικοί αριθμοί x_i, i=1,2,...,n ονομάζονται στοιχεία ή συντεταγμένες του διανύσματος. Με το συμβολισμό

$$x' = [x_1, x_2, \ldots, x_n]$$

παριστάνουμε το ίδιο διάνυσμα υπό μορφή γραμμής και οι διαστάσεις του x' είναι 1×n.

Στο σύνολο των διανυσμάτων που έχουν τις ίδιες διαστάσεις ορίζουμε τα εξής:

$$\begin{aligned} x=y &\Leftrightarrow^{*} x_i=y_i, \quad i=1,2,\ldots,n, \\ x>y &\Leftrightarrow x_i>y_i, \quad i=1,2,\ldots,n, \\ z=x+y &\Leftrightarrow z_i=x_i+y_i, \quad i=1,2,\ldots,n, \\ y=\lambda x &\Leftrightarrow y_i=\lambda x_i, \quad i=1,2,\ldots,n. \end{aligned} \tag{A.1.1}$$

* Το σύμβολο ⇔ διαβάζεται "τότε και μόνο τότε αν".

Ως εσωτερικό γινόμενο δύο n×1 διανυσμάτων x και y ορίζεται ο πραγματικός αριθμός:

$$x'y=x_1y_1+x_2y_2+\cdots+x_ny_n=\sum_{i=1}^{n} x_iy_i. \qquad (A.1.2)$$

Εύκολα προκύπτει ότι:

$$x'y=y'x$$
$$x'(y+z)=x'y+x'z \qquad (A.1.3)$$
$$x'(\lambda y)=\lambda(x'y).$$

Δύο n×1 διανύσματα x και y ονομάζονται "κάθετα (ή ορθογώνια)" τότε και μόνο τότε αν το εσωτερικό τους γινόμενο είναι μηδέν:

ΚΑΘΕΤΟΤΗΤΑ.

$$x \perp y \Leftrightarrow x'y=0. \qquad (A.1.4)$$

Το "μέτρο" $\|x\|$ ενός διανύσματος x είναι ο θετικός πραγματικός αριθμός

ΜΕΤΡΟ.

$$\|x\|=(x'x)^{1/2}=\sqrt{\sum_{i=1}^{n} x_i^2} \qquad (A.1.5)$$

και ισχύει

$$\boxed{\|\lambda x\|=|\lambda|\ \|x\|.} \qquad (A.1.6)$$

Αν $\|x\|=1$, τότε, το διάνυσμα x ονομάζεται "τυποποιημένο" διάνυσμα (normalized vector).

Αν x_i, $i=1,2,\ldots,k$ είναι k πραγματικά διανύσματα διαστάσεων n×1 και λ_i, $i=1,2,\ldots,k$ είναι πραγματικοί αριθμοί, όχι όλοι διάφοροι μεταξύ τους, τότε το διάνυσμα

$$y=\sum_{i=1}^{k}\lambda_ix_i=\lambda_1x_1+\lambda_2x_2+\cdots+\lambda_kx_k$$

ΓΡΑΜΜΙΚΟΣ ΣΥΝΔΥΑΣΜΟΣ ΔΙΑΝΥΣ.

ονομάζεται ένας "γραμμικός συνδυασμός" των διανυσμάτων $x_1, x_2, \ldots, x_k$.

Τα διανύσματα $x_1, x_2, \ldots, x_k$ είναι "γραμμικά ανεξάρτητα" μεταξύ τους τότε και μόνο τότε αν δεν υπάρχουν πραγματικοί αριθμοί (σταθερές) $\lambda_1, \lambda_2, \ldots, \lambda_k$, όχι όλοι ίσοι με μηδέν, έτσι ώστε να ισχύει:

ΑΝΕΞΑΡΤΗΣΙΑ.

$$\lambda_1x_1+\lambda_2x_2+\cdots+\lambda_kx_k=0. \qquad (A.1.7)$$

Αν υπάρχουν σταθερές $\lambda_1, \lambda_2, \ldots, \lambda_k$, όχι όλες διάφορες από το μηδέν, έτσι ώστε, να ισχύει η (A.1.7), τότε τα διανύσματα $x_1, x_2, \ldots, x_k$ είναι "γραμμικά εξαρτημένα" και στην περίπτωση αυτή γνωρίζουμε ότι ένα ή περισσότερα από τα $x_1, x_2, \ldots, x_k$ μπορούν να εκφραστούν ως γραμμικοί συνδιασμοί των υπόλοιπων, ενώ αντίθετα αν τα διανύσματα είναι γραμμικά ανεξάρτητα κανένα από τα $x_1, x_2, \ldots, x_k$ δεν μπορεί να εκφραστεί ως γραμμικός συνδιασμός των υπόλοιπων. Ακόμα, κάθε υποσύνολο γραμμικά ανεξάρτητων διανυσμάτων αποτελείται από διανύσματα που είναι επίσης γραμμικά ανεξάρτητα μεταξύ τους, ενώ, αν για το πλήθος k των n×1 διανυσμάτων $x_1, x_2, \ldots, x_k$ ισχύει $k>n$ τότε τα διανύσματα $x_1, x_2, \ldots, x_k$ είναι γραμμικά εξαρτημένα.

Ένα σύνολο n×1 πραγματικών διανυσμάτων L_n είναι ένας "διανυσματικός χώρος" τότε και μόνο τότε αν

$$\lambda(x_1+x_2)\in L_n$$

για κάθε $\lambda\in R$ και x_i, $x_j\in L_n$. Προφανώς, το σύνολο R_n όλων των n×1 πραγματικών διανυσμάτων είναι ένας διανυσματικός χώρος.

Αν $a_i\in L_n$, $i=1,2,\ldots,m$ $(m\geq n)$ και κάθε διάνυσμα $x\in L_n$ μπορεί να εκφραστεί ως γραμμικός συνδιασμός των a_i:

$$\boxed{x=\sum_{i=1}^{m}\lambda_ia_i,}$$

τότε, λέμε ότι το σύνολο των διανυσμάτων

$$\{a_1, a_2, \ldots, a_m\}$$

"καλύπτει" το διανυσματικό χώρο L_n.

Μια "βάση" στο διανυσματικό χώρο L_n είναι κάθε σύνολο με ελάχιστο πλήθος διανυσμάτων που καλύπτει τον L_n. Αν τα διανύσματα της βάσης είναι ανά δύο κάθετα μεταξύ τους τότε η βάση ονομάζεται "ορθογώνια".

Για το διανυσματικό χώρο R_n, που αναφέραμε πιο πάνω, είναι γνωστό ότι το σύνολο των διανυσμάτων $\{e_1, e_2, \ldots, e_n\}$, όπου $e_i'=(0,0,\ldots,1,\ldots,0)$ είναι το διάνυσμα που έχει όλα τα στοιχεία του ίσα με μηδέν εκτός από το στοιχείο i που είναι ίσο με τη μονάδα, ορίζει μια ορθογώνια βάση στο χώρο R_n.

Είναι φανερό ότι, αν το σύνολο των διανυσμάτων $A=\{a_1,a_2,\ldots,a_m\}$ καλύπτει το διανυσματικό χώρο L_n, τότε υπάρχει ένα υποσύνολο του A που είναι μια βάση για τον L_n. Ακόμη, αν τα στοιχεία του υποσυνόλου $\{a_1,a_2,\ldots,a_k\}$ του A $(k<m)$ είναι γραμμικά ανεξάρτητα, μπορούμε να βρούμε μια βάση του L_n που να τα περιέχει.

Μια βάση στο διανυσματικό χώρο L_n δεν ορίζεται μονοσήμαντα, όλες όμως οι βάσεις περιέχουν τον ίδιο αριθμό διανυσμάτων και κάθε διάνυσμα του L_n ορίζεται κατά ένα μόνο τρόπο ως γραμμικός συνδιασμός των διανυσμάτων μιας βάσης. Ο αριθμός των διανυσμάτων όλων των βάσεων που μπορούμε να ορίσουμε στο χώρο L_n καλείται "διάσταση" του διανυσματικού χώρου και συμβολίζεται με $d(L_n)$. Αν $d(L_n)=n$, τότε μπορούμε να δείξουμε ότι οποιαδήποτε $n+i$, $(i\geq 1)$, διανύσματα του L_n είναι γραμμικά εξαρτημένα και δεν μπορούμε να βρούμε ένα υποσύνολο του L_n με πλήθος στοιχείων μικρότερο του n που να καλύπτει τον L_n.

Α.2. ΠΙΝΑΚΕΣ: ΒΑΣΙΚΕΣ ΕΝΝΟΙΕΣ

Ένας πραγματικός πίνακας A, διαστάσεων $m\times n$, είναι ένα σύνολο $m\cdot n$ πραγματικών αριθμών διαταγμένων σε m γραμμές και n στήλες:

$$A=\begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \ldots & \alpha_{mn} \end{bmatrix}. \quad (A.2.1)$$

Οι πραγματικοί αριθμοί α_{ij} $(i=1,2,\ldots,m,\ j=1,2,\ldots,n)$ ονομάζονται στοιχεία του πίνακα A. Συχνά χρησιμοποιούμε για τον πίνακα A και το συμβολισμό $A=[\alpha_{ij}]$ για να δείξουμε ότι ο A είναι ο πίνακας με τυπικό στοιχείο το α_{ij}.

Αν m=n, δηλαδή ο αριθμός των γραμμών είναι ίσος με τον αριθμό των στηλών, τότε ο πίνακας A ονομάζεται "τετραγωνικός". τετραγωνικ

Αν ο τετραγωνικός πίνακας A είναι της μορφής

$$\begin{bmatrix} \alpha_{11} & & & \\ \alpha_{21} & \alpha_{22} & & 0 \\ \vdots & \vdots & & \\ \alpha_{n1} & \alpha_{n2} & \ldots & \alpha_{nn} \end{bmatrix} \quad \text{ή} \quad \begin{bmatrix} \alpha_{11} & \alpha_{12} & \ldots & \alpha_{1n} \\ & \alpha_{22} & & \alpha_{2n} \\ & & & \vdots \\ & 0 & & \alpha_{nn} \end{bmatrix}$$

τότε ονομάζεται "τριγωνικός". τριγωνικός

Είναι φανερό ότι κάθε $n\times 1$ διάνυσμα

$$a=\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix}$$

είναι ένας πίνακας με m γραμμές και μία στήλη. Αν λάβουμε υπόψη ότι οι n στήλες ενός $m\times n$ πίνακα ορίζουν n διανύσματα με m στοιχεία το καθένα τότε ο πίνακας A μπορεί να γραφτεί και με τη μορφή:

$$A=[a_1,a_2,\ldots,a_n]. \quad (A.2.2)$$

Στο σύνολο των $m\times n$ πινάκων ορίζουμε τα εξής:

$$\begin{aligned} A=B &\Leftrightarrow \alpha_{ij}=b_{ij} \\ A\leq B &\Leftrightarrow \alpha_{ij}\leq b_{ij} \\ C=A+B &\Leftrightarrow c_{ij}=\alpha_{ij}+b_{ij} \\ B=\lambda A &\Leftrightarrow b_{ij}=\lambda\alpha_{ij}. \end{aligned} \quad (A.2.3)$$

Ένας γραμμικός συνδιασμός B των πινάκων $A^1,A^2,\ldots,A^k$ ορίζεται ως εξής:

$$B=\lambda_1A^1+\lambda_2A^2+\cdots+\lambda_kA^k \Leftrightarrow b_{ij}=\lambda_1\alpha_{ij}^1+\lambda_2\alpha_{ij}^2+\cdots+\lambda_k\alpha_{ij}^k \quad (A.2.4)$$

όπου α_{ij}^{r} είναι το τυπικό στοιχείο του m×n πίνακα A^r, $r=1,2,\ldots,k$, και $\lambda_1,\lambda_2,\ldots,\lambda_k$ είναι σταθερές.

Από τους πιο πάνω ορισμούς και από τις ιδιότητες των πραγματικών αριθμών προκύπτουν τα εξής:

$$A+B=B+A$$ Αντιμεταθετική

$$(A+B)+C=A+(B+C)=A+B+C$$ Προσεταιριστική

$$\lambda(A+B)=\lambda A+\lambda B \qquad (A.2.5)$$

$$(\lambda+k)A=\lambda A+kA$$

$$\lambda(kA)=(\lambda k)A=(k\lambda)A=k(\lambda A).$$

Ο m×n πίνακας

$$0=\begin{bmatrix} 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & 0 \end{bmatrix}$$

ΜΗΔΕΝΙΚΟΣ ΠΙΝΑΚΑΣ
ουδέτερο στοιχείο πρόσθεσης.

ΜΗΔΕΝΙΚΟΣ

που έχει όλα τα στοιχεία του ίσα με μηδέν ονομάζεται "μηδενικός" πίνακας και είναι το ουδέτερο στοιχείο στην πρόσθεση των m×n πινάκων:

$$A+0=0+A=A.$$

ΑΝΤΙΘΕΤΟΣ

Για κάθε πίνακα A με τυπικό στοιχείο το α_{ij} ορίζεται ο αντίθετός του $-A$ με τυπικό στοιχείο το $-\alpha_{ij}$ και ισχύει:

$$A+(-A)=A-A=0.$$

ΓΙΝΟΜΕΝΟ

Ως γινόμενο του m×n πίνακα A επί τον n×q πίνακα B ορίζεται ο m×q πίνακας C του οποίου το τυπικό στοιχείο c_{ij} είναι το άθροισμα των γινομένων των στοιχείων της i γραμμής του πίνακα A επί τα στοιχεία της j στήλης του πίνακα B:

$$\boxed{C=AB \Leftrightarrow c_{ij}=\sum_{k=1}^{n}\alpha_{ik}b_{kj}} \qquad (A.2.6)$$

Αν λάβουμε υπόψη τον συμβολισμό (A.2.2) τότε η (A.2.6) γράφεται:

$$(c_1,c_2,\ldots,c_q)=A(b_1,b_2,\ldots,b_q)$$
$$=(Ab_1,Ab_2,\ldots,Ab_q), \qquad (A.2.7)$$

και

$$c_j=Ab_j$$
$$=(a_1,a_2,\ldots,a_n)b_j$$
$$=b_{1j}a_1+b_{2j}a_2+\cdots+b_{nj}a_n, \qquad (A.2.8)$$

$j=1,2,\ldots,q$, δηλαδή η j στήλη του πίνακα C είναι ένας γραμμικός συνδυασμός των στηλών του πίνακα A με σταθμιστές τα στοιχεία της j στήλης του πίνακα B.

Από τον ορισμό του γινομένου των πινάκων και τις ιδιότητες των πραγματικών αριθμών προκύπτουν τα εξής:

$$(AB)C=A(BC)=ABC$$
$$A(B+C)=AB+AC \qquad (A.2.9)$$
$$A(\lambda B)=\lambda(AB).$$

Ο n×n τετραγωνικός πίνακας

$$I_n=\begin{bmatrix} 1 & 0 & 0 & \ldots & 0 \\ 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix}$$

ΜΟΝΑΔΙΑΙΟΣ ΠΙΝΑΚΑΣ
ουδέτερο στοιχείο πολλαπλασιασμού.

που έχει όλα τα διαγώνια στοιχεία ίσα με τη μονάδα και τα μη διαγώνια στοιχεία ίσα με μηδέν ονομάζεται "μοναδιαίος" πίνακας και είναι το ουδέτερο στοιχείο του πολλαπλασιασμού των πινάκων, αν ο πολλαπλασιασμός είναι δυνατός. Ιδιαίτερα αν ο πίνακας A είναι τετραγωνικός τότε ισχύει:

$$AI=IA=A. \qquad (A.2.10)$$

Υπενθυμίζεται ότι στον πολλαπλασιασμό των πινάκων δεν ισχύει γενικά η μεταθετική ιδιότητα: το γινόμενο AB δεν είναι ίσο με το γινόμενο BA διότι, αν ορίζεται το AB, το BA μπορεί να μην ορίζεται αλλά και αν ορίζεται δεν είναι απαραίτητα ίσο με το AB.

Ο "ανάστροφος" του $m{\times}n$ πίνακα A είναι ο $n{\times}m$ πίνακας A'
ΝΑΣΤΡοφοΣ που προκύπτει από τον A με εναλλαγή των γραμμών και των στηλών του:

$$A=[\alpha_{ij}] \Leftrightarrow A'=[\alpha_{ji}]. \tag{A.2.11}$$

Για την αναστροφή των πινάκων ισχύουν οι ιδιότητες:

$$\begin{aligned} (A')' &= A \\ (A+B)' &= A'+B' \\ (AB)' &= B'A'. \end{aligned} \tag{A.2.12}$$

Ένας τετραγωνικός πίνακας λέγεται "συμμετρικός" όταν ισχύει $A=A'$, δηλαδή όταν είναι ίσος με τον ανάστροφό του.

Αν A είναι ένας $n{\times}n$ τετραγωνικός πίνακας τότε ως "ίχνος" (trace) του A ορίζεται η πραγματική συνάρτηση:

$$\mathrm{tr}(A)=\alpha_{11}+\alpha_{22}+\cdots+\alpha_{nn}=\sum_{i=1}^{n}\alpha_{ii}, \tag{A.2.13}$$

δηλαδή το άθροισμα των διαγωνίων στοιχείων του A. Εύκολα αποδεικνύονται οι εξής ιδιότητες:

$$\begin{aligned} &\mathrm{tr}(A+B)=\mathrm{tr}(A)+\mathrm{tr}(B) \\ &\mathrm{tr}(A')=\mathrm{tr}(A) \\ &\mathrm{tr}(\lambda A)=\lambda\,\mathrm{tr}(A) \\ &\mathrm{tr}(AB)=\mathrm{tr}(BA)=\sum_{i=1}^{n}\sum_{j=1}^{m}\alpha_{ij}b_{ji}, \quad (AB \text{ τετρ. πίνακας}) \\ &\mathrm{tr}(I_n)=n \\ &\mathrm{tr}\,(0)=0. \end{aligned} \tag{A.2.14}$$

Η k δύναμη ενός τετραγωνικού πίνακα A ορίζεται από την ισότητα:

$$A^k=\underbrace{A\cdot A\ldots A}. \tag{A.2.15}$$

Ισχύουν οι εξής ιδιότητες των δυνάμεων:

$$\begin{aligned} A^0 &= I \\ A^kA^\lambda &= A^{k+\lambda} \\ (A^k)^\lambda &= A^{k\lambda} \\ I_n^k &= I_n. \end{aligned} \tag{A.2.16}$$

Αν για ένα τετραγωνικό πίνακα A ισχύει $A^2=A$, τότε ο πίνακας A καλείται "αυτοδύναμος" (idempotent). Είναι προφανές ότι κάθε μοναδιαίος πίνακας είναι αυτοδύναμος. Αν ο A είναι αυτοδύναμος τότε εύκολα αποδεικνύεται ότι $A^k=A$, $k\geq 1$.

Αν τα n διανύσματα που ορίζουν οι στήλες (ή οι γραμμές) ενός $n{\times}n$ τετραγωνικού πίνακα A είναι τυποποιημένα και ανά δύο ορθογώνια τότε ο πίνακας A καλείται "ορθογώνιος". Στην περίπτωση αυτή εύκολα αποδεικνύεται ότι:

$$AA'=A'A=I_n. \tag{A.2.17}$$

Είναι φανερό ότι κάθε μοναδιαίος πίνακας I_n είναι ένας ορθογώνιος πίνακας.

Συχνά διευκολύνει τις πράξεις ο διαμερισμός ενός πίνακα σε δύο (ή περισσότερους) υποπίνακες:

$$A=[A_1 \;\; A_2]$$

όπου A είναι ένας $m{\times}n$ πίνακας και A_1 και A_2 είναι $m{\times}n_1$ και $m{\times}n_2$ πίνακες $(n_1+n_2=n)$. Ο ανάστροφος του πίνακα A γράφεται στην περίπτωση αυτή:

$$A'=[A_1 \;\; A_2]'=\begin{bmatrix} A_1' \\ A_2' \end{bmatrix}. \tag{A.2.18}$$

Αν B είναι ένας $n{\times}q$ πίνακας και θεωρήσουμε το διαμερισμό του B:

$$B=\begin{bmatrix} B_1 \\ B_2 \end{bmatrix}$$

όπου B_1 είναι $n_1{\times}q$ και B_2 είναι $n_2{\times}q$ $(n_1+n_2=n)$, τότε το γινόμενο $C=AB$ μπορεί να γραφτεί ως εξής:

$$C=AB=[A_1 \;\; A_2]\begin{bmatrix} B_1 \\ B_2 \end{bmatrix}=A_1B_1+A_2B_2 . \qquad (A.2.19)$$

Μπορούμε ακόμη να θεωρήσουμε λεπτότερες διαμερίσεις ενός $m\times n$ πίνακα A:

$$A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

όπου A_{11} είναι $m_1\times n_1$, A_{12} είναι $m_1\times n_2$, A_{21} είναι $m_2\times n_1$ και A_2 είναι $m_2\times n_2$ πίνακες ($m_1+m_2=m$ και $n_1+n_2=n$). Στην περίπτωση ενός τέτοιου διαμερισμού ο ανάστροφος του A είναι:

$$A'=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}=\begin{bmatrix} A'_{11} & A'_{21} \\ A'_{12} & A'_{22} \end{bmatrix}. \qquad (A.2.20)$$

Αν τώρα B είναι ένας $n\times q$ πίνακας και θεωρήσουμε το διαμερισμό:

$$B=\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

όπου B_{11} είναι $n_1\times q_1$, B_{12} είναι $n_1\times q_2$, B_{21} είναι $n_2\times q_1$ και B_{22} είναι $n_2\times q_2$ ($q_1+q_2=q$), τότε το γινόμενο $C=AB$ υπολογίζεται ως εξής:

$$C=\begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$$

$$=\begin{bmatrix} A_{11}B_{11}+A_{12}B_{21} & A_{11}B_{12}+A_{12}B_{22} \\ A_{21}B_{11}+A_{22}B_{21} & A_{21}B_{12}+A_{22}B_{22} \end{bmatrix}, \qquad (A.2.21)$$

όπου C_{11} είναι $m_1\times q_1$, C_{12} είναι $m_1\times q_2$, C_{21} είναι $m_2\times q_1$ και C_{22} είναι $m_2\times q_2$ πίνακες.

Το κατά "Kronecker" γινόμενο του $m\times n$ πίνακα A επί τον $p\times q$ πίνακα B ορίζεται ως εξής:

$$A\otimes B=\begin{bmatrix} \alpha_{11}B & \alpha_{12}B & \dots & \alpha_{1n}B \\ \alpha_{21}B & \alpha_{22}B & \dots & \alpha_{2n}B \\ \vdots & \vdots & & \vdots \\ \alpha_{n1}B & \alpha_{n2}B & \dots & \alpha_{nn}B \end{bmatrix}. \qquad (A.2.22)$$

Ο κατά Kronecker πολλαπλασιασμός των πινάκων έχει τις εξής ιδιότητες:

$$(A+B)\otimes C=A\otimes C+B\otimes C$$

$$A\otimes(B+C)=A\otimes B+A\otimes C$$

$$A\otimes(B\otimes C)=(A\otimes B)\otimes C \qquad (A.2.23)$$

$$(A\otimes B)'=A'\otimes B'$$

$$(A\otimes B)(C\otimes D)=AC\otimes BD$$

$$tr(A\otimes B)=tr(A)tr(B) .$$

A.3. ΟΡΙΖΟΥΣΑ ΚΑΙ ΑΝΤΙΣΤΡΟΦΟΣ ΤΕΤΡΑΓΩΝΙΚΟΥ ΠΙΝΑΚΑ

Σε κάθε τετραγωνικό πίνακα A αντιστοιχίζεται ένας πραγματικός αριθμός ο οποίος καλείται "ορίζουσα" του A και συμβολίζεται με $|A|$. Αν ο A είναι $n\times n$ τότε η $|A|$ ονομάζεται ορίζουσα τάξεως n.

Αν $n=1$ τότε $A=[\alpha]$ και $|A|=\alpha$, ενώ, αν $n=2$, τότε,

$$A=\begin{bmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{12} \end{bmatrix} \quad \text{και} \quad |A|=\begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix}=\alpha_{11}\alpha_{12}-\alpha_{21}\alpha_{12}$$

Αν $n\geqslant 2$, η ορίζουσα του $A_{n\times n}$ πίνακα μπορεί να ορισθεί με τη βοήθεια των οριζουσών των $(n-1)\times(n-1)$ υποπινάκων του A: αν "ελάσσονα ορίζουσα" ℓ_{ij} του στοιχείου α_{ij} του πίνακα A καλέσουμε την ορίζουσα του $(n-1)\times(n-1)$ υποπίνακα που προκύπτει από τον A μετά την αφαίρεση της i γραμμής και της j στήλης και "συντελεστή" c_{ij} του στοιχείου α_{ij} του A καλέσουμε το γινόμενο:

$$c_{ij}=(-1)^{i+j}\ell_{ij} \qquad (A.3.1)$$

τότε η ορίζουσα του n×n πίνακα A ορίζεται ως το άθροισμα των γινομένων των στοιχείων μιας γραμμής (ή μιας στήλης) του A επί τους αντίστοιχους συντελεστές τους:

$$|A| = \sum_{k=1}^{n} \alpha_{ik} c_{ik} \qquad (A.3.2)$$

για οποιοδήποτε i=1,2,...,n.

Αν π.χ.

$$A=\begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 5 \\ -2 & 4 & 6 \end{bmatrix}$$

τότε

$$c_{11}=(-1)^{1+1}\begin{vmatrix} 2 & 5 \\ 4 & 6 \end{vmatrix}=2\cdot 6-4\cdot 5=-8$$

$$c_{12}=(-1)^{1+2}\begin{vmatrix} 1 & 5 \\ -2 & 6 \end{vmatrix}=-[1\cdot 6-(-2)\cdot 5]=-16 \text{ κ.ο.κ.}$$

ή τελικά

$$C=[c_{ij}]=\begin{bmatrix} -8 & -16 & 8 \\ -2 & 20 & -14 \\ 7 & -6 & 1 \end{bmatrix}.$$

Σύμφωνα με την (A.3.2):

$$|A|=2(-8)+3(-16)+4\cdot 8, \text{ ή,}$$
$$=1(-2)+2\cdot 20+5(-14), \text{ ή,}$$
$$=-2\cdot 7+4(-6)+6\cdot 1$$
$$=-32 .$$

Από τον ορισμό των συντελεστών c_{ij} προκύπτει ότι:

$$\sum_{k=1}^{n} \alpha_{ik} c_{jk}=0, \quad i\neq j. \qquad (A.3.3)$$

δηλαδή, το άθροισμα των γινομένων των στοιχείων μιας γραμμής (ή μιας στήλης) του A επί τους συντελεστές των στοιχείων μιας άλλης γραμμής (ή στήλης) του A είναι πάντοτε μηδέν.

Μερικές από τις ιδιότητες των οριζουσών είναι οι εξής:

• i) Αν εναλλάξουμε δύο γραμμές (ή δύο στήλες) του πίνακα A, η ορίζουσά του αλλάζει πρόσημο.

• ii) Αν ο πίνακας A έχει δύο ίδιες γραμμές (ή στήλες), η ορίζουσά του είναι ίση με μηδέν.

• iii) Αν τα στοιχεία μιας γραμμής (ή στήλης) του πίνακα A είναι όλα μηδέν, η ορίζουσά του είναι ίση με μηδέν.

• iv) Αν τα στοιχεία μιας γραμμής (ή στήλης) του πίνακα A πολλαπλασιαστούν επί λ, η ορίζουσά του πολλαπλασιάζεται επί λ.

• v) Αν στα στοιχεία της γραμμής (ή της στήλης) k του πίνακα A προσθέσουμε τα αντίστοιχα στοιχεία της γραμμής (ή της στήλης) r πολλαπλασιασμένα επί λ, η τιμή της ορίζουσας του A δε μεταβάλλεται.

• vi) $|A|=|A'|$ (A.3.4)

• vii) $|A+B|\neq|A|+|B|$

• viii) $|AB|=|A|\;|B|$

• ix) $|I_n|=1$

• x) $|0|=0$

• xi) $|A\otimes B|=|A|^n|B|^m$, όπου $A_{n\times n}$ και $B_{m\times m}$ πίνακες τετραγωνικοί

• xii) Αν ο πίνακας A είναι διαγώνιος ή τριγωνικός τότε

$$|A|=\alpha_{11}\,\alpha_{22}\cdots\alpha_{nn}.$$

• xiii) Αν ο πίνακας A είναι ορθογώνιος τότε

$$|A|=(\mp 1)^n.$$

"Βαθμός" $r(A)$ ενός $m\times n$ πίνακα A είναι η τάξη της μεγαλύτερης μη μηδενικής ορίζουσας που περιέχεται στον A. Ισοδύναμα, ως βαθμός του A, ορίζεται ο μεγαλύτερος αριθμός των γραμμικά ανεξάρτητων γραμμών (ή στηλών) του. Είναι προφανές ότι ο βαθμός ενός πίνακα είναι ένας φυσικός αριθμός. Ισχύουν οι εξής ιδιότητες:

$$0\leqslant r(A)\leqslant \min(m,n),\ \text{όπου } A \text{ είναι ένας } m\times n \text{ πίνακας}$$

$$r(I_n)=n$$

$$r(O)=0$$

$$r(A)=r(A')=r(A'A)=r(AA')$$

$$r(A+B)\leqslant r(A)+r(B) \qquad (A.3.5)$$

$$r(AB)\leqslant \min[r(A),r(B)]$$

$$r(A\otimes B)=r(A)r(B),\ \text{αν οι } A \text{ και } B \text{ είναι τετραγωνικοί πίνακες.}$$

$$r(A)=tr(A),\quad \text{αν ο } A \text{ είναι αυτοδύναμος.}$$

Αν ο πίνακας A είναι διαγώνιος τότε $r(A)$ είναι ο αριθμός των μη μηδενικών στοιχείων της διαγωνίου του.

Ένας τετραγωνικός πίνακας $A_{n\times n}$ ονομάζεται "μη ιδιάζων" τότε και μόνο τότε αν είναι πλήρους βαθμού, δηλαδή αν

$$r(A)=n,\ \text{ή ισοδύναμα, αν}\ |A|\neq 0,$$

ενώ, αντίθετα, αν δεν είναι πλήρους βαθμού ονομάζεται "ιδιάζων", οπότε

$$r(A)<n,\ \text{ή ισοδύναμα, αν}\ |A|=0.$$

Ο βαθμός ενός πίνακα δε μεταβάλλεται αν τον πολλαπλασιάσουμε από αριστερά ή από δεξιά επί ένα μη ιδιάζοντα πίνακα. Συνεπώς, αν υπάρχουν μη ιδιάζοντες πίνακες $D_{m\times m}$ και $F_{n\times n}$ τέτοιοι ώστε

$$DAF=\begin{bmatrix} I_k & O \\ O & O \end{bmatrix} \qquad (A.3.6)$$

τότε $r(A)=k$.

Αν A είναι ένας $n\times n$ μη ιδιάζων τετραγωνικός πίνακας τότε υπάρχει ο αντίστροφός του A^{-1}, ορίζεται μονοσήμαντα και ισχύει:

$$AA^{-1}=A^{-1}A=I.$$

Ο πίνακας A^{-1} υπολογίζεται από τον τύπο:

$$A^{-1}=\frac{C'}{|A|}, \qquad (A.3.7)$$

όπου,

$$C=[c_{ij}]=[(-1)^{i+j}l_{ij}]$$

είναι ο πίνακας των συντελεστών των στοιχείων του A που ορίσαμε στην (A.3.1).

Π.χ. για τον πίνακα

$$A=\begin{bmatrix} 2 & 3 & 4 \\ 1 & 2 & 5 \\ -2 & 4 & 6 \end{bmatrix}$$

βρήκαμε ότι $|A|=-32\neq 0$. Άρα είναι μη ιδιάζων, υπάρχει ο αντίστροφός του και, σύμφωνα με την (A.3.7), είναι ο

$$A^{-1}=\frac{1}{-32}\begin{bmatrix} -8 & -2 & 7 \\ -16 & 20 & -6 \\ 8 & -14 & 1 \end{bmatrix}$$

$$=\begin{bmatrix} -\frac{1}{4} & \frac{1}{16} & -\frac{7}{32} \\ \frac{1}{2} & -\frac{5}{8} & \frac{3}{16} \\ -\frac{1}{4} & \frac{7}{16} & -\frac{1}{32} \end{bmatrix}.$$

Μερικές ιδιότητες των αντιστρόφων πινάκων είναι οι εξής:

$$\left.\begin{array}{l} I^{-1}=I \\ (A^{-1})^{-1}=A \\ (A')^{-1}=(A^{-1})' \\ |A^{-1}|=|A|^{-1}=\frac{1}{|A|} \\ (AB)^{-1}=B^{-1}A^{-1} \end{array}\right\} \quad (A.3.8)$$

$A^{-1}=A'$, τότε και μόνο τότε αν ο A είναι οθρογώνιος και στην περίπτωση αυτή οι A^{-1} και A' είναι επίσης ορθογώνιοι.

Αν $A=A'$ τότε και $(A^{-1})'=A^{-1}$.

Αν ο A είναι διαγώνιος, τότε και ο A^{-1} είναι διαγώνιος και τα στοιχεία του A^{-1} είναι τα αντίστροφα των αντίστοιχων στοιχείων του A.

$(A\otimes B)^{-1}=A^{-1}\otimes B^{-1}$, αν οι A και B είναι τετραγωνικοί πίνακες.

Για το διαμερισμένο πίνακα

$$A=\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

ο αντίστροφος A^{-1} ορίζεται από τη σχέση:

$$A^{-1}=\begin{bmatrix} B^{-1} & -B^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}B^{-1} & A_{22}^{-1}(I+A_{21}B^{-1}A_{12}A_{22}^{-1}) \end{bmatrix} \quad (A.3.9)$$

όπου,

$B=A_{11}-A_{12}A_{22}^{-1}A_{21}$ και οι A_{22} και B είναι μη ιδιάζοντες πίνακες.

Αν A είναι ένας μη αρνητικός ($\alpha_{ij}\geqslant 0$, $i,j=1,2,\ldots,n$) τετραγωνικός πίνακας, τότε ο αντίστροφος του πίνακα (I_n-A)

είναι επίσης μη αρνητικός τότε και μόνο τότε αν όλες οι ελάσσονες ορίζουσες του A είναι θετικές:

$$|\alpha_{11}|>0,\ \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix}>0,\ldots,\ \begin{vmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \alpha_{nn} \end{vmatrix}>0 \quad (A.3.10)$$

Στην περίπτωση αυτή ο αντίστροφος του (I_n-A) μπορεί να γραφτεί ως ανάπτυγμα των δυνάμεων του A:

$$(I-A)^{-1}=I+A+A^2+\cdots \quad (A.3.11)$$

Δύο τετραγωνικοί πίνακες $A_{n\times n}$ και $B_{n\times n}$ καλούνται <u>"όμοιοι"</u> (similar) τότε και μόνο τότε αν υπάρχει μη ιδιάζων πίνακας $C_{n\times n}$ τέτοιος ώστε

$$B=C^{-1}AC\,. \quad (A.3.12)$$

Μερικές χρήσιμες ιδιότητες των ομοίων πινάκων A και B είναι οι εξής:

$$\left.\begin{array}{l} \mathrm{tr}(A)=\mathrm{tr}(B) \\ |A|=|B| \\ \mathrm{r}(A)=\mathrm{r}(B) \\ B^k=C^{-1}A^kC,\ \text{όπου } B^k \text{ και } A^k \text{ είναι οι} \\ \qquad k \text{ δυνάμεις των } A \text{ και } B \end{array}\right\} \quad (A.3.13)$$

Α.4. ΓΡΑΜΜΙΚΑ ΣΥΣΤΗΜΑΤΑ

Ένα γραμμικό σύστημα m εξισώσεων με n αγνώστους:

$$\begin{array}{l} \alpha_{11}x_1+\alpha_{12}x_2+\cdots+\alpha_{1n}x_n=c_1 \\ \alpha_{21}x_1+\alpha_{22}x_2+\cdots+\alpha_{2n}x_n=c_2 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ \alpha_{m1}x_1+\alpha_{m2}x_2+\cdots+\alpha_{mn}x_n=c_m \end{array} \quad (A.4.1)$$

γράφεται με τη βοήθεια των πινάκων και των διανυσμάτων ως εξής:

$$Ax=c \qquad (A.4.2)$$

όπου $A=[a_{ij}]$ είναι ο $m\times n$ πίνακας των συντελεστών, x είναι το $n\times 1$ διάνυσμα των αγνώστων και c είναι το $m\times 1$ διάνυσμα των σταθερών όρων.

<u>Το σύστημα (A.4.2) έχει λύση τότε και μόνο τότε αν</u>

$$r(A)=r(A:c)=k \qquad (A.4.3)$$

και η λύση είναι μοναδική τότε και μόνο τότε αν $k=n$. Αν το σύστημα έχει λύση αλλά $k<n$, τότε, $n-k$ από τους αγνώστους μπορούν να οριστούν αυθαίρετα ενώ οι υπόλοιποι k άγνωστοι ορίζονται μονοσήμαντα.

Επειδή $r(A)\leqslant \min(m,n)$, το σύστημα μπορεί να μην έχει μοναδική λύση και στην περίπτωση που ο αριθμός των εξισώσεων είναι μικρότερος από τον αριθμό των αγνώστων $(m<n)$ αλλά $r(A)=r(A:c)=k\leqslant m$. Και στην περίπτωση αυτή οι $n-k$ από τους αγνώστους προσδιορίζονται αυθαίρετα και οι υπόλοιποι ορίζονται, στη συνέχεια, μονοσήμαντα.

Αν ο πίνακας A είναι τετραγωνικός (ο αριθμός των εξισώσεων είναι ίσος με τον αριθμό των αγνώστων) και μη ιδιάζων (οι m εξισώσεις είναι ανεξάρτητες), δηλαδή, αν

$$m=n=r(A)$$

τότε η λύση προσδιορίζεται μονοσήμαντα από την εξίσωση

$$x=A^{-1}c. \qquad (A.4.4)$$

Οι τιμές των αγνώστων μπορούν επίσης να προσδιοριστούν με τον κανόνα του <u>Cramer:</u>

$$x_i=\frac{|A_i|}{|A|}, \quad i=1,2,\ldots,n, \qquad (A.4.5)$$

όπου $|A_i|$ είναι η ορίζουσα του πίνακα που προκύπτει από τον A αν αντικαταστήσουμε την i στήλη με το διάνυσμα c των σταθερών όρων.

Αν στο σύστημα (A.4.2) αντικαταστήσουμε το $m\times 1$ διάνυσμα c των σταθερών όρων με το $m\times 1$ μηδενικό διάνυσμα τότε προκύπτει το αντίστοιχο <u>"ομογενές σύστημα":</u>

$$Ax=0 \qquad (A.4.6)$$

το οποίο έχει πάντα τη λύση $x=0$. Ικανή και αναγκαία συνθήκη για να έχει το ομογενές σύστημα (A.4.6) και άλλη λύση, εκτός από τη μηδενική, είναι η

$$r(A)=k<n.$$

Προφανώς η μη μηδενική λύση δεν ορίζεται μονοσήμαντα αφού $n-k$ από τους αγνώστους μπορούν να οριστούν αυθαίρετα. Αν όμως $r(A)=n-1$ τότε η λύση είναι μοναδική με την έννοια ότι αν προσδιορίσουμε αυθαίρετα την τιμή ενός αγνώστου τότε οι τιμές των υπολοίπων ορίζονται μονοσήμαντα.

Η μέθοδος "Ελαχίστων Τετραγώνων" εισάγει μια ιδιαίτερη μέθοδο για την επίλυση ενός γραμμικού συστήματος στην περίπτωση που ο αριθμός των εξισώσεων m είναι μεγαλύτερος από τον αριθμό των αγνώστων n $(m>n)$ και οι m εξισώσεις δεν είναι συμβιβαστές.

<u>Ας θεωρήσουμε το γενικό γραμμικό υπόδειγμα:</u>

$$y=Xb+e \qquad (A.4.7)$$

όπου e είναι το διάνυσμα των αποκλίσεων των στοιχείων του διανύσματος y από τα αντίστοιχα στοιχεία του διανύσματος Xb. Η επίλυση του συστήματος (A.4.7) με τη μέθοδο ελαχίστων τετραγώνων συνίσταται στην επιλογή του διανύσματος b έτσι ώστε να ελαχιστοποιείται το άθροισμα των τετραγώνων των αποκλίσεων e_i:

$$\Sigma e_i^2=e'e=(y-Xb)'(y-Xb). \qquad (A.4.8)$$

Αν ο πίνακας X είναι πλήρους βαθμού [$r(X)=k+1$, όπου $k+1$

είναι ο αριθμός των στηλών του] τότε η λύση ελαχίστων τετραγώνων του συστήματος (A.4.7) ορίζεται μονοσήμαντα και είναι η

$$\hat{b}=(X'X)^{-1}X'y \qquad (A.4.9)$$

και εύκολα προκύπτει για τις εκτιμήσεις $\hat{e}_i$ των αποκλίσεων ότι:

$$\begin{aligned}\hat{e}&=y-X\hat{b}=y-X(X'X)^{-1}X'y\\&=[I-X(X'X)^{-1}X']y\\&=My\end{aligned} \qquad (A.4.10)$$

και

$$\hat{e}'\hat{e}=y'M'My=y'M^2y=yMy \qquad (A.4.11)$$

διότι για τον "στοιχειώδη πίνακα" M των ελαχίστων τετραγώνων εύκολα αποδεικνύεται ότι:

$$\left.\begin{array}{ll}M=M' & (\text{συμμετρικός})\\ M^2=M & (\text{αυτοδύναμος})\\ r(M)=tr(M)=n-(k+1) & \\ MX=0 & \end{array}\right\}. \qquad (A.4.12)$$

A.5. ΧΑΡΑΚΤΗΡΙΣΤΙΚΕΣ ΡΙΖΕΣ ΚΑΙ ΧΑΡΑΚΤΗΡΙΣΤΙΚΑ ΔΙΑΝΥΣΜΑΤΑ ΠΙΝΑΚΑ

Αν A είναι ένας $n\times n$ τετραγωνικός πίνακας, λ ένας πραγματικός αριθμός και x ένα $n\times 1$ μη μηδενικό διάνυσμα και ισχύει η σχέση:

$$Ax=\lambda x \qquad (A.5.1)$$

τότε ο πραγματικός αριθμός λ καλείται μία "χαρακτηριστική ρίζα" και το διάνυσμα x το αντίστοιχο "χαρακτηριστικό διάνυσμα" του πίνακα A. Προφανώς τα χαρακτηριστικά διανύσματα δεν ορίζονται μονοσήμαντα διότι, όπως προκύπτει από την (A.5.1), αν x είναι ένα χαρακτηριστικό διάνυσμα του A τότε και το kx, k σταθερή, είναι ένα χαρακτηριστικό διάνυσμα του A.

Η σχέση (A.5.1) γράφεται

$$(A-\lambda I)x=0 \qquad (A.5.2)$$

και ορίζει ένα ομογενές γραμμικό σύστημα. Η αναγκαία και ικανή συνθήκη για να έχει η (A.5.2) μη μηδενική λύση είναι η

$$|A-\lambda I|=0 \qquad (A.5.3)$$

η οποία ορίζει μία εξίσωση n βαθμού ως προς λ:

$$\lambda^n+\alpha_1\lambda^{n-1}+\alpha_2\lambda^{n-2}+\dots+\alpha_{n-1}\lambda+\alpha_n=0. \qquad (A.5.4)$$

Η εξίσωση (A.5.4) καλείται "χαρακτηριστική εξίσωση" του πίνακα A και, σύμφωνα με το θεμελιώδες θεώρημα της άλγεβρας, έχει n λύσεις $\lambda_1,\lambda_2,\dots,\lambda_n$ οι οποίες δεν είναι όλες πραγματικές ή διάφορες μεταξύ τους. Σε κάθε μία από τις χαρακτηριστικές ρίζες λ_i, $i=1,2,\dots,n$ αντιστοιχεί ένα χαρακτηριστικό διάνυσμα x_i, $i=1,2,\dots,n$ και, όπως επισημάνθηκε και πιο πάνω, και όλα τα διανύσματα kx_i είναι χαρακτηριστικά διανύσματα που αντιστοιχούν στη χαρακτηριστική ρίζα λ_i.

Π.χ. αν

$$A=\begin{bmatrix}6 & 3\\ -4 & -1\end{bmatrix}$$

η χαρακτηριστική εξίσωση είναι

$$\lambda^2-5\lambda+6=0$$

και οι ρίζες της είναι οι $\lambda_1=2$ και $\lambda_2=3$. Τα χαρακτηριστικά διανύσματα που αντιστοιχούν στις χαρακτηριστικές ρίζες λ_1 και λ_2 είναι

$$x_1=\begin{bmatrix}k\\ -\frac{3}{4}k\end{bmatrix}\ \text{και}\ \ x_2=\begin{bmatrix}\zeta\\ -\zeta\end{bmatrix},\quad k,\zeta=\text{σταθερές}. \qquad (A.5.5)$$

Μπορούμε να απαλείψουμε τις σταθερές k, ζ αν "τυποποιήσουμε" τα διανύσματα x_1 και x_2 δηλαδή αν αναζητήσουμε μέσα στις κλάσεις των διανυσμάτων που ορίζει η (Α.5.5) εκείνα που έχουν μέτρο τη μονάδα:

$$\|x_1\| = \sqrt{k^2+\frac{9}{16}k^2}=\frac{5k}{4}=1$$

$$\|x_2\| = \sqrt{\zeta^2+\zeta^2}=\zeta\sqrt{2}=1$$

από τις οποίες προκύπτει ότι $k=\frac{4}{5}$ και $\zeta=\frac{\sqrt{2}}{2}$. Άρα τα τυποποιημένα χαρακτηριστικά διανύσματα του A ορίζονται μονοσήμαντα και είναι τα

$$x_1=\begin{bmatrix}\frac{4}{5}\\ -\frac{3}{5}\end{bmatrix} \quad \text{και} \quad x_2=\begin{bmatrix}\frac{\sqrt{2}}{2}\\ -\frac{\sqrt{2}}{2}\end{bmatrix}$$

Μερικές από τις ιδιότητες των χαρακτηριστικών ριζών ενός πίνακα A είναι οι εξής:

- i) Το άθροισμα των χαρακτηριστικών ριζών του A ισούται με $tr(A)$: $\sum_{i=1}^{n}\lambda_i=tr(A)$.
- ii) Το γινόμενο των χαρακτηριστικών ριζών του A ισούται με την ορίζουσα του A: $\prod_{i=1}^{n}\lambda_i=|A|$.
- iii) Αν ο πίνακας A είναι "διαγώνιος" τότε οι χαρακτηριστικές ρίζες του είναι τα στοιχεία της διαγωνίου του.
- iv) Αν ο πίνακας A είναι "ορθογώνιος" τότε οι χαρακτηριστικές ρίζες του παίρνουν μόνο τις τιμές 0 ή 1 και ο βαθμός του A ισούται με το άθροισμα των χαρακτηριστικών ριζών του: $\lambda_i=0$ ή 1 και $r(A)=\sum_{i=1}^{n}\lambda_i[=tr(A)]$.
- v) Αν οι πίνακες A και B είναι "όμοιοι" -δηλαδή αν $B=C^{-1}AC$, όπου C είναι ένας μη ιδιάζων τετραγωνικός πίνακας, τότε έχουν τις ίδιες χαρακτηριστικές ρίζες και τα ίδια χαρακτηριστικά διανύσματα.

(Α.5.6)

- vi) Αν λ_i είναι μία χαρακτηριστική ρίζα του A και k είναι ένας θετικός ακέραιος, τότε ο αριθμός λ_i^k είναι χαρακτηριστική ρίζα του πίνακα A^k. Αν ο A είναι μη ιδιάζων τότε ο αριθμός λ_i^{-k} είναι χαρακτηριστική ρίζα του A^{-k}.
- vii) Αν ο πίνακας A είναι "συμμετρικός" τότε οι χαρακτηριστικές του ρίζες είναι όλες πραγματικοί αριθμοί και ο βαθμός του ισούται με τον αριθμό των μη μηδενικών χαρακτηριστικών ριζών του. Ακόμη για τις χαρακτηριστικές ρίζες των συμμετρικών πινάκων ισχύουν οι εξής ιδιότητες:

  Αν x_i και x_j είναι το χαρακτηριστικό διανύσματα που αντιστοιχούν στις μη μηδενικές χαρακτηριστικές ρίζες λ_i και λ_j $(\lambda_i\neq\lambda_j)$ τότε $x_i'x_j=0$. δηλαδή τα χαρακτηριστικά διανύσματα x_i και x_j είναι ορθογώνια.

  Αν x_i είναι χαρακτηριστικό διάνυσμα που αντιστοιχεί στη χαρακτηριστική ρίζα λ_i τότε $x_i'Ax_i=\lambda_i$.

  Ο πίνακας A είναι όμοιος με το διαγώνιο πίνακα V του οποίου τα διαγώνια στοιχεία είναι οι χαρακτηριστικές ρίζες του A.

  Αν M είναι ο πίνακας των "τυποποιημένων" χαρακτηριστικών διανυσμάτων του A τότε ο M είναι "ορθογώνιος" και ισχύουν τα εξής:

(Α.5.6)

$$M'=M^{-1} \quad \text{και} \quad M'AM=M^{-1}AM=V.$$

Α.6. ΓΡΑΜΜΙΚΟΙ ΜΕΤΑΣΧΗΜΑΤΙΣΜΟΙ, ΓΡΑΜΜΙΚΕΣ ΚΑΙ ΤΕΤΡΑΓΩΝΙΚΕΣ ΜΟΡΦΕΣ ΠΙΝΑΚΑ

Κάθε m×n πίνακας A ορίζει ένα μετασχηματισμό από τον ευκλείδιο χώρο

$$R^n=\{x/x'=[x_1,x_2,\dots,x_n]\}$$

στον ευκλείδιο χώρο

$$R^m=\{y/y'=[y_1,y_2,\dots,y_m]\}$$

διότι κάθε n×1 διάνυσμα $x\in R^n$ πολλαπλασιαζόμενο από αριστερά επί τον πίνακα $A_{m\times n}$ δίνει ως γινόμενο ένα m×1 διάνυσμα $y\in R^m$

$$A_{m\times n}x_{n\times 1}=y_{m\times 1}. \qquad (A.6.1)$$

Ο μετασχηματισμός (A.6.1) είναι γραμμικός διότι ισχύουν οι ιδιότητες

$$A(x_1+x_2)=Ax_1+Ax_2$$
$$A(\lambda x_1)=\lambda(Ax_1), \qquad (A.6.2)$$

όπου x_1 και x_2 είναι διανύσματα του χώρου R^n και λ μία σταθερή. Ο γραμμικός μετασχηματισμός (A.6.1) απεικονίζει την αρχή των αξόνων του χώρου R^n στην αρχή των αξόνων του χώρου R^m:

$$A_{m\times n}O_{n\times 1}=O_{m\times 1}.$$

Κάθε διάνυσμα y που ορίζεται από την (A.6.1) ονομάζεται μία "γραμμική μορφή" ως προς τα στοιχεία του x και κάθε στοιχείο του y είναι ένας γραμμικός συνδιασμός των στοιχείων του x με σταθμιστές τα στοιχεία του πίνακα A:

$$y_i=\sum_{j=1}^{n}\alpha_{ij}x_j, \quad i=1,2,\dots,m. \qquad (A.6.3)$$

Αν ο πίνακας A είναι ένας ορθογώνιος πίνακας τότε η γραμμική μορφή $y=Ax$ ονομάζεται "ορθογώνια".

Αν A είναι ένας συμμετρικός (άρα και τετραγωνικός) πίνακας τότε ο πραγματικός αριθμός

$$Q=x'Ax \qquad (A.6.4)$$

που ορίζεται για κάθε n×1 διάνυσμα x, ονομάζεται μία "τετραγωνική μορφή" ως προς τα στοιχεία του x και είναι μία ομογενής συνάρτηση δευτέρου βαθμού ως προς τα στοιχεία του x:

$$Q=x'Ax=[x_1,x_2,\dots,x_n]\begin{bmatrix}\alpha_{11} & \alpha_{12} & \dots & \alpha_{1n}\\ \alpha_{21} & \alpha_{22} & \dots & \alpha_{2n}\\ \vdots & \vdots & & \vdots\\ \alpha_{n1} & \alpha_{n2} & \dots & \alpha_{nn}\end{bmatrix}\begin{bmatrix}x_1\\ x_2\\ \vdots\\ x_n\end{bmatrix} \qquad (A.6.5)$$

$$=\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_{ij}x_ix_j$$

$$=\sum_{i=1}^{n}\alpha_{ii}x_i^2+\sum_{\substack{i=1\\(i\neq j)}}^{n}\sum_{j=1}^{n}\alpha_{ij}x_ix_j$$

$$=\alpha_{11}x_1^2+\dots+\alpha_{nn}x_n^2+2\alpha_{12}x_1x_2+2\alpha_{13}x_1x_3+\dots+2\alpha_{n-1,n}x_{n-1}x_n.$$

Ο πίνακας A ονομάζεται:

- i) "θετικός πεπερασμένος" τότε και μόνο τότε αν $x'Ax>0$ για κάθε $x\neq 0$.
- ii) "αρνητικός πεπερασμένος" τότε και μόνο τότε αν $x'Ax<0$ για κάθε $x\neq 0$.
- iii) "μη αρνητικός πεπερασμένος" τότε και μόνο τότε αν $x'Ax\geqslant 0$ για κάθε $x\neq 0$.

Μερικές από τις ιδιότητες των θετικών (αρνητικών) πεπερασμένων πινάκων είναι οι εξής:

- i) Αν ο A είναι θετικός (αρνητικός) πεπερασμένος τότε ο $-A$ είναι αρνητικός (θετικός) πεπερασμένος.
- ii) Ο A είναι θετικός (αρνητικός) πεπερασμένος τότε και μόνο τότε αν όλες οι χαρακτηριστικές του ρίζες είναι θετικές (αρνητικές).
- iii) Αν ο A είναι θετικός πεπερασμένος τότε όλες οι αρχικές ορίζουσές του είναι θετικές (άρα και $|A|>0$), ο A είναι μη ιδιάζων και $r(A)=n$. Αν ο A είναι αρνητικός πεπερασμένος τότε οι αρχικές ορίζουσές του έχουν εναλλασσόμενα πρόσημα (αρχίζοντας από αρνητικό πρόσημο) ο A είναι μη ιδιάζων και $r(A)=n$.
- iv) Αν ο A είναι θετικός (αρνητικός) πεπερασμένος και P είναι ένας μη ιδιάζων πίνακας τό-

(A.6.6)

τε ο πίνακας $P'AP$ είναι θετικός πεπερασμένος.

v) Αν ο A είναι θετικός (αρνητικός) πεπερασμένος τότε ο A^{-1} είναι θετικός (αρνητικός) πεπερασμένος.

vi) Αν ο A είναι ένας m×n πίνακας με $r(A)=n<m$ τότε ο $A'A$ είναι θετικός πεπερασμένος και μη ιδιάζων ενώ ο AA' είναι μη αρνητικός πεπερασμένος.

vii) Αν ο A είναι θετικός πεπερασμένος τότε υπάρχει ένας μη ιδιάζων πίνακας K τέτοιος ώστε $K'K=A$ και $KA^{-1}K'=I$.

viii) Αν M είναι ο πίνακας των τυποποιημένων χαρακτηριστικών διανυσμάτων του A και $x=My$ τότε

$$x'Ax=y'M'AMy=y'Vy=\sum_{j=1}^{n}\lambda_j y_j^2$$

όπου λ_j, i=1,2,...,n, είναι οι χαρακτηριστικές ρίζες του A. Έτσι, μια τετραγωνική μορφή $x'Ax$ μπορεί πάντα να γραφτεί ως σταθμικό άθροισμα τετραγώνων με σταθμιστές τις χαρακτηριστικές ρίζες του A.

ix) Αν $A=B+C$, όπου ο B είναι θετικός πεπερασμένος και ο C μη αρνητικός πεπερασμένος, τότε, (α) ο A είναι θετικός πεπερασμένος, (β) $|B|\leqslant|A|$ και (γ) ο πίνακας $B^{-1}-A^{-1}$ είναι μη αρνητικός πεπερασμένος.

(A.6.6)

Α.7. ΔΙΑΦΟΡΙΚΟΣ ΛΟΓΙΣΜΟΣ ΤΩΝ ΔΙΑΝΥΣΜΑΤΩΝ ΚΑΙ ΤΩΝ ΠΙΝΑΚΩΝ

Αν τα στοιχεία του n×1 διανύσματος x είναι συναρτήσεις της μεταβλητής t τότε η "παράγωγος" του διανύσματος x ως προς t είναι το διάνυσμα:

$$\frac{dx}{dt}=\begin{bmatrix}\frac{dx_1}{dt}\\ \frac{dx_2}{dt}\\ \vdots\\ \frac{dx_n}{dt}\end{bmatrix} \qquad (A.7.1)$$

Αν τα στοιχεία του m×n πίνακα $A=[\alpha_{ij}]$ είναι συναρτήσεις της μεταβλητής t τότε η παράγωγος του A ως προς t είναι ο πίνακας

$$\frac{dA}{dt}=\left[\frac{d\alpha_{ij}}{dt}\right] \qquad (A.7.2)$$

και ισχύουν οι εξής ιδιότητες:

$$\frac{dAB}{dt}=A\frac{dB}{dt}+\frac{dA}{dt}B \qquad (A.7.3)$$

$$\frac{dA^{-1}}{dt}=-A^{-1}\frac{dA}{dt}A^{-1}. \qquad (A.7.4)$$

Αν $y=f(x)$, όπου $x=[x_1,x_2,\ldots,x_n]'$, είναι μία πραγματική συνάρτηση των στοιχείων $x_1,x_2,\ldots,x_n$ του διανύσματος x, τότε το διάνυσμα των παραγώγων πρώτης τάξης:

$$\frac{\partial y}{\partial x}=\begin{bmatrix}\frac{\partial y}{\partial x_1}\\ \frac{\partial y}{\partial x_2}\\ \vdots\\ \frac{\partial y}{\partial x_n}\end{bmatrix}=\nabla y \qquad (A.7.5)$$

ονομάζεται "ανάδελτα" (gradient vector) της συνάρτησης y, ενώ ο πίνακας των παραγώγων δεύτερης τάξης:

$$\frac{\partial^2 y}{\partial x^2}=\frac{\partial(\partial y/\partial x)}{\partial x}=\begin{bmatrix}\frac{\partial^2 y}{\partial x_1^2}, & \frac{\partial^2 y}{\partial x_1\partial x_2}, & \ldots, & \frac{\partial^2 y}{\partial x_1\partial x_n}\\ \frac{\partial^2 y}{\partial x_2\partial x_1}, & \frac{\partial^2 y}{\partial x_2^2}, & \ldots, & \frac{\partial^2 y}{\partial x_2\partial x_n}\\ \vdots & \vdots & & \vdots\\ \frac{\partial^2 y}{\partial x_n\partial x_1} & \frac{\partial^2 y}{\partial x_n\partial x_2}, & \ldots, & \frac{\partial^2 y}{\partial x_n^2}\end{bmatrix} \qquad (A.7.6)$$

ονομάζεται "πίνακας του Hesse" (Hessian matrix) της συνάρτησης y.

Αν $y=f(x)=[f_1(x), f_2(x), \ldots, f_m(x)]$ είναι μία "διανυσματική συνάρτηση" του διανύσματος $x'=[x_1, x_2, \ldots, x_n]$, τότε ο πίνακας

$$\frac{\partial y}{\partial x}=\left[\frac{\partial f_1(x)}{\partial x}, \frac{\partial f_2(x)}{\partial x}, \ldots, \frac{\partial f_m(x)}{\partial x}\right]$$

$$=\begin{bmatrix} \frac{\partial f_1(x)}{\partial x_1}, & \frac{\partial f_2(x)}{\partial x_1}, & \ldots, & \frac{\partial f_m(x)}{\partial x_1} \\ \frac{\partial f_1(x)}{\partial x_2}, & \frac{\partial f_2(x)}{\partial x_2}, & \ldots, & \frac{\partial f_m(x)}{\partial x_2} \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial f_1(x)}{\partial x_n}, & \frac{\partial f_2(x)}{\partial x_n}, & \ldots, & \frac{\partial f_m(x)}{\partial x_n} \end{bmatrix} \qquad (A.7.7)$$

ονομάζεται "Ιακωβιανός πίνακας" (Jacobian matrix) της συνάρτησης y, ενώ η ορίζουσα του Ιακωβιανού πίνακα ονομάζεται "Ιακωβιανή".

Σχετικά με την παραγώγιση των γραμμικών και "τετραγωνικών" μορφών ισχύουν τα εξής:

$$\left.\begin{array}{lll} \text{i)} & \frac{\partial a'x}{\partial x}=a' & \\ \text{ii)} & \frac{\partial Ax}{\partial x}=A' & \\ \text{iii)} & \frac{\partial x'Ax}{\partial x}=2x'A & \\ \text{iv)} & \frac{\partial y'Ax}{\partial x}=y'A, & A \text{ συμμετρικός} \\ \text{v)} & \frac{\partial x'Ax}{\partial A}=xx' & \\ \text{vi)} & \frac{\partial y'Ax}{\partial A}=xy', & A \text{ συμμετρικός.} \end{array}\right\} \qquad (A.7.8)$$

Ακόμα αποδεικνύεται ότι:

$$\begin{array}{ll} \text{vii)} & \frac{\partial tr(A)}{\partial A}=I \\ \text{viii)} & \frac{\partial |A|}{\partial A}=|A|A^{-1} \\ \text{ix)} & \frac{\partial \ln|A|}{\partial A}=A^{-1}, \quad |A|>0. \end{array} \qquad (A.7.9)$$

A.8. ΜΕΓΙΣΤΑ ΚΑΙ ΕΛΑΧΙΣΤΑ ΣΥΝΑΡΤΗΣΕΩΝ

Ας θεωρήσουμε το πρόβλημα:

$$\min f(x), \quad \text{ως προς } x \in X \subseteq R^n.$$

Το σημείο $x^* \in X$ ορίζει ένα "καθολικό ελάχιστο" για την $f(x)$, αν

$$f(x^*) \leq f(x), \quad \text{για κάθε } x \in X.$$

Αν η συνάρτηση $f(x)$ είναι συνεχής στο σύνολο X και το X είναι "κλειστό" και "φραγμένο" τότε το καθολικό ελάχιστο υπάρχει.

Το σημείο $x^0 \in X$ ορίζει ένα "τοπικό ελάχιστο" για την $f(x)$, αν

$$f(x^0) \leq f(x), \quad \text{για κάθε } x \in N_\alpha(x^0)$$

όπου $N_\alpha(x^0)$ είναι μία περιοχή του x^0 με ακτίνα $\alpha>0$. Το καθολικό ελάχιστο είναι και ένα τοπικό ελάχιστο αλλά όχι και αντίστροφα.

Αν το "όρισμα" x της συνάρτησης $f(x)$ δεν υπόκειται σε περιορισμούς τότε $X=R^n$ και οι αναγκαίες και ικανές συνθήκες για να είναι το σημείο x^0 ένα τοπικό ελάχιστο είναι οι εξής:

i) $\frac{\partial f}{\partial x}(x^0)=0$,

ii) οι μερικές παράγωγοι πρώτης και δεύτερης τάξης της $f(x)$ είναι συνεχείς συναρτήσεις σε μια περιοχή $N_\alpha(x^0)$ και (A.8.1)

iii) ο πίνακας του Hesse των μερικών παραγώγων δεύτερης τάξης, υπολογιζόμενος στο σημείο x^0:

$$\frac{\partial^2 f}{\partial x^2}(x^0)$$

είναι θετικός πεπερασμένος.

Το πρόβλημα της αναζήτησης του μέγιστου της συνάρτησης $f(x)$ ανάγεται στην αναζήτηση του ελάχιστου της συνάρτησης $-f(x)$: $\max f(x)=\min[-f(x)]$.

Αν το όρισμα x της συνάρτησης $f(x)$ υπόκειται σε "m ανεξάρτητους γραμμικούς περιορισμούς":

$$g(x)=b \qquad (A.8.2)$$

τότε το σύνολο X είναι το

$$X=\{x/g(x)=b\}\subseteq R^n$$

και στην περίπτωση αυτή το πρόβλημα τίθεται ως εξής:

$$\min f(x), \quad \text{κάτω από τους περιορισμούς } g(x)=b.$$

Η κλασική μέθοδος για την επίλυση του προβλήματος είναι η μέθοδος των πολλαπλασιαστών του Lagrange. Η συνάρτηση του Lagrange για τη συνάρτηση $f(x)$ είναι η

$$L(x,\lambda)=f(x)+\lambda[b-g(x)] \qquad (A.8.3)$$

όπου λ είναι το διάνυσμα των πολλαπλασιαστών του Lagrange. Οι αναγκαίες και ικανές συνθήκες για να είναι το σημείο x ένα τοπικό ελάχιστο της $f(x)$ είναι οι εξής:

i) $\frac{\partial L}{\partial x}(x^0)=\frac{\partial f}{\partial x}(x^0)-\lambda\frac{\partial g}{\partial x}(x^0)=0$, (n εξισώσεις)

ii) $\frac{\partial L}{\partial \lambda}(x^0)=b-g(x^0)=0$, (m εξισώσεις) (A.8.4)

iii) οι μερικές παράγωγοι πρώτης και δεύτερης τάξης της $f(x)$ και οι μερικές παράγωγοι πρώτης τάξης της $g(x)$ να είναι συνεχείς συναρτήσεις σε μια περιοχή $N_\alpha(x^0)$ του x^0.

iv) ο πίνακας των μερικών παραγώγων της $g(x)$, υπολογιζόμενος στο x^0:

$$\frac{\partial g}{\partial x'}(x^0)=\begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \frac{\partial g_1}{\partial x_2} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \frac{\partial g_2}{\partial x_1} & \frac{\partial g_2}{\partial x_2} & \cdots & \frac{\partial g_2}{\partial x_n} \\ \vdots & \vdots & & \vdots \\ \frac{\partial g_m}{\partial x_1} & \frac{\partial g_m}{\partial x_2} & \cdots & \frac{\partial g_m}{\partial x_n} \end{bmatrix}_{x_0} \qquad (A.8.4)$$

είναι πλήρους βαθμού: $r\left[\frac{\partial g}{\partial x'}(x^0)\right]=m$, και

v) ο πίνακας των παραγώγων δεύτερης τάξης της $f(x)$, υπολογιζόμενος στο x^0:

$$\frac{\partial^2 f}{\partial x^2}(x^0)$$

να είναι θετικός πεπερασμένος.

A.9. ΑΡΙΘΜΗΤΙΚΕΣ ΜΕΘΟΔΟΙ ΥΠΟΛΟΓΙΣΜΟΥ ΤΩΝ ΑΚΡΟΤΑΤΩΝ ΜΙΑΣ ΣΥΝΑΡΤΗΣΗΣ

Τα ακρότατα μιας συνάρτησης προκύπτουν από την επίλυση του συστήματος των εξισώσεων (A.8.1.i) ή (A.8.4.i). Αν οι εξισώσεις αυτές είναι μη γραμμικές και μάλιστα υψηλού βαθμού, τότε η αναλυτική επίλυσή τους είναι δύσκολη ή αδύνατη. Σε τέτοιες περιπτώσεις τα ακρότατα υπολογίζονται με "επαναληπτικές προσεγγιστικές διαδικασίες" στον ηλεκτρονικό υπολογιστή.

Περιγράφουμε παρακάτω μερικές βασικές μεθόδους για τον "αριθμητικό" υπολογισμό των ακροτάτων μιας συνάρτησης.

Στη γενική περίπτωση το πρόβλημα τίθεται ως εξής:

$$\min_x f(x)=\max_x[-f(x)] \qquad (A.9.1)$$

με ή χωρίς περιορισμούς για τα στοιχεία του διανύσματος x.

Στην Οικονομετρία η συνάρτηση $f(x)$ είναι συνήθως το άθροισμα $S(b)$ των τετραγώνων των σφαλμάτων όταν χρησιμοποιούμε τη μη γραμμική μέθοδο ελαχίστων τετραγώνων ή η συνάρτηση πιθανοφάνειας $L(b)$ όταν χρησιμοποιούμε τη μέθοδο της μέγιστης πιθανοφάνειας, ενώ b είναι το διάνυσμα των προς εκτίμηση παραμέτρων. (σελ. 130)

Η σημαντικότερη κλάση "αριθμητικών μεθόδων" για την επίλυση του προβλήματος (4.9.1) είναι εκείνη που βασίζεται στις μεθόδους διαφορικού λογισμού και ακολουθεί την εξής διαδικασία: ξεκινά από ένα αρχικό διάνυσμα εκτιμήσεων x_0 και προσεγγίζει το μέγιστο (ή ελάχιστο) με επαναληπτικές προσεγγιστικές διαδικασίες σύμφωνα με το βασικό σχήμα:

$$x_{p+1}=x_p+h_p d_p \qquad (A.9.2)$$

όπου x_p είναι η προσέγγιση του μεγίστου κατά την p επαναληπτική διαδικασία, d_p είναι το διάνυσμα που ορίζει την "κατεύθυνση" προς την οποία θα αναζητήσουμε το μέγιστο κατά την (p+1) επαναληπτική διαδικασία και h_p είναι ένας θετικός αριθμός που ορίζει το "μήκος του βήματος" προς την κατεύθυνση d_p. Οι επαναληπτικές διαδικασίες συνεχίζονται μέχρι να ικανοποιηθεί κάποιο κριτήριο σύγκλισης.

Οι επιμέρους μέθοδοι της κλάσης αυτής διαφέρουν ως προς την επιλογή των h_p, d_p και του κριτηρίου σύγκλισης των επαναληπτικών διαδικασιών.

Αν η συνάρτηση $f(x)$ είναι συνεχής και έχει συνεχείς παραγώγους πρώτης και δεύτερης τάξης τότε οι αναγκαίες συνθήκες για την ύπαρξη μεγίστου είναι οι

$$\frac{\partial f}{\partial x}=g(x)=0 \qquad (A.9.3)$$

και ο πίνακας του Hesse:

$$\frac{\partial^2 f}{\partial x \partial x'}=G \quad \text{να είναι αρνητικός πεπερασμένος.} \qquad (A.9.4)$$

Ας καλέσουμε g_p την τιμή του διανύσματος (A.9.3) και

G_p την τιμή του πίνακα (A.9.4) στο σημείο $x'_p=[x_1^p, x_2^p, \dots, x_k^p]$. Ας υποθέσουμε ακόμα ότι μας δίνεται μια αρχική εκτίμηση $x'_0=[x_1, x_2, \dots, x_k]$ του διανύσματος x και ότι αναζητούμε το μέγιστο της $f(x)$ σύμφωνα με τις επαναληπτικές διαδικασίες που ορίζει το σχήμα (A.9.2).

Στην κλάση των αριθμητικών μεθόδων που βασίζονται στις μεθόδους του διαφορικού λογισμού το διάνυσμα d_p που ορίζει τη νέα κατεύθυνση προς την οποία θα αναζητήσουμε το μέγιστο εκφράζεται από τη σχέση:

$$d_p=B_p g_p \qquad (A.9.5)$$

όπου B_p είναι ένας θετικός πεπερασμένος πίνακας σταθμιστών.

i) Η μέθοδος της ταχύτερης ανόδου στο μέγιστο (the method of steepest-ascent).

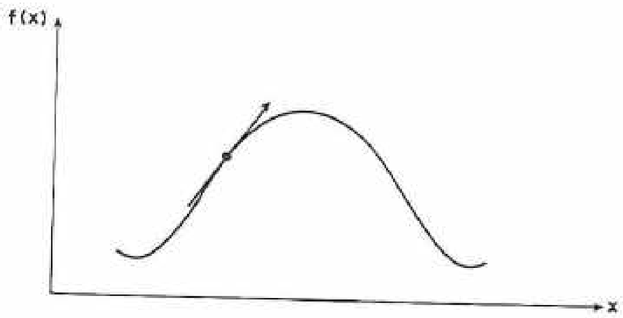
Στη μέθοδο αυτή λαμβάνεται

$$B=I \qquad (A.9.6)$$

ή,

$$d_p=g_p \qquad (A.9.7)$$

και η αιτιολογία για την επιλογή αυτή είναι ότι τα σημεία του διαφορικού (gradient points) αυξάνουν την τιμή της καλύτερης γραμμικής προσέγγισης της $f(x)$ προς την κατεύθυνση του μεγίστου (Βλ. Σχήμα A.9.1).



Σχήμα (A.9.1)

Ως h_p μπορούμε να επιλέξουμε κάποια σταθερή, αλλά η αυθαίρετη επιλογή της σταθερής είναι δυνατό να μην αποδόσει καλά αποτελέσματα: αν η τιμή της h_p είναι μικρή η σύγκλιση μπορεί να είναι αργή και δαπανηρή ενώ αν η τιμή της h_p είναι μεγάλη τότε το σχήμα (Α.9.2) μπορεί να μη συγκλίνει καθόλου. Η κατάλληλη τιμή της σταθερής h_p μπορεί να προσδιοριστεί αν ζητήσουμε να είναι τέτοια ώστε να μεγιστοποιεί τη βελτίωση της συνάρτησης f σε κάθε επαναληπτική διαδικασία: υποθέτοντας ότι η $f(x)$ επιδέχεται ανάπτυξη δευτέρου βαθμού κατά Taylor γύρω από το σημείο x_p έχουμε:

$$f(x_{p+1})=f(x_p)+(x_{p+1}-x_p)'g_p+\frac{1}{2}(x_{p+1}-x_p)'G_p(x_{p+1}-x_p)+R \quad (A.9.8)$$

ή, λαμβάνοντας υπόψη τις (Α.9.2) και (Α.9.7):

$$f(x_{p+1})-f(x)=h_p g_p' g_p+\frac{1}{2}h_p^2 g_p' G_p g_p+R. \quad (A.9.9)$$

Για να μεγιστοποιήσουμε το αριστερό μέλος της (Α.9.9) ως προς h_p, θέτουμε την πρώτη παράγωγο του δευτέρου μέλους ως προς h_p ίση με μηδέν, αγνοώντας το υπόλοιπο R, και λαμβάνουμε:

$$h_p=-(g_p' G_p g_p)^{-1} g_p' g_p, \quad (A.9.10)$$

οπότε το σχήμα των επαναληπτικών διαδικασιών (Α.9.2) παίρνει τη μορφή:

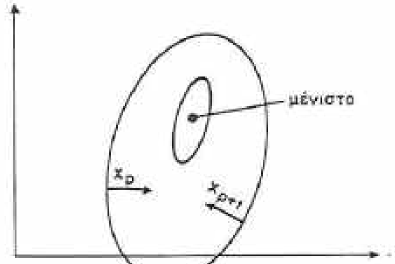
$$x_{p+1}=x_p-(g_p' G_p g_p)^{-1} g_p' g_p g_p. \quad (A.9.11)$$

Η σταθερή (Α.9.10) θα μεγιστοποιεί το πρώτο μέλος της (Α.9.9) αν η δεύτερη παράγωγος ως προς h_p είναι αρνητική, δηλαδή αν

$$g_p' G_p g_p < 0, \quad (A.9.12)$$

η οποία βέβαια σημαίνει ότι ο πίνακας G_p του Hesse, σε κάθε επαναληπτική διαδικασία, πρέπει να είναι "αρνητικός πεπερασμένος". Αλλά αναγκαία και ικανή συνθήκη για να είναι ο πίνακας G_p αρνητικός πεπερασμένος είναι η συνάρτηση $f(x)$ να

είναι "κυρτή" σε κάποια περιοχή του x_p. Άρα η μέθοδος που περιγράψαμε δε θα είναι αποτελεσματική αν το σημείο x_p δεν είναι αρκετά κοντά στο μέγιστο έτσι ώστε να εξασφαλιστεί η κυρτότητα της $f(x)$. Μια πρόσθετη δυσκολία είναι η πιθανότητα που υπάρχει οι επαναληπτικές διαδικασίες να οδηγήσουν σε ένα "σημείο καμπής" και όχι στο μέγιστο. Το σημαντικότερο όμως πρόβλημα της μεθόδου είναι ο αναλυτικός υπολογισμός του πίνακα G_p του Hesse σε κάθε επαναληπτική διαδικασία και αυτό είναι γενικά πολύ δύσκολο. Τέλος, ένα πρακτικό μειονέκτημα της μεθόδου αυτής είναι το ότι μπορεί να παλινδρομεί κατά μήκος του μεγίστου -αν το μέγιστο βρίσκεται σε μια πολύ "στενή" περιοχή- κάνοντας έτσι τη σύγκλιση πολύ αργή και δαπανηρή (Βλ. Σχ. Α.9.2).



Η κατεύθυνση αναζήτησης του μεγίστου είναι σχεδόν κάθετη προς την επιθυμητή κατεύνθυνση.

Σχήμα (Α.9.2)

ii) Η μέθοδος Newton-Raphson

Το βασικό χαρακτηριστικό της μεθόδου των Newton-Raphson είναι ότι μεγιστοποιεί την (Α.9.8) ως προς το άγνωστο διάνυσμα x_{p+1}. Παραγωγίζοντας ως προς x_{p+1}, εύκολα προκύπτουν οι συνθήκες πρώτης τάξης:

---

1. Η συνάρτηση $f(x)$ ονομάζεται "κυρτή" στο $S \in R^n$ αν για κάθε x_1 και $x_2 \in S$ ισχύει:

$$f[\lambda x_1+(1-\lambda)x_2] \leqslant \lambda f(x_1)+(1-\lambda)f(x_2), \quad 0 \leqslant \lambda \leqslant 1.$$

και "καθαρά κυρτή" αν ισχύει μόνο η ανισότητα.

$$g_p+G_p(x_{p+1}-x_p)=0 \quad (A.9.13)$$

ή

$$x_{p+1}=x_p-G_p^{-1}g_p\,. \quad (A.9.14)$$

Αν η συνάρτηση $f(x)$ είναι δευτέρου βαθμού, τότε η προσέγγιση (A.9.8) είναι ακριβής και η (A.9.14) προσδιορίζει το μέγιστο σε ένα βήμα. Αν η $f(x)$ είναι ανωτέρου βαθμού αλλά το x_p βρίσκεται κοντά στο μέγιστο, τότε η προσέγγιση (A.9.14) θα είναι πολύ καλή και η σύγκλιση προς το μέγιστο αναμένεται να είναι γρήγορη. Αντίθετα, αν το x_p είναι αρκετά μακρυά από το μέγιστο έτσι ώστε η $f(x)$ να μην είναι κυρτή σε μια περιοχή του σημείου αυτού, τότε η μέθοδος N-R δε θα συγκλίνει και θα εξακολουθήσει να κινείται προς λανθασμένη κατεύθυνση. Ο λόγος για τον οποίο δε θα συγκλίνει η μέθοδος N-R στην περίπτωση αυτή είναι ο ίδιος με εκείνον που αναφέραμε και στην προηγούμενη μέθοδο: η συνθήκη δεύτερης τάξης για την ύπαρξη του μεγίστου απαιτεί ο πίνακας G_p του Hesse να είναι αρνητικός πεπερασμένος.

Οι πιο σύγχρονες τεχνικές αρχίζουν την αναζήτηση του μεγίστου με άλλες μεθόδους και όταν το πλησιάσουν αρκετά τότε χρησιμοποιούν τη μέθοδο N-R για να επιτύχουν την ταχύτερη σύγκλιση.

iii) Η μέθοδος της τετραγωνικής αναρρύχισης και ορισμένες παραλλαγές της

Σύμφωνα με τη μέθοδο αυτή (που είναι παραλλαγή της μεθόδου N-R) επιχειρούμε, σε κάθε επαναληπτική διαδικασία, εκείνο το βήμα που μεγιστοποιεί την προσέγγιση δευτέρου βαθμού της $f(x)$ πάνω σε μια σφαίρα κατάλληλης ακτίνας διότι σε μια τέτοια περίπτωση εξασφαλίζεται η κυρτότητα της $f(x)$ και συνεπώς το αρνητικό πεπερασμένο του πίνακα G_p του Hesse. Αυτό μπορεί να γίνει ως εξής: ας υποθέσουμε ότι ο πίνακας G, σε κάποια από τις επαναληπτικές διαδικασίες, δεν είναι αρνητικός πεπερασμένος. Αν εκφράσουμε τον πίνακα G ως

$$G=MAM'\,. \quad (A.9.15)$$

όπου M είναι ένας ορθογώνιος πίνακας και A είναι ο διαγώνιος πίνακας του οποίου τα στοιχεία είναι οι χαρακτηριστικές τιμές του πίνακα G, τότε μερικές χαρακτηριστικές τιμές του G θα είναι θετικές. Αν τώρα αντί του πίνακα G χρησιμοποιήσουμε τον πίνακα

$$G^*=MA^*M' \quad (A.9.16)$$

όπου ο A^* προκύπτει από τον A αν αλλάξουμε τα πρόσημα των θετικών διαγωνίων στοιχείων του, δηλαδή των θετικών χαρακτηριστικών τιμών του G, αφήνοντας τις αρνητικές χαρακτηριστικές τιμές ίδιες, τότε είναι φανερό ότι ο πίνακας G^* θα είναι αρνητικός πεπερασμένος. Η μέθοδος αυτή δίνει, κατά κανόνα, πολύ καλά αποτελέσματα στις πρακτικές εφαρμογές.

Γενικά, στη μέθοδο της τετραγωνικής αναρρύχισης μπορούμε να αντικαταστήσουμε τον πίνακα G_p με τον πίνακα $G_p-\lambda_p I$, $\lambda_p \geq 0$, έτσι ώστε να έχουμε:

$$x_{p+1}=x_p-(G_p-\lambda_p I)^{-1}g_p \quad (A.9.17)$$

από την οποία εύκολα προκύπτει ότι αν $\lambda_p=0$ τότε έχουμε τη μέθοδο N-R.

Μια άλλη παραλλαγή της μεθόδου της τετραγωνικής αναρρύχισης είναι η χρησιμοποίηση του σχήματος:

$$x_{p+1}=x_p-(G_p-S_p)^{-1}g_p \quad (A.9.18)$$

όπου S_p είναι ένας θετικός πεπερασμένος πίνακας τέτοιος ώστε ο πίνακας G_p-S_p να είναι αρνητικός πεπερασμένος.

iv) Μέθοδοι των "συζυγών διευθύνσεων"

Οι μέθοδοι που ανήκουν στην ευρεία αυτή κλάση έχουν ως βασικό χαρακτηριστικό την αναζήτηση του μεγίστου κατά μήκος των λεγόμενων συζυγών διευθύνσεων. Οι συζυγείς διευθύνσεις ορίζονται ως εξής:

Δύο διευθύνσεις d_i και d_j ονομάζονται συζυγείς ως προς την τετραγωνική μορφή $x'Ax+b'x+c$ (A.9.19) αν ισχύει $d_i'Ad_j=0$.

Στην πράξη, φυσικά, η συνάρτηση που πρέπει να μεγιστοποιήσουμε δεν είναι τετραγωνική και συνεπώς οι διάφορες μέθοδοι της κατηγορίας αυτής βασίζονται σε ψευδοσυζυγείς διευθύνσεις. Μεταξύ άλλων, οι διαφορετικές προσεγγίσεις που χρησιμοποιούν συζυγείς διευθύνσεις διαφέρουν μεταξύ τους στον τρόπο επιλογής των συζυγών διευθύνσεων οι οποίες δεν ορίζονται μονοσήμαντα. Από αυτές οι συζυγείς "διαφορικές" διευθύνσεις αποδείχτηκαν αρκετά επιτυχείς στην πράξη αν και, θεωρητικά, αυτό δεν έχει αιτιολογηθεί πλήρως.

Μια από τις πιο αποτελεσματικές μεθόδους του τύπου αυτού η οποία δεν απαιτεί τον υπολογισμό των παραγώγων της συνάρτησης $f(x)$ είναι η μέθοδος των "διαφορικών" συζυγών διευθύνσεων του Powell. Όπως και άλλα μέλη της κλάσης αυτής, η μέθοδος του Powell βασίζεται στην επίλυση μιας σειράς μονοδιάστατων μεγιστοποιήσεων: σε κάθε επαναληπτική διαδικασία η συνάρτηση $f(x)$ μεγιστοποιείται διαδοχικά κατά μήκος n συζυγών διευθύνσεων. Στο τέλος κάθε τέτοιας επαναληπτικής διαδικασίας είτε διατηρούμε τις n προηγούμενες συζυγείς διευθύνσεις, είτε μια από αυτές αντικαθίσταται με μια νέα συζυγή διεύθυνση.

Τα ιδιαίτερα χαρακτηριστικά της μεθόδου του Powell μπορούμε να τα επισημάνουμε αν παρακολουθήσουμε τα βήματα μιας επαναληπτικής διαδικασίας.

Έστω $f(x)=f(x_1,x_2,\ldots,x_n)$ η συνάρτηση της οποίας αναζητούμε το μέγιστο. Ας υποθέσουμε ότι έχουμε συμπληρώσει p επαναληπτικές διαδικασίες και έστω $d_1^p, d_2^p, \ldots, d_n^p$ n γραμμικά ανεξάρτητες διευθύνσεις αναζήτησης του μεγίστου από την πιο πρόσφατη εκτίμηση του μεγίστου $x'^p=[x_1^p, x_2^p, \ldots, x_n^p]$. Τα βήματα της επόμενης επαναληπτικής διαδικασίας είναι τα εξής:

α) Για $r=1,2,\ldots,n$ υπολογίζουμε, διαδοχικά, τις σταθερές λ_r έτσι ώστε να μεγιστοποιείται η

$$f(x+\sum_{j=1}^{r}\lambda_j d_j^p). \tag{A.9.20}$$

β) Έστω

$$\bar{x}^p=\sum_{j=1}^{n}\lambda_j d_j^p+x^p. \tag{A.9.21}$$

Η μετατόπιση

$$f=\sum_{j=1}^{n}\lambda_j d_j^p \tag{A.9.22}$$

χρησιμοποιείται ως διεύθυνση αναζήτησης του μεγίστου, δηλαδή η τιμή του λ επιλέγεται ώστε να μεγιστοποιείται η

$$f(\bar{x}^p+\lambda f). \tag{A.9.23}$$

γ) Ως σημείο αφετηρίας για την επόμενη επαναληπτική διαδικασία λαμβάνεται το

$$x^{p+1}=\bar{x}^p+\lambda f. \tag{A.9.24}$$

δ) Ως διευθύνσεις αναζήτησης του νέου μεγίστου λαμβάνονται οι εξής:

$$\begin{aligned} d_i^{p+1}&=d_{i+1}^p, \quad i=1,2,\ldots,n-1 \\ d_n^{p+1}&=f. \end{aligned} \tag{A.9.25}$$

Αρχικά, ως διευθύνσεις αναζήτησης του μεγίστου λαμβάνονται οι διευθύνσεις των καρτεσιανών αξόνων. Δοθέντος του τρόπου επιλογής των νέων διευθύνσεων , που περιγράψαμε πιο πάνω, αποδεικνύεται ότι, αν η συνάρτηση που θα μεγιστοποιήσουμε είναι τετραγωνική, τότε, μετά από p επαναληπτικές διαδικασίες, οι τελευταίες p διευθύνσεις $(p\leqslant n)$ που έχουν επιλεγεί για το επόμενο βήμα είναι ανά δύο συζυγείς. Συνεπώς, μετά από n επαναληπτικές διαδικασίες όλες οι διευθύνσεις είναι ανά δύο συζυγείς και μπορεί να αποδειχτεί ότι αν η συνάρτηση $f(x)$ είναι τετραγωνική τότε απαιτούνται το πολύ n επαναληπτικές διαδικασίες για τον υπολογισμό του μεγίστου.

Η μέθοδος των συζυγών διευθύνσεων του Powell έχει το μεγάλο πλεονέκτημα ότι δεν χρειάζεται ο υπολογισμός παραγώγων και ότι εξασφαλίζει τη σύγκλιση τετραγωνικών μορφών.

v) Η μέθοδος Davidon-Fletcher-Powell

Όπως ήδη έχουμε αναφέρει ένα βασικό μειονέκτημα των μεθόδων που χρησιμοποιούν το διάνυσμα g των πρώτων παραγώγων και τον πίνακα G των δευτέρων παραγώγων της $f(x)$ είναι ο αναλυτικός υπολογισμός του πίνακα G -ή των παραλλαγών του- και αυτό δεν είναι πάντα εύκολο. Για την αντιμετώπιση του προβλήματος αυτού, εκτός από τη μέθοδο των συζυγών διευθύνσεων που περιγράψαμε, επινοήθηκαν και άλλες μέθοδοι που δεν απαιτούν τον αναλυτικό υπολογισμό του πίνακα G. Η απόδοση των μεθόδων αυτών στις πρακτικές εφαρμογές είναι αξιοσημείωτη -αν και δεν έχει δικαιολογηθεί επαρκώς θεωρητικά- ακόμα και σε περιπτώσεις που η αρχική προσέγγιση βρίσκεται μακρυά από το μέγιστο.

Η πιο δημοφιλής από τις μεθόδους αυτές είναι η μέθοδος του Davidon-Flether-Powell (DFP). Είδαμε ότι στη μέθοδο Newton -Raphson η αναζήτηση του μεγίστου γίνεται σύμφωνα με την

$$x_{p+1}=x_p-G_p^{-1}g_p.$$

Αν H είναι μια αρχική προσέγγιση του G^{-1} και θέσουμε

$$d=Hg \tag{A.9.26}$$

τότε η αρχική υπόδειξη του Davidon είναι να χρησιμοποιήσουμε, αντί του H, την ακόλουθη παραλλαγμένη μορφή του:

$$H^*=H-\frac{1}{c'd}dd'-\frac{1}{c'Hc}Hcc'H' \tag{A.9.27}$$

όπου

$$c=g(x-d)-g(x). \tag{A.9.28}$$

Σύμφωνα με μεταγενέστερη υπόδειξη του Powell η απόδοση της μεθόδου είναι καλύτερη αν αντί της (A.9.27) χρησιμοποιηθεί η

$$H^*=H-\frac{(d+Hc)(d+Hc)'}{c'd+c'Hc}. \tag{A.9.29}$$

ΒΙΒΛΙΟΓΡΑΦΙΑ

Ανδρεαδάκης, Σ.Α., "Μαθήματα Γραμμικής Άλγεβρας", Αθήνα 1974.

Γκλάβας, Χ.Β., "Θεωρία Μητρών, Οριζουσών και Γραμμικής Αλγεβρας", Αθήνα 1973.

Δονάτος, Γ.Σ., "Εισαγωγή στα Μαθηματικά της Οικονομικής Αναλύσεως", Αθήνα 1974.

Δρεττάκης, Μ., "Γραμμική Άλγεβρα", Αθήνα 1974.

Καζαντζίδης, Γ.Σ., "Βασική Γραμμική Άλγεβρα", Αθήνα 1972.

Arnold, S.F., *The Theory of Linear Models and Multivariate Analysis*, Wiley, New York 1981.

Brand, Y., *Non-Linear Parameter Estimation*, Academic Press, New York 1974.

Bellman, R., *Introduction to Matrix Analysis*, McGraw-Hill, New York 1960.

Gantmatcher, F.R., *Matrix Theory, Vol. I.*, Chelsea Pub.Co., New York 1959.

Goldfeld S.M. and Quandt R.E., *Nonlinear Methods in Econometrics*, North-Holland, Amsterdam 1972.

Graybill, F.A., *Introduction to Matrices with Applications in Statistics*, Wadsworth, Belmot California, 1969.

Hadley, G., *Linear Algebra*, Addison-Wesley, Reading Mass. 1961.

Hadley, G., *Linear Programming*, Addison-Wesley, Reading Mass. 1962.

Heal G., G. Hughes and R. Tarling, *Linear Algebra and Linear Economics*, Macmillan, London 1974.

Horst, P., *Matrix Algebra for Social Scientists*, Holt, Rinehart and Winston, New York 1963.

Intriligator, M.D., *Mathematical Optimization and Economic theory*, Prentice-Hall, Englewood Cliffs, N.J. 1971.