Εισαγωγή στην οικονομετρία (1404)

Διάλεξη 9

Άγγελος Αλεξόπουλος
Οικονομικό Πανεπιστήμιο Αθηνών - Τμήμα Οικονομικής Επιστήμης
angelos@aueb.gr

Περίγραμμα Διάλεξης

✓ Απλό γραμμικό υπόδειγμα
- Συνάρτηση παλινδρόμησης

✓ Κλασικές υποθέσεις του γραμμικού υποδείγματος
- Γραμμικότητα, εξωγένεια, ομοσκεδαστικότητα, απουσία συσχέτισης, κανονικότητα διαταρακτικών όρων

Ανάλυση Παλινδρόμησης

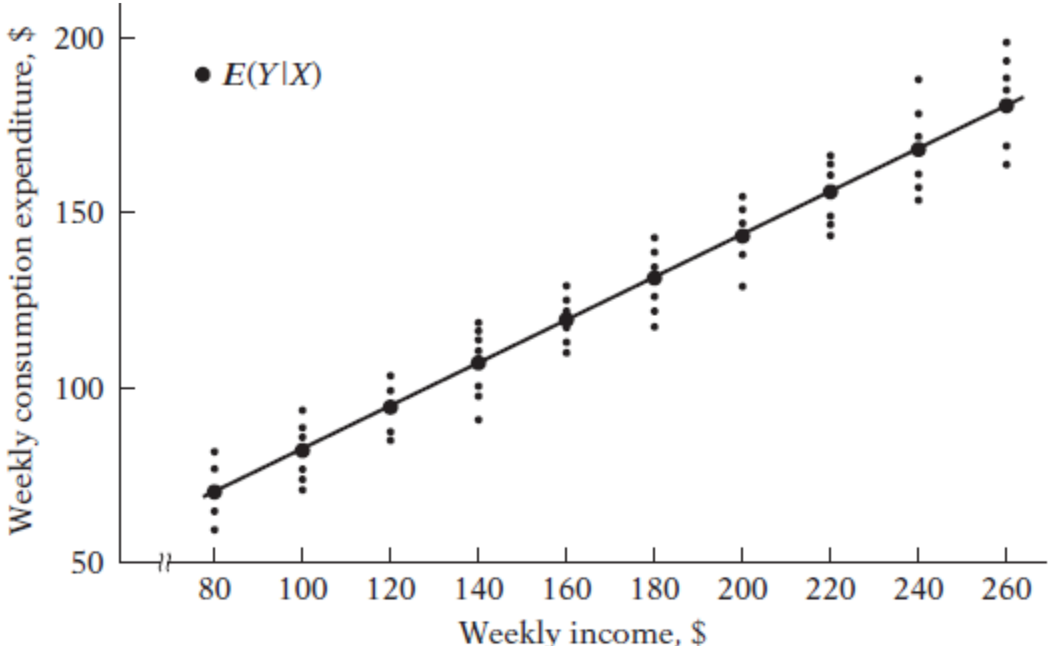
Ανάλυση παλινδρόμησης: αναφέρεται στην εκτίμηση/πρόβλεψη της μέσης τιμής (πληθυσμού) μιας εξαρτημένης μεταβλητής γνωρίζοντας τις τιμές μιας ερμηνευτικής μεταβλητής (ή περισσότερων)

Παράδειγμα: εβδομαδιαίο εισόδημα (X) και κατανάλωση (Y) νοικοκυριών σε $

- εβδομαδιαίο εισόδημα σε κατηγορίες
- για κάθε κατηγορία έχουμε $E(Y|X)$

- Κατά μέσο όρο η εβδομαδιαία κατανάλωσ αυξάνεται καθώς αυξάνεται το εισόδημα

- Αν ενώσουμε τα σημεία του δεσμευμένου μέσου $E(Y|X)$ για κάθε κατηγορία εισοδήματος προκύπτει η ***Γραμμή παλινδρόμησης πληθυσμού***



Συνάρτηση παλινδρόμησης πληθυσμού

✓Συνάρτηση παλινδρόμησης πληθυσμού (population regression function, PRF): ο δεσμευμένος μέσος $E(Y|X_i)$ είναι συνάρτηση του X_i , όπου X_i μια συγκεκριμένη τιμή της X

$$E(Y|X_i) = f(X_i)$$

$f(X_i)$ συνάρτηση της ερμηνευτικής μεταβλητής X

- Αν υποθέσουμε ότι $E(Y|X_i)$ είναι μια γραμμική συνάρτηση της X_i τότε

$$E(Y|X_i) = \beta_0 + \beta_1 X_i$$

β_0, β_1 είναι άγνωστοι παράμετροι της γραμμής παλινδρόμησης πληθυσμού

Συνάρτηση παλινδρόμησης πληθυσμού

- Για κάθε κατηγορία εισοδήματος X_i, η κατανάλωση των νοικοκυριών συγκεντρώνεται γύρω από τη μέση κατανάλωση όλων των νοικοκυριών στο X_i, δηλαδή στο δεσμευμένο μέσο
- Η απόκλιση του Y_i από την αναμενόμενη τιμή ορίζεται ως

$u_i = Y_i - E(Y|X_i)$ ή $Y_i = E(Y|X_i) + u_i$

όπου u_i τυχαία μεταβλητή που ονομάζεται

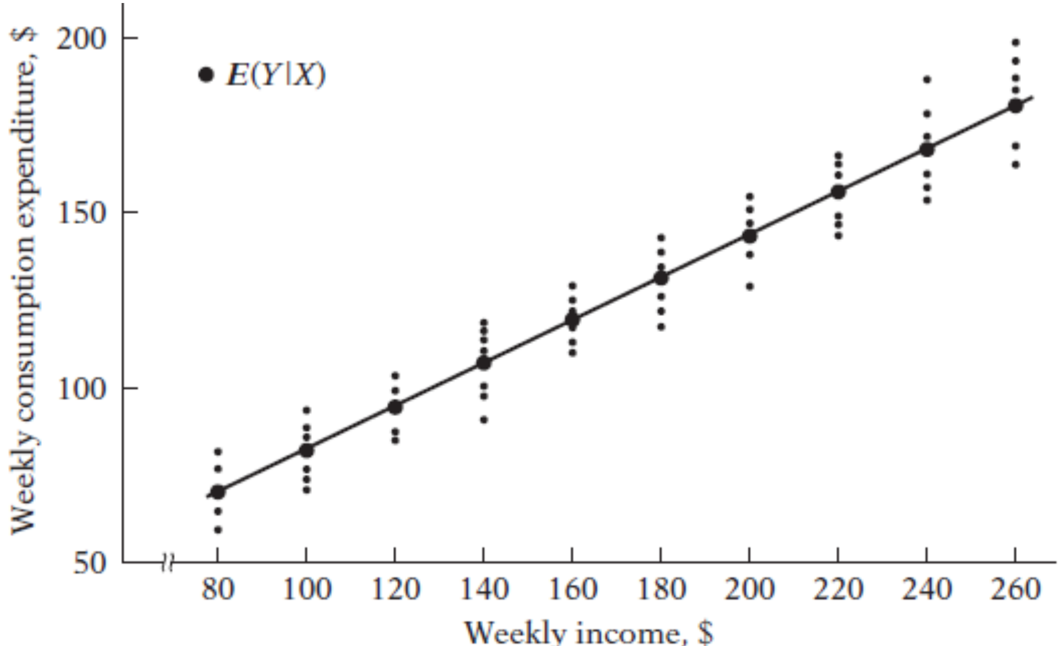
<u>διαταρακτικός όρος</u> ή <u>σφάλμα</u>

- Αν υποθέσουμε ότι $E(Y|X_i)$

είναι μια γραμμική συνάρτηση της X_i τότε

$$Y_i = E(Y|X_i) + u_i$$
$$= \beta_0 + \beta_1 X_i + u_i$$

<u>Υπόδειγμα γραμμικής παλινδρόμησης</u>



Απλό γραμμικό υπόδειγμα

Υπόδειγμα γραμμικής παλινδρόμησης με μία ερμηνευτική μεταβλητή ή απλό γραμμικό υπόδειγμα (linear regression model)

$$Y_i = \beta_0 + \beta_1 X_i + u_i, i = 1, \ldots, n$$

- Y_i είναι η ερμηνευμένη μεταβλητή (explained) του οικονομετρικού υποδείγματος, που αναφέρεται ως <u>εξαρτημένη μεταβλητή</u> (dependent variable)
- X_i η ερμηνευτική (explanatory) μεταβλητή, που αναφέρεται ως <u>ανεξάρτητη μεταβλητή</u> (independent variable)
  - στοχαστική δηλαδή τυχαία μεταβλητή και άρα περιγράφεται από μια κατανομή
  - σταθερή ή προκαθορισμένη σε επαναλαμβανόμενα δείγματα, $E(X_i) = X_i$ για όλες τις παρατηρήσεις του δείγματος

Απλό γραμμικό υπόδειγμα

$$Y_i = \beta_0 + \beta_1 X_i + u_i, i = 1, \dots, n$$

- u_i είναι ο διαταρακτικός όρος ή το σφάλμα, που αποτελεί το **ανερμήνευτο** μέρος του υποδείγματος παλινδρόμησης
  - αποτελεί μια τυχαία μεταβλητή που αντιπροσωπεύει τυχαίες (μη συστηματικές) αποκλίσεις των προβλέψεων του οικονομικού υποδείγματος για την παρατηρούμενη τιμή
  - εμπεριέχει όλους τους άλλους παράγοντες εκτός από την X που καθορίζουν την τιμή της εξαρτημένης μεταβλητής Y για μια συγκεκριμένη παρατήρηση i
  - σφάλματα στη μέτρηση οικονομικών μεγεθών, απρόβλεπτες διαρθρωτικές αλλαγές στην οικονομία, φυσικές καταστροφές, παράληψη μεταβλητών

Απλό γραμμικό υπόδειγμα

$$Y_i = \beta_0 + \beta_1 X_i + u_i, i = 1, \dots, n$$

- β_0, β_1 είναι οι συντελεστές (coefficients) ή παράμετροι (parameters) της γραμμής παλινδρόμησης πληθυσμού

- β_0 είναι ο σταθερός όρος της γραμμής παλινδρόμησης (intercept)
  - Η τιμή της γραμμής παλινδρόμησης όταν $X = 0$
- β_1 είναι η κλίση της γραμμής παλινδρόμησης (slope)
  - Η μεταβολή στην Y που σχετίζεται με μια μεταβολή της X κατά μία μονάδα
  - $\Delta Y = \beta_1 \Delta X + \Delta u \Rightarrow \frac{\Delta Y}{\Delta X} = \beta_1$, αν $\Delta u = 0$ (ceteris paribus)
  - ceteris paribus: οι παράγοντες που εισέρχονται στο διαταρακτικό όρο δεν σχετίζονται με την X_i (είναι σταθεροί)

Απλό γραμμικό υπόδειγμα

$$Y_i = \beta_0 + \beta_1 X_i + u_i, i = 1, \ldots, n$$

- $\beta_0 + \beta_1 X_i$ αποτελεί το **ερμηνευμένο** μέρος από την οικονομική θεωρία της οικονομετρικού υποδείγματος (παλινδρόμησης), γραμμή ή συνάρτηση παλινδρόμησης πληθυσμού (population regression line or function)
    - $E(Y_i|X_i) = f(X_i) = \beta_0 + \beta_1 X_i$ γραμμική συνάρτηση παλινδρόμησης πληθυσμού

Υποθέσεις γραμμικού υποδείγματος

✓Γραμμικότητα (linearity) του υποδείγματος

- Αναφέρεται στους συντελεστές του υποδείγματος και όχι στις ανεξάρτητες μεταβλητές

$$Y_i = \beta_0 + \beta_1 X_i + u_i, i = 1, \ldots, n$$

- Η ιδιότητα αυτή ικανοποιείται αν

$$\frac{\partial Y_i}{\partial \beta_0} = 1 \text{ και } \frac{\partial Y_i}{\partial \beta_1} = \mathrm{X}_i$$

οι παράγωγοι της Y_i ως προς τις παραμέτρους β_0 και β_1 είναι ανεξάρτητες των παραμέτρων

Υποθέσεις γραμμικού υποδείγματος

✓Το γραμμικό σφάλμα ικανοποιεί την συνθήκη της **ισχυρής εξωγένειας**

$$E(u_i|X_i) = 0$$

$$\begin{aligned} E(u_i|X_i) &= E([Y_i - E(Y_i|X_i)]|X_i) \\ &= E(Y_i|X_i) - E(E(Y_i|X_i)|X_i) \\ &= E(Y_i|X_i) - E(Y_i|X_i) = 0 \end{aligned}$$

συνεπάγεται ότι **(1)** $E(u_i) = 0$

$$E(u_i) = E\big(E(u_i|X_i)\big) = E(0) = 0$$

- χρήση Ν.Ε.Π. $E(Y) = E[E(Y|X)]$

και **(2)** $E(X_i u_i) = 0$

- οποιαδήποτε συνάρτηση των X_i δεν σχετίζεται με τον διαταρακτικό όρο

Υποθέσεις γραμμικού υποδείγματος

(συνέχεια)

- Εξωγένεια: $E(u_i) = 0$ (όταν η ερμηνευτική μεταβλητή είναι σταθερή σε επαναλαμβανόμενα δείγματα)

- Ισχυρή εξωγένεια: $E(u_i|X_i) = 0$ (όταν η ερμηνευτική μεταβλητή είναι στοχαστική δηλαδή τυχαία μεταβλητή)

Υποθέσεις γραμμικού υποδείγματος

✓ Διακύμανση διαταρακτικού όρου

- Ομοσκεδαστικότητα: η διακύμανση του u_i είναι σταθερή για κάθε παρατήρηση i

$$Var(u_i) = \sigma^2 \ \forall i$$

ή όταν η ερμηνευτική μεταβλητή είναι στοχαστική

$$Var(u_i|X_i) = \sigma^2 \ \forall i$$

Υποθέσεις γραμμικού υποδείγματος

✓ Απουσία συσχέτισης μεταξύ των διαταρακτικών όρων

$$Cov(u_i, u_j) = E(u_i u_j) - E(u_i)\mathrm{E}(u_j) = 0 \ \forall i \neq j$$

ή

$$Cov(u_i, u_j | X_i) = 0 \ \forall i \neq j$$

- Οι υποθέσεις του διαταρακτικού όρου αναφορικά με την διακύμανση κ τη συνδιακύμανση αναφέρονται ως η υπόθεση της σφαιρικότητας

$$E(u_i u_j) = \begin{cases} \sigma_u^2, i = j, \forall i = j \\ 0, i \neq j, \forall i \neq j \end{cases}$$

Υποθέσεις γραμμικού υποδείγματος

✓Κατανομή του διαταρακτικού όρου

$$u_i \sim N(0, \sigma^2), \forall i$$

ή

$$u_i | X_i \sim N(0, \sigma^2), \forall i$$

- Η κατανομή του διαταρακτικού όρου ακολουθεί την κανονική κατανομή με μέσο μηδέν και διακύμανση σ^2

Άλλες ιδιότητες

- Υπόθεση για τις τιμές της ανεξάρτητης μεταβλητής όταν X_i προκαθορισμένη

$$cov(X_i, u_i) = E(X_i u_i) - E(X_i)\mathrm{E}(u_i) = E(X_i u_i) = X_i E(u_i) = 0$$

- $E(X_i) = X_i$ *και* $\mathrm{E}(u_i) = 0$

και $cov(X_i, u_i | X_i) = 0$ (X_i στοχαστική)

Άλλες ιδιότητες

- Ιδιότητες εξαρτημένης μεταβλητής

Μέσος:

$$E(Y_i) = E(\beta_0 + \beta_1 X_i + u_i) = \beta_0 + \beta_1 E(X_i) + E(u_i) = \beta_0 + \beta_1 X_i$$

Διακύμανση:

$$Var(Y_i) = Var(\beta_0 + \beta_1 X_i + u_i)$$
$$= Var(\beta_0 + \beta_1 X_i) + Var(u_i) + 2Cov(\beta_0 + \beta_1 X_i, u_i)$$
$$= Var(\beta_0) + Var(\beta_1 X_i) + Var(u_i) + 2Cov(\beta_0, \beta_1 X_i) + 2Cov(\beta_0, u_i)$$
$$+ 2Cov(\beta_1 X_i, u_i)$$
$$= 0 + 0 + \sigma^2 + 0 + 0 + 0 = \sigma^2 \ \forall i$$

Άλλες ιδιότητες

- Ιδιότητες εξαρτημένης μεταβλητής (συνέχεια)

Συνδιακύμανση:

$$Cov(Y_i, Y_j) = cov\left(\beta_0 + \beta_1 X_i + u_i, \beta_0 + \beta_1 X_j + u_j\right) = 0$$

Κατανομή:

$$Y_i \sim N[E(Y_i), Var(Y_i)] \text{ επομένως } Y_i \sim N(\beta_0 + \beta_1 X_i, \sigma^2)\ \forall i$$

Κλασικές υποθέσεις - σύνοψη

Αν η X_i είναι στοχαστική, δηλαδή τυχαία μεταβλητή

- Πιο ρεαλιστική

1. Γραμμικότητα ως προς τις παραμέτρους
2. $E(u_i|X_i) = 0, \forall i$
3. $Var(u_i|X_i) = E\left(u_i^2|X_i\right) = \sigma^2, \forall i$
4. $Cov\left(u_i, u_j|X_i\right) = 0, \forall i \neq j$
5. $u_i|X_i \sim N(0, \sigma^2), \forall i$

Κλασικές υποθέσεις - σύνοψη

Αν X_i είναι σταθερή σε επαναλαμβανόμενα δείγματα

1. Γραμμικότητα ως προς τις παραμέτρους
2. $E(u_i) = 0, \forall i$
3. $Var(u_i) = E(u_i^2) = \sigma^2, \forall i$
4. $Cov(u_i, u_j) = 0, \forall i \neq j$
5. $u_i \sim N(0, \sigma^2), \forall i$

Περίληψη

- Απλό γραμμικό υπόδειγμα

- Κλασικές υποθέσεις
  - Γραμμικότητα
  - Μηδενική (δεσμευμένη) αναμενόμενη τιμή του διαταρακτικού όρου, εξωγένεια
  - Ομοσκεδαστικότητα
  - Απουσία συσχέτισης
  - Κανονικότητα διαταρακτικών όρων

Βιβλιογραφία

- ✓ Stock & Watson, κεφ. 4 (4.1 & 4.4)
- ✓ Gujarati & Porter, κεφ. 2 & 3
- ✓ Τζαβαλής, κεφ. 2

Ευχαριστώ!

Ερωτήσεις?