

Σταθμισμένα ελάχιστα τετράγωνα (WLS, Weighted Least Squares)

Πρόταση. Έστω το γραμμικό μοντέλο $\mathbf{Y} = \mathbf{X} \cdot \boldsymbol{\beta} + \mathbf{e}$, $\mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{V})$, όπου $\boldsymbol{\beta} = (\beta_0, \dots, \beta_{p-1})$. Η καλύτερη γραμμική αμερόληπτη εκτιμήτρια του $\boldsymbol{\beta}$ είναι η

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} \mathbf{Y}$$

με πίνακα διασποράς $V(\hat{\boldsymbol{\beta}}) = \sigma^2 (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1}$.

Απόδειξη. Στο κλασσικό γραμμικό μοντέλο ($\mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{I}_n)$), ως γνωστό παίρνουμε ως εκτιμήτρια ελαχίστων τετραγώνων (ή ισοδύναμα MLE) του $\boldsymbol{\beta}$ την

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}.$$

Αν όμως $\mathbf{e} \sim N(\mathbf{0}, \sigma^2 \mathbf{V})$, θα πρέπει να μετασχηματίσουμε το μοντέλο $\mathbf{Y} = \mathbf{X} \cdot \boldsymbol{\beta} + \mathbf{e}$ έτσι ώστε τα σφάλματα να είναι ανεξάρτητα για να μπορέσουμε στην συνέχεια να χρησιμοποιήσουμε την παραπάνω εκτιμήτρια ελαχίστων τετραγώνων. Ο πίνακας $\mathbf{V}$ ως πίνακας διασποράς είναι συμμετρικός και ημιοριστικά θετικός (και άρα έχει μη αρνητικές ιδιοτιμές). Επομένως θα υπάρχει πίνακας $\mathbf{Q}$ αντιστρέψιμος ώστε $\mathbf{V} = \mathbf{Q} \cdot \mathbf{Q}^T$ (βλ. υποσημείωση¹).

Στο μοντέλο τώρα $\mathbf{Y} = \mathbf{X} \cdot \boldsymbol{\beta} + \mathbf{e}$ πολλαπλασιάζουμε με $\mathbf{Q}^{-1}$ και παίρνουμε:

$$\mathbf{Q}^{-1} \mathbf{Y} = \mathbf{Q}^{-1} \mathbf{X} \cdot \boldsymbol{\beta} + \mathbf{Q}^{-1} \mathbf{e} \Leftrightarrow \mathbf{Y}^* = \mathbf{X}^* \cdot \boldsymbol{\beta} + \mathbf{e}^*$$

όπου τώρα

$$V(\mathbf{e}^*) = V(\mathbf{Q}^{-1} \mathbf{e}) = \mathbf{Q}^{-1} \cdot V(\mathbf{e}) \cdot (\mathbf{Q}^{-1})^T = \sigma^2 \mathbf{Q}^{-1} \cdot \mathbf{Q} \cdot \mathbf{Q}^T \cdot (\mathbf{Q}^T)^{-1} = \sigma^2 \mathbf{I}$$

και επομένως έχουμε το κλασσικό γραμμικό μοντέλο

$$\mathbf{Y}^* = \mathbf{X}^* \cdot \boldsymbol{\beta} + \mathbf{e}^*, \quad \mathbf{e}^* \sim N(0, \sigma^2 \mathbf{I})$$

από το οποίο παίρνουμε την γνωστή εκτιμήτρια ελαχ. τετραγώνων:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^{*T} \mathbf{X}^*)^{-1} \mathbf{X}^{*T} \mathbf{Y}^* = ((\mathbf{Q}^{-1} \mathbf{X})^T \mathbf{Q}^{-1} \mathbf{X})^{-1} (\mathbf{Q}^{-1} \mathbf{X})^T \mathbf{Q}^{-1} \mathbf{Y} = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} \mathbf{Y}$$

με

$$V(\hat{\boldsymbol{\beta}}) = (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} \mathbf{X}^T \mathbf{V}^{-1} \cdot V(\mathbf{Y}) \cdot \mathbf{V}^{-1} \mathbf{X} (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1} = \sigma^2 (\mathbf{X}^T \mathbf{V}^{-1} \mathbf{X})^{-1}.$$

Όταν ο πίνακας $\mathbf{V}$ είναι διαγώνιος,

$$\mathbf{V} = \begin{bmatrix} v_1 & 0 & \dots & 0 \\ 0 & v_2 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \dots & 0 & v_n \end{bmatrix} = \text{diag}(v_1, v_2, \dots, v_n),$$

δηλαδή τα σφάλματα ε_i είναι ανεξάρτητα αλλά δεν έχουν ίδια διασπορά (είναι «σταθμισμένα», $V(\varepsilon_i) = \sigma^2 v_i$, με v_i γνωστά), τότε η εκτίμηση των β_i από το SPSS γίνεται χρησιμοποιώντας Linear Regression με Weights (WLS) τη στήλη με τα βάρη $1/v_1, 1/v_2, \dots, 1/v_n$.

¹ Έστω λ_i μία ιδιοτιμή του $\mathbf{V}$ και $\mathbf{w}_i$ ένα ιδιοδιάνυσμα που αντιστοιχεί στην λ_i . Εξ ορισμού θα ισχύει ότι $\mathbf{V} \mathbf{w}_i = \lambda_i \mathbf{w}_i$, $i = 1, 2, \dots, u$, και επομένως

$$\mathbf{V} \cdot [\mathbf{w}_1, \dots, \mathbf{w}_u] = [\mathbf{w}_1, \dots, \mathbf{w}_u] \cdot \text{diag}(\lambda_1, \dots, \lambda_u) \Rightarrow \mathbf{V} = \mathbf{W} \cdot \text{diag}(\lambda_1, \dots, \lambda_u) \cdot \mathbf{W}^{-1}.$$

Επειδή τα ιδιοδιανύσματα είναι κάθετα μεταξύ τους (ισχύει ότι $\mathbf{w}_i^T \mathbf{V} \mathbf{w}_j = \lambda_i \mathbf{w}_i^T \mathbf{w}_j = \lambda_j \mathbf{w}_j^T \mathbf{w}_i \Rightarrow \lambda_j \mathbf{w}_i^T \mathbf{w}_j = \lambda_i \mathbf{w}_i^T \mathbf{w}_j \Rightarrow \mathbf{w}_i^T \mathbf{w}_j = 0$ για $\lambda_i \neq \lambda_j$ - η καθετότητα αποδεικνύεται περαιτέρω και για ιδιοδιανύσματα που αντιστοιχούν στην ίδια ιδιοτιμή) και έχουν μέτρο 1 (αν π.χ. κάποιο $\mathbf{w}_i$: $\|\mathbf{w}_i\|^2 = \mathbf{w}_i^T \mathbf{w}_i \neq 1$ παίρνουμε στην θέση του το $\mathbf{w}_i / \|\mathbf{w}_i\|$) θα ισχύει ότι ο πίνακας $\mathbf{W}$ είναι ορθογώνιος, δηλαδή $\mathbf{W} \cdot \mathbf{W}^T = \mathbf{I}$ και επομένως $\mathbf{W}^{-1} = \mathbf{W}^T$. Άρα τελικά θα ισχύει ότι

$$\mathbf{V} = \mathbf{W} \cdot \text{diag}(\lambda_1, \dots, \lambda_u) \cdot \mathbf{W}^T = (\mathbf{W} \cdot \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_u})) \cdot (\mathbf{W} \cdot \text{diag}(\sqrt{\lambda_1}, \dots, \sqrt{\lambda_u}))^T = \mathbf{Q} \cdot \mathbf{Q}^T$$

($\lambda_i \geq 0$ λόγω του ότι ο $\mathbf{V}$ είναι ημιοριστικά θετικός).