Το Γραμμικό Υπόδειγμα Student-t

Μέγιστη Πιθανοφάνεια, και Παχείς Ουρές

Στέλιος Αρβανίτης

Οικονομετρία II

0.1. Το γραμμικό υπόδειγμα Student-t

Θεωρούμε το γνωστό από τα προηγούμενα γραμμικό υπόδειγμα

$$Y_{(i)} = X'_{(i)}\beta_0 + \varepsilon_{(i)}, \qquad i = 1, \ldots, n.$$

Υποθέτουμε ότι, δεσμευμένα ως προς τη μήτρα παλινδρομητών X, οι διαταρακτικοί όροι είναι ανεξάρτητοι και ισόνομα κατανεμημένοι:

$$\varepsilon_{(i)} \mid X \overset{i.i.d.}{\sim} t_\nu,$$

όπου: $\nu > 0$ είναι οι βαθμοί ελευθερίας.

Το πλαίσιο αυτό συνεπάγεται δεσμευμένη ομοιογένεια:

$$\text{Var}(\varepsilon_{(i)} \mid X) = \begin{cases} \sigma^2 \dfrac{\nu}{\nu - 2}, & \nu > 2, \\ \text{δεν υπάρχει}, & 0 < \nu \leq 2. \end{cases}$$

και επομένως δεσμευμένη i.i.d. δομή για τα σφάλματα.

Δεσμευμένη πυκνότητα. Η δεσμευμένη πυκνότητα του $Y_{(i)}$, δεδομένου του X_n, εφόσον η πραγματική κατανομή αντιστοιχούσε στην τιμή β, γράφεται:

$$f(y_{(i)} \mid X_{(i)}; \beta, \nu) = \frac{\Gamma\left(\frac{\nu+1}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)\sqrt{\nu\pi}} \left(1 + \frac{(y_{(i)} - X'_{(i)}\beta)^2}{\nu}\right)^{-(\nu+1)/2}.$$

Παρατηρούμε ότι η πυκνότητα εξαρτάται από το: $u_{(i)}(\beta) = y_{(i)} - X'_{(i)}\beta$, και σε αντίθεση με το Γκαουσσιανό υπόδειγμα, ο τρόπος με τον οποίο μεγάλες τιμές του $u_{(i)}(\beta)$ επηρεάζουν την παραπάνω συνάρτηση είναι πολυωνυμικός και όχι εκθετικός, γεγονός που δημιουργεί βαρύτερες ουρές.

Συνάρτηση πιθανοφάνειας. Υπό την παραπάνω δεσμευμένη ανεξαρτησία:

$$L_n(\beta, \nu) = \prod_{i=1}^{n} f(y_{(i)} \mid X_{(i)}; \beta, \nu).$$

Λαμβάνοντας λογαρίθμους και διαιρώντας μέ n:

$$\ell_n(\beta, \sigma^2, \nu) = \frac{1}{n} \sum_{i=1}^{n} \log f(y_{(i)} \mid X_{(i)}; \beta, \nu).$$

Άρα:

$$\begin{aligned} \ell_n = \log \Gamma\left(\frac{\nu+1}{2}\right) - \log \Gamma\left(\frac{\nu}{2}\right) - \frac{1}{2} \log(\nu \pi \sigma^2) \\ - \frac{\nu+1}{2} \frac{1}{n} \sum_{i=1}^{n} \log\left(1 + \frac{(y_{(i)} - X'_{(i)}\beta)^2}{\nu}\right). \end{aligned}$$

Ο εκτιμητής μέγιστης πιθανοφάνειας ορίζεται ως:

$$(\hat{\beta}, \hat{\nu}) \in \arg\max_{(\beta,\nu)} \ell_n(\beta, \sigma^2, \nu).$$

Σχέση με τον OLSE. Στο Gaussian γραμμικό υπόδειγμα ο MLE συμπίπτει με τον OLSE, διότι η λογαριθμική πιθανοφάνεια είναι ισοδύναμη με την ελαχιστοποίηση: $\sum_{i=1}^{n} (y_{(i)} - X'_{(i)}\beta)^2$.

Αντίθετα, στο Student-t υπόδειγμα εμφανίζεται το:

$$\frac{1}{n} \sum_{i=1}^{n} \log\left(1 + \frac{(y_{(i)} - X'_{(i)}\beta)^2}{\nu}\right),$$

το οποίο δεν είναι τετραγωνικό, κάτι το όποίο αποτελεί ένδειξη του ότι: $\hat{\beta}_{\text{MLE}} \neq \hat{\beta}_{\text{OLS}}$ γενικά, ακόμη και όταν το ν είναι γνωστό.

Οικονομετρική σημασία. Το Student-t υπόδειγμα είναι μπορεί να είναι χρήσιμο όταν π.χ. υπάρχουν ενδείξεις ότι η δεσμευμένη κατανομή των σφαλμάτων παρουσιάζει παχείς ουρές, με την μη ύπαρξη της δεύτερης ή και της πρώτης ροπής.

Η μορφή της πιθανοφάνειας δημιουργεί μεγαλύτερη ανθεκτικότητα σε ακραία κατάλοιπα, καθώς η συνεισφορά τους επηρεάζει λογαριθμικά αντί τετραγωνικά την συνάρτηση πιθανοφάνειας.

Ροπές και βαθμοί ελευθερίας. Από τις ιδιότητες της κατανομής Student-t, η ύπαρξη ροπών εξαρτάται από το : ν.

Συγκεκριμένα αποδεικνύεται εύκολα ότι:

$$\mathbb{E}|\varepsilon_{(i)}|^k < \infty \iff \nu > k.$$

Άρα:

- αν $\nu \leq 2$, τότε η διακύμανση δεν υπάρχει,
- ενώ αν $0 < \nu \leq 1$, τότε δεν υπάρχει ούτε ο μέσος.

Παρά ταύτα, η παραπάνω παραμετρική εξειδίκευση μπορεί ακόμη να επιτρέπει την κατασκευή εκτιμητών με κλασικούς ασυμπτωτικούς ρυθμούς και Γκαουσσίανή ασυμπτωτική κατανομή υπό κατάλληλες συνθήκες.

Το Γκαουσιανό υπόδειγμα ως οριακή περίπτωση. Αποδεικνύεται ότι καθώς $\nu \to \infty$, ισχύει:

$$t_\nu(0, \sigma^2) \rightsquigarrow N(0, \sigma^2).$$

Συνεπώς το Γκαουσσιανό γραμμικό υπόδειγμα εμφανίζεται ως οριακή περίπτωση του εν λόγω Student-t υποδείγματος.

Υπολογιστικές παρατηρήσεις. Το πρόβλημα μεγιστοποίησης της πιθανοφάνειας δεν διαθέτει κλειστή μορφή λύσης και απαιτεί αριθμητική βελτιστοποίηση χρησιμποιώντας π.χ. αλγορίθμους όπως:

- Newton–Raphson,
- Fisher scoring,
- quasi-Newton μεθόδους (BFGS),
- κ.λ.π.

Όταν: ν είναι άγνωστο, η βελτιστοποίηση γίνεται δυσκολότερη, διότι:

- η πιθανοφάνεια γίνεται πιο επίπεδη,
- εμφανίζονται ισχυρότερες μη γραμμικότητες,
- η ταυτοποίηση του ν είναι ασθενέστερη για μεγάλα ν.

Επεκτάσεις. Το υπόδειγμα επεκτείνεται φυσικά:

- σε πλαίσια δεσμευμένης ετερογένειας:

$$\nu_{(i)}^2 = \nu(X_{(i)}),$$

- σε δυναμικά υποδείγματα χρονοσειρών,
- σε Student-t GARCH υποδείγματα,
- κ.ο.κ.