

Οικονομετρία II

Διάλεξη: Ο Εκτιμητής LAD

Διάμεσοι, Απόλυτες Αποκλίσεις και το Γραμμικό Υπόδειγμα

Στέλιος Αρβανίτης

Τμήμα Οικονομικής Επιστήμης, Ο.Π.Α.

Άνοιξη 2026

Περιεχόμενα

1	Αφετηρία: από τον OLSE στον LADE	2
2	Η έννοια της διαμέσου	2
3	Η διάμεσος ως ελαχιστοποιητής απόλυτης απόκλισης	3
4	Επιστροφή στο γραμμικό υπόδειγμα	4
5	Το πληθυσμιακό LAD κριτήριο	5
6	Ο LADE	6
7	LADE, OLSE και βαριές ουρές	7
8	Υπολογισμός του LADE ως γραμμικό πρόγραμμα	7
9	Στοιχεία ασυμπτωτικής θεωρίας	8
10	Σύντομη σύνοψη	8

1 Αφετηρία: από τον OLSE στον LADE

Στο βασικό γραμμικό υπόδειγμα γράφουμε

$$Y_n = X_n\beta_0 + \varepsilon_n,$$

ή, γραμμή προς γραμμή,

$$Y_{(i)} = X_{(i)}\beta_0 + \varepsilon_{(i)}, \quad i = 1, \dots, n.$$

Στον OLSE η παράμετρος εκτιμάται μέσω της ελαχιστοποίησης του μέσου τετραγωνικού σφάλματος:

$$\hat{\beta}_n^{OLS} \in \arg \min_{\beta \in \Theta} \frac{1}{n} \sum_{i=1}^n (Y_{(i)} - X_{(i)}\beta)^2.$$

Το τετραγωνικό κριτήριο είναι φυσικό όταν θέλουμε να περιγράψουμε την υπό συνθήκη μέση τιμή:

$$\mathbb{E}(Y_{(i)} \mid X_n) = X_{(i)}\beta_0.$$

Στη διάλεξη αυτή εξετάζουμε μια εναλλακτική εκτίμηση, τον εκτιμητή ελαχίστων απόλυτων αποκλίσεων:

$$\hat{\beta}_n^{LAD} \in \arg \min_{\beta \in \Theta} \frac{1}{n} \sum_{i=1}^n |Y_{(i)} - X_{(i)}\beta|.$$

Ο εκτιμητής αυτός ονομάζεται Least Absolute Deviations Estimator ή LADE.

Διαισθητική ερμηνεία

Ο OLSE στηρίζεται στη λογική της μέσης τιμής. Ο LADE στηρίζεται στη λογική της διαμέσου. Έτσι, ενώ ο OLSE περιγράφει τη σχέση των X 's με την υπό συνθήκη μέση τιμή του $Y_{(i)}$, ο LADE περιγράφει τη σχέση των X 's με την υπό συνθήκη διάμεσο του $Y_{(i)}$.

2 Η έννοια της διαμέσου

Έστω πραγματική τυχαία μεταβλητή ε με συνάρτηση κατανομής

$$F(x) = \mathbb{P}(\varepsilon \leq x).$$

Ορισμός 2.1 (Διάμεσος). Ένας αριθμός $m \in \mathbb{R}$ ονομάζεται διάμεσος της ε αν

$$\mathbb{P}(\varepsilon \leq m) \geq \frac{1}{2} \quad \text{και} \quad \mathbb{P}(\varepsilon \geq m) \geq \frac{1}{2}.$$

Το σύνολο όλων των διαμέσων συμβολίζεται με

$$\text{Med}(\varepsilon).$$

Όταν η κατανομή είναι συνεχής και αυστηρά αύξουσα γύρω από τη διάμεσο, η παραπάνω συνθήκη ισοδυναμεί με

$$F(m) = \frac{1}{2}.$$

Προσοχή

Η εξίσωση $F(m) = 1/2$ δεν είναι ο γενικός ορισμός της διαμέσου. Είναι μια βολική μορφή που ισχύει όταν η κατανομή δεν έχει άλματα στη διάμεσο. Σε διακριτές κατανομές μπορεί να υπάρχει ολόκληρο διάστημα διαμέσων.

Παράδειγμα 2.1 (Διακριτή κατανομή με μη μοναδική διάμεσο). Έστω

$$\varepsilon = \begin{cases} -1, & \text{με πιθανότητα } 1/2, \\ 1, & \text{με πιθανότητα } 1/2. \end{cases}$$

Τότε κάθε

$$m \in [-1, 1]$$

είναι διάμεσος. Πράγματι, για κάθε τέτοιο m ,

$$\mathbb{P}(\varepsilon \leq m) \geq \frac{1}{2} \quad \text{και} \quad \mathbb{P}(\varepsilon \geq m) \geq \frac{1}{2}.$$

Πρόταση 2.1 (Ικανή συνθήκη μοναδικότητας). *Μια συνήθης ικανή συνθήκη για μοναδικότητα είναι το να υπάρχει η συνάρτηση πυκνότητας f , και για κάποια τιμή στης διαμέσου, έστω m να ισχύει ότι*

$$f(m) > 0.$$

3 Η διάμεσος ως ελαχιστοποιητής απόλυτης απόκλισης

Το βασικό θεωρητικό γεγονός πίσω από τον LADE είναι ότι η διάμεσος ελαχιστοποιεί την αναμενόμενη απόλυτη απόκλιση.

Θεωρούμε τη συνάρτηση

$$\psi(c) = \mathbb{E}|\varepsilon - c|.$$

Για να είναι καλά ορισμένη και πεπερασμένη, αρκεί

$$\mathbb{E}|\varepsilon| < \infty.$$

Πρόταση 3.1 (Η διάμεσος ελαχιστοποιεί την αναμενόμενη απόλυτη απόκλιση). *Αν $\mathbb{E}|\varepsilon| < \infty$, τότε*

$$\arg \min_{c \in \mathbb{R}} \mathbb{E}|\varepsilon - c| = \text{Med}(\varepsilon).$$

Αν η διάμεσος είναι μοναδική, τότε ο μοναδικός ελαχιστοποιητής είναι αυτή η διάμεσος.

Απόδειξη. Για $c_2 > c_1$, μπορούμε να γράψουμε τη διαφορά

$$\psi(c_2) - \psi(c_1) = \mathbb{E}(|\varepsilon - c_2| - |\varepsilon - c_1|).$$

Η συνάρτηση $c \mapsto |\varepsilon - c|$ είναι κυρτή. Επομένως και η ψ είναι κυρτή ως αναμενόμενη τιμή κυρτών συναρτήσεων.

Αν η κατανομή είναι συνεχής στο c , μπορούμε να παραγωγίσουμε:

$$\psi'(c) = \mathbb{P}(\varepsilon < c) - \mathbb{P}(\varepsilon > c).$$

Η συνθήκη πρώτης τάξης γίνεται

$$\mathbb{P}(\varepsilon < c) = \mathbb{P}(\varepsilon > c),$$

δηλαδή, στην συνεχή περίπτωση,

$$\mathbb{P}(\varepsilon \leq c) = \frac{1}{2}.$$

Άρα το ελάχιστο επιτυγχάνεται στη διάμεσο.

Στη γενική περίπτωση, η ίδια ιδέα γράφεται με υποπαραγώγους:

$$0 \in \partial\psi(c) \iff \mathbb{P}(\varepsilon \leq c) \geq \frac{1}{2} \text{ και } \mathbb{P}(\varepsilon \geq c) \geq \frac{1}{2}.$$

Αυτή είναι ακριβώς η γενική συνθήκη διαμέσου. □

Σημείωση

Η απόδειξη μέσω υποπαραγώγων είναι πιο γενική, διότι δεν χρειάζεται να αγνοήσουμε τα σημεία μη παραγωγισιμότητας της απόλυτης τιμής. Για την οικονομετρική διαίσθηση, αρκεί να θυμόμαστε ότι η ελαχιστοποίηση της αναμενόμενης απόλυτης απόκλισης οδηγεί στη διάμεσο.

4 Επιστροφή στο γραμμικό υπόδειγμα

Υποθέτουμε το γραμμικό υπόδειγμα

$$Y_{(i)} = X_{(i)}\beta_0 + \varepsilon_{(i)}.$$

Η LAD λογική δεν χρειάζεται την υπόθεση

$$\mathbb{E}(\varepsilon_{(i)} \mid X_n) = 0.$$

Αντί αυτής, χρησιμοποιεί υπόθεση μηδενικής υπό συνθήκη διαμέσου:

$$0 \in \text{Med}(\varepsilon_{(i)} \mid X_n), \quad i = 1, \dots, n.$$

Όταν η υπό συνθήκη διάμεσος είναι μοναδική, γράφουμε απλά

$$\text{Med}(\varepsilon_{(i)} \mid X_n) = 0.$$

Τότε, επειδή

$$Y_{(i)} = X_{(i)}\beta_0 + \varepsilon_{(i)},$$

έχουμε

$$\text{Med}(Y_{(i)} \mid X_n) = X_{(i)}\beta_0.$$

Διαισθητική ερμηνεία

Η υπόθεση μηδενικής υπό συνθήκη διαμέσου σημαίνει ότι, υπό τις πληροφορίες του X_n , το σφάλμα έχει ίση πιθανότητα να είναι θετικό ή αρνητικό. Άρα το $X_{(i)}\beta_0$ δεν είναι η υπό συνθήκη μέση τιμή του $Y_{(i)}$, αλλά η υπό συνθήκη διάμεσος του $Y_{(i)}$.

5 Το πληθυσμιακό LAD κριτήριο

Θεωρούμε το υπό συνθήκη πληθυσμιακό κριτήριο

$$\mu_n^*(\beta) = \mathbb{E} \left[\frac{1}{n} \sum_{i=1}^n |Y_{(i)} - X_{(i)}\beta| \mid X_n \right].$$

Από την γραμμικότητα του ολοκληρώματος,

$$\mu_n^*(\beta) = \frac{1}{n} \sum_{i=1}^n \mathbb{E} [|Y_{(i)} - X_{(i)}\beta| \mid X_n].$$

Αντικαθιστώντας το DGP,

$$Y_{(i)} = X_{(i)}\beta_0 + \varepsilon_{(i)},$$

λαμβάνουμε

$$\mu_n^*(\beta) = \frac{1}{n} \sum_{i=1}^n \mathbb{E} [|\varepsilon_{(i)} - X_{(i)}(\beta - \beta_0)| \mid X_n].$$

Για κάθε i , το εσωτερικό criterion ελαχιστοποιείται όταν

$$X_{(i)}(\beta - \beta_0)$$

είναι διάμεσος της υπό συνθήκη κατανομής του $\varepsilon_{(i)} \mid X_n$. Υπό μηδενική και μοναδική διάμεσο, αυτό σημαίνει

$$X_{(i)}(\beta - \beta_0) = 0.$$

Αν η συνθήκη αυτή ισχύει για όλα τα i , τότε

$$X_n(\beta - \beta_0) = 0.$$

Αν επιπλέον

$$\text{rank}(X_n) = p,$$

τότε αναγκαστικά (γιατί;)

$$\beta = \beta_0.$$

Πρόταση 5.1 (Ταυτοποίηση στο LAD κριτήριο). *Υποθέστε ότι*

$$0 = \text{Med}(\varepsilon_{(i)} \mid X_n)$$

είναι μοναδική υπό συνθήκη διάμεσος για κάθε i , και ότι

$$\text{rank}(X_n) = p.$$

Τότε το $\mu_n^(\beta)$ ελαχιστοποιείται μοναδικά στο β_0 .*

Απόδειξη. Για κάθε i , η ποσότητα

$$\mathbb{E} [|\varepsilon_{(i)} - c| \mid X_n]$$

ελαχιστοποιείται μοναδικά στο $c = 0$. Στο LAD κριτήριο ο ρόλος του c παίζεται από

$$c_i(\beta) = X_{(i)}(\beta - \beta_0).$$

Άρα κάθε όρος ελαχιστοποιείται όταν $c_i(\beta) = 0$. Όλοι οι όροι ελαχιστοποιούνται ταυτόχρονα όταν

$$X_n(\beta - \beta_0) = 0.$$

Η πλήρης στήλη τάξη του X_n συνεπάγεται ότι αυτό ισχύει μόνο για $\beta = \beta_0$. □

6 Ο LADE

Το θεωρητικό κριτήριο $\mu_n^*(\beta)$ δεν είναι παρατηρήσιμο. Αντικαθίσταται από το δειγματικό ανάλογο:

$$\mu_n(\beta) = \frac{1}{n} \sum_{i=1}^n |Y_{(i)} - X_{(i)}\beta|.$$

Ορισμός 6.1 (LADE). Ο εκτιμητής ελαχίστων απόλυτων αποκλίσεων ορίζεται ως

$$\hat{\beta}_n^{LAD} \in \arg \min_{\beta \in \Theta} \mu_n(\beta) = \arg \min_{\beta \in \Theta} \frac{1}{n} \sum_{i=1}^n |Y_{(i)} - X_{(i)}\beta|.$$

Ο LADE είναι M -εκτιμητής, με κριτήριο

$$m(Y_{(i)}, X_{(i)}, \beta) = |Y_{(i)} - X_{(i)}\beta|.$$

Σημείωση

Σε αντίθεση με τον OLSE, το LAD κριτήριο είναι κυρτό αλλά όχι παντού παραγωγίσιμο. Η μη παραγωγισιμότητα εμφανίζεται όταν κάποιο κατάλοιπο είναι μηδέν:

$$Y_{(i)} - X_{(i)}\beta = 0.$$

Αυτό επηρεάζει τον τρόπο υπολογισμού του εκτιμητή.

7 LAD, OLSE και βαριές ουρές

Η διαδικασία βελτιστοποίησης που οδηγεί στον OLSE τιμωρεί τα μεγάλα κατάλοιπα τετραγωνικά:

$$r_i^2, \quad r_i = Y_{(i)} - X_{(i)}\beta.$$

Η διαδικασία βελτιστοποίησης που οδηγεί στον LAD τα τιμωρεί γραμμικά:

$$|r_i|.$$

Διαισθητική ερμηνεία

Αυτός είναι ο λόγος που ο LAD είναι συχνά πιο ανθεκτικός σε ακραίες παρατηρήσεις και παχείς ουρές. Ένα πολύ μεγάλο κατάλοιπο μπορεί να κυριαρχήσει στο OLS κριτήριο, αλλά έχει πολύ πιο περιορισμένη επίδραση στο LAD κριτήριο.

Αν τα σφάλματα έχουν κατανομή t_ν με μικρούς βαθμούς ελευθερίας, τότε μπορεί να μην υπάρχουν αρκετές ροπές για να στηριχθούν οι συνηθισμένες OLS προσεγγίσεις. Για παράδειγμα:

- για $\nu > 2$, η διακύμανση είναι πεπερασμένη,
- για $\nu \leq 2$, η διακύμανση δεν είναι πεπερασμένη,
- για $\nu = 1$, έχουμε Cauchy κατανομή, όπου δεν υπάρχει ούτε πεπερασμένη μέση τιμή.

Το LAD criterion απαιτεί πολύ ηπιότερη πληροφορία σχετικά με τις ουρές, εφόσον η υπό συνθήκη διάμεσος είναι καλά ορισμένη.

8 Υπολογισμός του LAD ως γραμμικό πρόγραμμα

Το LAD πρόβλημα μπορεί να γραφεί ως γραμμικό πρόγραμμα. Θέτουμε

$$r_i = Y_{(i)} - X_{(i)}\beta.$$

Εισάγουμε μη αρνητικές μεταβλητές $u_i, v_i \geq 0$ τέτοιες ώστε

$$r_i = u_i - v_i, \quad |r_i| = u_i + v_i.$$

Τότε το LAD πρόβλημα είναι ισοδύναμο με:

$$\min_{\beta, u, v} \frac{1}{n} \sum_{i=1}^n (u_i + v_i)$$

υπό τους περιορισμούς

$$Y_{(i)} - X_{(i)}\beta = u_i - v_i, \quad u_i \geq 0, \quad v_i \geq 0, \quad i = 1, \dots, n.$$

Σημείωση

Αυτή η αναπαράσταση εξηγεί γιατί ο LADE μπορεί να υπολογιστεί με μεθόδους γραμμικού προγραμματισμού. Επίσης εξηγεί γιατί ο LADE δεν υπολογίζεται μέσω των συνήθων κανονικών εξισώσεων του OLS.

9 Στοιχεία ασυμπτωτικής θεωρίας

Υπό κατάλληλες συνθήκες, ο LADE είναι συνεπής για β_0 . Η βασική λογική είναι ίδια με τη γενική θεωρία M -εκτιμητών:

1. το δειγματικό κριτήριο $\mu_n(\beta)$ συγκλίνει σε πληθυσμιακό κριτήριο,
2. το πληθυσμιακό κριτήριο έχει μοναδικό ελάχιστο στο β_0 ,
3. άρα ο ελαχιστοποιητής του δειγματικού κριτηρίου συγκλίνει στο β_0 .

Για την ασυμπτωτική κανονικότητα χρειάζονται πιο ειδικές συνθήκες. Σε iid πλαίσιο, και όταν η υπό συνθήκη πυκνότητα του σφάλματος στο μηδέν είναι θετική, η ασυμπτωτική διακύμανση του LADE περιλαμβάνει τη συνάρτηση πυκνότητας στο μηδέν. Σε απλή μορφή, η πληροφορία για τη διάμεσο γίνεται ισχυρότερη όταν η πυκνότητα γύρω από τη διάμεσο είναι μεγαλύτερη.

Προσοχή

Η εξαγωγή της πλήρους ασυμπτωτικής θεωρίας του LADE είναι πιο εκλεπτισμένη από του OLSE, επειδή το κριτήριο δεν είναι διαφορίσιμο. Η ανάλυση χρησιμοποιεί υποπαραγώγους, εμπειρικές ανελίξεις (empirical processes) ή ειδικά αναπτύγματα γύρω από τη διάμεσο.

10 Σύντομη σύνοψη

- Ο OLSE συνδέεται με την υπό συνθήκη μέση τιμή.
- Ο LADE συνδέεται με την υπό συνθήκη διάμεσο.
- Η διάμεσος ελαχιστοποιεί την αναμενόμενη απόλυτη απόκλιση.
- Το LAD κριτήριο είναι κυρτό αλλά μη διαφορίσιμο.
- Ο LADE είναι συχνά πιο ανθεκτικός σε παχείς ουρές και ακραίες παρατηρήσεις.

- Ο υπολογισμός του μπορεί να γίνει μέσω γραμμικού προγραμματισμού–δείτε και την επόμενη διάλεξη.