ΜΕΜ-205 Περιγραφική Στατιστική

Τμήμα Μαθηματικών και Εφ. Μαθηματικών, Πανεπιστήμιο Κρήτης

Κώστας Σμαραγδάκης (kesmarag@pm.me)

22-03-2023

$\underset{\sim}{x} \to y$

$\underset{\sim}{x} = [x^{(1)}, x^{(2)}, x^{(3)}, \ldots, x^{(k)}]$

Πολλαπλή Γραμμική Παλινδρόμηση

$a^{T} b = \sum_{i=1}^{K} a_i b_i = b^{T} a$

$\sum_{i=1}^{K} x^{(i)} B^{(i)}$

$$\mathrm{y} = \underbrace{\mathrm{A} + \mathbf{x}^{\mathrm{T}}\mathbf{B}}_{\mu_{y|x}} + \epsilon$$

$$\mathbf{x} = \begin{bmatrix} \mathrm{x}^{(1)} \\ \mathrm{x}^{(2)} \\ \vdots \\ \mathrm{x}^{(\mathrm{K})} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} \mathrm{B}^{(1)} \\ \mathrm{B}^{(2)} \\ \vdots \\ \mathrm{B}^{(\mathrm{K})} \end{bmatrix}$$

Μοντέλο

~~Ευθεία~~ παλινδρόμησης για τον πληθυσμό

$$\mu_{\mathrm{y}|\mathbf{x}} = \mathrm{A} + \mathbf{x}^{\mathrm{T}}\mathbf{B}$$

$y = A + x^T B + \varepsilon$

Δειγματικό μοντέλο απλής γραμμικής παλινδρόμησης

$$\hat{y} = a + \mathbf{x}^{T}\mathbf{b}$$

- a είναι δειγματική προσέγγιση του A
- $\mathbf{b} = [b^{(1)}, b^{(2)}, \ldots, b^{(K)}]^{T}$ είναι δειγματική προσέγγιση του **B**
- $\hat{y}$ είναι η εκτιμώμενη τιμή του y για δοσμένο $\mathbf{x} = [x^{(1)}, x^{(2)}, \ldots, x^{(K)}]^{T}$

Τυχαίο σφάλμα του δειγματικού μοντέλου απλής γραμμικής παλινδρόμησης

$$e = y - \hat{y}$$

Πολλαπλή Γραμμική Παλινδρόμηση

$$\{(\overset{x^{(1)}}{1},\overset{x^{(2)}}{2},\underset{y}{3}),(2,3,5)\}$$

Έστω το τυχαίο δείγμα

$$\{(\underbrace{x_1^{(1)},\dots,x_1^{(K)}},y_1),(x_2^{(1)},\dots,x_2^{(K)},y_2),\dots,(x_N^{(1)},\dots,x_N^{(K)},y_N)\}$$

Για το τυχαίο σφάλμα του δειγματικού μοντέλου πολλαπλής γραμμικής παλινδρόμησης έχουμε:

$$e_n = y_n - \hat{y}_n, \quad n = 1,\dots,N$$

όπου η προσέγγιση του κάθε y_n δίνεται ως

$$\hat{y}_n = a + \mathbf{x}_n^T\mathbf{b} \rightarrow \hat{y}_n = a + [x_n^{(1)},\dots,x_n^{(K)}]\begin{bmatrix} b^{(1)} \\ b^{(2)} \\ \vdots \\ b^{(k)} \end{bmatrix} = \underbrace{[1 \;\; x_n^{(1)} \cdots x_n^{(K)}]}_{X_n}\underbrace{\begin{bmatrix} a \\ b^{(1)} \\ \vdots \\ b^{(k)} \end{bmatrix}}_{=\ \underset{\sim}{\beta}}$$

Άθροισμα τετραγωνικών σφαλμάτων

$$\mathsf{SSE} = \sum_{n=1}^{N} e_n^2 \xrightarrow{\min} a, \underset{\sim}{b}$$

$$\hat{\mathbf{y}} = \begin{bmatrix} \hat{y}_1 \\ \vdots \\ \hat{y}_N \end{bmatrix} = \mathbf{X}\mathbf{p}$$

$$\mathbf{p} = \begin{bmatrix} a \\ b^{(1)} \\ b^{(2)} \\ \vdots \\ b^{(K)} \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_1^{(1)} & x_1^{(2)} & \dots & x_1^{(K)} \\ 1 & x_2^{(1)} & x_2^{(2)} & \dots & x_2^{(K)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_N^{(1)} & x_N^{(2)} & \dots & x_N^{(K)} \end{bmatrix}, \quad \mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}$$

Προσέγγιση ελαχίστων τετραγώνων

$$\mathbf{e}^T = \mathbf{y} - \hat{\mathbf{y}} = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix} - \begin{bmatrix} X \end{bmatrix} \begin{bmatrix} a \\ b^{(1)} \\ \vdots \\ b^{(K)} \end{bmatrix}$$

$$\mathcal{Q}(\mathbf{p}) = \mathbf{e}^{\mathrm{T}}\mathbf{e} = \mathbf{y}^{\mathrm{T}}\mathbf{y} - 2\mathbf{p}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{y} + \mathbf{p}^{\mathrm{T}}\mathbf{X}^{\mathrm{T}}\mathbf{X}\mathbf{p} \qquad \mathbf{y} - \mathbf{X}\mathbf{p}$$

$$\mathbf{p} = \arg\min_{\mathbf{p}'} \mathcal{Q}(\mathbf{p}')$$

$$\mathbf{p} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}$$

$$Q(p) = \sum_{i=1}^{N} e_i^2 = \underset{\sim}{e}^T \underset{\sim}{e} = (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{p})^T (\underset{\sim}{y} - \underset{\sim}{X}\underset{\sim}{p}) = [y^T - p^T X^T](y - Xp) =$$

$$= y^T y - \underbrace{y^T X p} - p^T X^T y + p^T X^T X p$$

$$y^T X p = \underbrace{y^T (Xp)}_{y \cdot Xp} = (Xp)^T y = p^T X^T y$$

$$\frac{\partial}{\partial \alpha_1} \alpha^T b = \frac{\partial}{\partial \alpha_1} \left(\sum \alpha_k b_k \right) = \frac{\partial}{\partial \alpha_1} (\underbrace{\alpha_1 b_1} + \alpha_2 b_2 + \cdots) = b_1$$

$$= y^T y - 2 p^T X^T y + p^T X^T X p$$

$$\frac{\partial Q}{\partial p_j} = 0 \quad \forall j = 1, \ldots, K+1$$

$$\frac{\partial}{\partial p_j} \left(\underbrace{p^T} X^T y \right) = \frac{\partial}{\partial p_j} \left(\sum_{k=1}^{K+1} p_k (X^T y)_k \right) =$$

$$= (X^T y)_j \qquad X^T y$$

$$\frac{\partial (p^T X^T X p)}{\partial p_j} = \frac{\partial}{\partial p_j}\left[(Xp)^T Xp\right] = \frac{\partial}{\partial p_j} \sum_{k=1}^{K+1} (Xp)_k^2 = 2 \sum_{k=1}^{K+1} (Xp)_k \cdot$$

$$(Xp)_i = \sum_{l=1}^{K+1} X_{il} p_l$$

$$\cdot \sum_{l=1}^{K+1} X_{kl} \frac{\partial p_l}{\partial p_j} = 2 \sum_{k=1}^{K+1} X_{kj} (Xp)_k = 2 X^T Xp$$

$$\frac{\partial Q}{\partial p} = \begin{bmatrix} \frac{\partial Q}{\partial p_1} \\ \vdots \\ \frac{\partial Q}{\partial p_{K+1}} \end{bmatrix} = 0 = -\cancel{2} X^T y + \cancel{2} X^T Xp = 0 \Rightarrow X^T Xp = X^T y \Rightarrow$$

$$\Rightarrow p = (X^T X)^{-1} X^T y$$

Παράδειγμα

Να βρεθεί το δειγματικό μοντέλο γραμμικής παλινδρόμησης για το σύνολο δεδομένων

$$\{(1, -1, 1), (0, -1, -1), (2, 0, 2), (1, 1, 2)\}$$

$$\underset{\sim}{X} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & -1 \\ 1 & 2 & 0 \\ 1 & 1 & 1 \end{bmatrix} \qquad y = \begin{bmatrix} 1 \\ -1 \\ 2 \\ 2 \end{bmatrix} \qquad X^T X \in \mathbb{R}^{3\times 3}$$

$$\underset{\sim}{p} = \begin{bmatrix} \alpha \\ b^{(1)} \\ b^{(2)} \end{bmatrix} = (X^T X)^{-1} X^T y$$

Άσκηση

Δείξτε ότι η εκτίμηση ελαχίστων τετραγώνων

$$\mathbf{p} = (\mathbf{X}^{\mathrm{T}}\mathbf{X})^{-1}\mathbf{X}^{\mathrm{T}}\mathbf{y}$$

στη περίπτωση της απλής γραμμικής παλινδρόμησης οδηγεί, όπως περιμένουμε, στις εκτιμήσεις των παραμέτρων:

$$\mathrm{b} = \frac{\mathrm{SS_{xy}}}{\mathrm{SS_{xx}}}, \quad \mathrm{a} = \bar{\mathrm{Y}} - \mathrm{b}\bar{\mathrm{X}}$$