Παλινδρόμηση 11/11/16

Πολλαπλή γραμμική παλινδρόμηση

$Y = B_0 + B_1 X + \varepsilon$: μοντέλο

Έχω μια τυχαία μεταβλητή και θέλω να δω πως επηρεάζεται όχι από μία αλλά από p-μεταβλητές δηλ:

$$Y = B_0 + B_1X_1 + B_2X_2 + \dots + B_pX_p + \varepsilon$$

→ παράμετροι

→ σφάλματα

→ εξαρτημένη μετ. ή απόκριση

→ επεξηγηματικές ή ανεξάρτητες

Μορφή δεδομένων

Πειραματικές Μονάδες	Y	X_1	X_2	...	X_p
1	Y_1	x_{11}	x_{12}	...	x_{1p}
2	Y_2	x_{21}	x_{22}	...	x_{2p}
⋮	⋮	⋮	⋮		⋮
n	Y_n	x_{n1}	x_{n2}	...	x_{np}

Μια ισοδύναμη γραφή στα δεδομένα $\Longleftrightarrow$

$$Y_i = B_0 + B_1X_{i1} + B_2X_{i2} + \dots + B_pX_{ip} + \varepsilon_i, \quad i=1,\dots,n$$

$\Updownarrow$

$$\boxed{\underset{\sim}{Y} = X\underset{\sim}{B} + \underset{\sim}{\varepsilon}}$$

n×1 διαν

παράδ

αν ορίσω κατάλληλα του πίνακα X $\underset{\sim}{Y} = \begin{pmatrix} Y_1 \\ \vdots \\ Y_n \end{pmatrix}$ και $\underset{\sim}{X} = \begin{pmatrix} 1 & x_{11} & x_{12} & \dots & x_{1p} \\ \vdots & & & & \\ 1 & x_{n1} & & & x_{np} \end{pmatrix}$

και $\underset{\sim}{B} = \begin{pmatrix} B_0 \\ \vdots \\ B_p \end{pmatrix}$ $\underset{\sim}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$

Αρα το μοντέλο είναι

$$\boxed{\underset{n\times 1}{\underset{\sim}{Y}} = \underset{n\times(p+1)}{X}\,\underset{n\times 1}{\underset{\sim}{B}} + \underset{n\times 1}{\underset{\sim}{\varepsilon}}}$$

Y — Διανυσμα παρατηρήσεων της εξαρτημ. μετ.

$\underset{\sim}{B}$ ~ διάνυσμα παραμέτρων

$\underset{\sim}{\varepsilon}$ → διάνυσμα σφαλμάτων

X — πίνακας σχεδιασμού

Πρώτο Βήμα για μοντελοποίηση

- **Εκτιμητές Ελαχίστων Τετραγώνων των παραμέτρων $\underset{\sim}{B}$ του μοντέλου** (όλοια με απλή οραμ)

Ε.Ε.Τ στηρίζονται στην ελαχιστοποίηση $S(\underset{\sim}{B}) = \sum_{i=1}^{n} \varepsilon_i^2$

$$= \underset{\sim}{\varepsilon}'\underset{\sim}{\varepsilon}$$

$$= (\underset{\sim}{Y} - X\underset{\sim}{B})^T(\underset{\sim}{Y} - X\underset{\sim}{B})$$

$$= (\underset{\sim}{Y}' - X'\underset{\sim}{B}')(\underset{\sim}{Y} - X\underset{\sim}{B})$$

$$= \underset{\sim}{Y}'Y - \underset{\sim}{Y}'X\underset{\sim}{B} - B'X'\underset{\sim}{Y} + B'X'X\underset{\sim}{B}$$

Επειδη $\underset{\sim}{Y}'X\underset{\sim}{B}$
είναι 1x1 πίνα. δηλ αριθμ.
αρα Βαθμωτό μέγεθος
οπως και το $(Y'XB)' = \underset{\sim}{B}'X'\underset{\sim}{Y}$ ~~αριθμός~~ και άρα

$$S(\underset{\sim}{B}) = Y'Y - 2\underset{\sim}{B}'X\underset{\sim}{Y} + \underset{\sim}{B}'X'X\underset{\sim}{B}$$

και αρα οι ΕΕΤ θα προκύψουν απο ελαχ. του $S(\underset{\sim}{B})$

- $\dfrac{\partial(\underset{\sim}{B}'\underset{\sim}{a})}{\partial \underset{\sim}{B}} = \underset{\sim}{a}$ και
- $\dfrac{\partial(\underset{\sim}{B}'A\underset{\sim}{B})}{\partial \underset{\sim}{B}} = 2AB$ όπου $\dfrac{\partial}{\partial \underset{\sim}{B}} = \nabla \underset{\sim}{B}$
$= \left(\dfrac{\partial}{B_0}, \ldots, \dfrac{\partial}{B_p}\right)$

και άρα

$$\frac{\partial}{\partial \underset{\sim}{B}} S(\underset{\sim}{B}) = -2X'\underset{\sim}{Y} + 2X'X\underset{\sim}{B}$$ και μηδενιζω δια υποψήφιο ακρότατο

$$\frac{\partial}{\partial \underset{\sim}{B}} S(\underset{\sim}{B}) = 0 \Rightarrow X'X\underset{\sim}{B} = X'\underset{\sim}{Y} \;\underset{\text{και πολλ. θα εχω}}{\overset{\text{Αν } \exists (X'X)^{-1} \text{ τοτε}}{\Longleftrightarrow}}\; \underset{\sim}{B} = (X'X)^{-1}X'\underset{\sim}{Y}$$

(2)

Επομένως αν ο $(X'X)$ είναι αντιστρέψιμος τότε οι ΕΕΤ είναι

$$\boxed{\hat{\underset{\sim}{B}} = (X'X)^{-1} X' Y}$$

- ως παρατήρηση ο $(X'X)$ είναι αντιστρέψιμος τότε το μοντέλο λέγεται πλήρους βαθμίδας
- Αν ο $X'X$ δεν είναι αντιστρέψιμος τότε το μοντέλο λέγεται μη πλήρους βαθμίδας, οι ΕΕΤ δεν είναι μοναδικοί δίνονται όμως απο την σχέση $\hat{\underset{\sim}{B}} = (X'X)^{-} X' \underset{\sim}{Y}$ οπου $(X'X)^{-}$ είναι ένας γενικευμένος αντίστροφος του $X'X$

οπότε έχοντας τους ΕΕΤ μπορούμε να να εκτιμήσουμε το

▶ Εκτιμώμενο μοντέλο πολλαπλής γραμμικής παλινδρόμησης

$$\hat{\underset{\sim}{Y}} = X \hat{\underset{\sim}{B}}$$

Υπόλοιπα

$$\underset{\sim}{e} = \underset{\sim}{Y} - \hat{\underset{\sim}{Y}} = Y - X\hat{\underset{\sim}{B}}$$

▶ Ιδιότητες των ΕΕΤ $\hat{\underset{\sim}{B}}$

Υποθέσεις στα σφάλματα (απλή ορολογία)

1. $E(\varepsilon_i) = 0 \quad i = 1, \dots, n$
2. $Var(\varepsilon_i) = \sigma^2 \quad i = 1, \dots, n$
3. $Cov(\varepsilon_i, \varepsilon_j) = 0 \quad i \neq j = 1, \dots, n$
4. $\varepsilon_i \sim N(0, \sigma^2) \quad i = 1, \dots, n$

Εστω τ.δ $\underset{\sim}{w} = \begin{pmatrix} w_1 \\ w_n \end{pmatrix}$ τοτε $E(\underset{\sim}{w}) = \begin{pmatrix} E(w_1) \\ E(w_n) \end{pmatrix}$

οποτε αν $\underset{\sim}{\varepsilon} = \begin{pmatrix} \varepsilon_1 \\ \vdots \\ \varepsilon_n \end{pmatrix}$ τοτε $E(\underset{\sim}{\varepsilon}) = \begin{pmatrix} E\varepsilon_1 \\ \vdots \\ E(\varepsilon_n) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = \underset{\sim}{0}$

▶ <u>Υποθέσεις σφαλμάτων ε</u>

1. $E(\underset{\sim}{\varepsilon}) = \underset{\sim}{0}$
2. $Var(\underset{\sim}{\varepsilon}) = \sigma^2 I_n$

▶ <u>Συνέπειες για $\underset{\sim}{y}$</u> $(\underset{\sim}{y} = X\underset{\sim}{\beta} + \underset{\sim}{\varepsilon})$

1. $E(\underset{\sim}{y}) = X\underset{\sim}{\beta}$
2. $Var(y) = \sigma^2 I_n$

$$Var(\underset{\sim}{w}) \overset{op}{=} \Sigma = \begin{pmatrix} Var(w_1) & & Cov(w_i, w_j) \\ Cov(w_j, w_i) & Var(w_2) & \\ & & \ddots \\ & & & Var(w_n) \end{pmatrix}$$

ιδιότητες $Var(\underset{\sim}{w})$

1. ο $Var(\underset{\sim}{w})$ συμμετρικός
2. ο $Var(\underset{\sim}{w})$ θετικά ημιορισμένος (η διασπορά είναι πάντα θετ.)

<u>οπότε</u>

$$Var(\underset{\sim}{\varepsilon}) = \begin{pmatrix} Var(\varepsilon_1) & & Cov(\varepsilon_i, \varepsilon_j) \\ & \ddots & \\ & & Var(\varepsilon_n) \end{pmatrix}_{n \times n} = \begin{pmatrix} \sigma^2 & 0 & 0 \\ & \ddots & \\ & & \sigma^2 \end{pmatrix} = \sigma^2 \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ & \ddots & \\ & & 1 \end{pmatrix}$$

όμως $Cov(\varepsilon_i, \varepsilon_j) = 0$ και $Var(\varepsilon_i) = 0$

③

Οπότε έχοντας πλέον αυτές τις υποθέσεις μπορώ να επιλύσω τις ιδιότητες Ε.Τ.

ΙΔΙΟΤΗΤΑ 1. $E(\hat{\underset{\sim}{\beta}}) = \underset{\sim}{\beta}$ (αμερόληπτος)

Αποδ

$$E(\hat{\underset{\sim}{\beta}}) = E[(X^TX)^{-1}X^T\underset{\sim}{Y}]$$
$$= (X'X)^{-1}X^TE(\underline{Y})$$
$$= (X'X)^{-1}X^TX\underset{\sim}{\beta}$$
$$= I_{p+1}\underset{\sim}{\beta} = \underset{\sim}{\beta}$$

ΙΔΙΟΤΗΤΑ 2. $Var(\hat{\underset{\sim}{\beta}}) = \sigma^2(X'X)^{-1}$ (Θέλω να βρώ τον πίνακα ιδιοτιμών, διότι όσο πιο μικρή διασ. τόσο καλύτερος εκτιμητής) Αυτό μπορώ να το κρίνω είτε με det ή με το ίχνος ή με οποιαδ συναρ. ιδιοτιμών.)

Αποδ Αν $\underset{\sim}{w}$ είναι τ.δ και A ένας πίνακας κατάλληλης διάστασης ώστε να ισχύουν οι πράξεις τότε

$$Var(A\underset{\sim}{w}) = A\,Var(\underset{\sim}{w})A'$$

οπότε

$$Var(\hat{\underset{\sim}{\beta}}) = Var(\underbrace{(X^TX)^{-1}X^T}_{A}Y)$$
$$= (X^TX)^{-1}X^T(Var(\underset{\sim}{Y}))[(X^TX)^{-1}X^T]^T$$

[και επειδή $(AB)^T = B^TA^T$
$(A^{-1})^T = (A^T)^{-1}$]

$$= (X^TX)^{-1}X^T(\sigma^2I_n)X(X^TX)^{-1}$$
$$= \sigma^2\,\cancel{(X^TX)^{-1}}\,\cancel{X^TX}(X^TX)^{-1}$$
$$= \sigma^2(X^TX)^{-1}.$$

ΙΔΙΟΤΗΤΑ 3.

Διακύμανση της πρόβλεψης.

Έστω $\underset{\sim}{x_0}^T = (1, x_{01}, x_{02}, , x_{pp})$ ένα γνωστό διάνυσμα για δεδομένες τιμές

x_{0i} των ανεξάρτητων μεταβλ. x_i, $i=1, , p$

(Θέλω να δω πως συμπεριφέρονται το Y για συγκεκριμένες τιμές)

Η εκτιμώμενη τιμή του Y είναι $\hat{Y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_{01} \quad + \hat{\beta}_p x_{0p}$

το Θέμα είναι πόσο καλή είναι αυτή η πρόβλεψη.

Η πρόβλεψη, οπότε, είναι καλή αν η $Var(\hat{Y}_0)$ είναι μικρή!

Θ.α.ο η $Var(\hat{Y}_0) = \sigma^2 \underset{\sim}{x_0}^T (X^T X)^{-1} \underset{\sim}{x_0}$ Κάθε τετραγωνική μορφή είναι ένας αριθμός

$Var(\hat{Y}_0) = Var(x_0^T \underset{\sim}{\hat{\beta}})$

$\boxed{\underline{\text{Ισχύει}} \quad Var(a' \underset{\sim}{w}) = \underset{\sim}{a}^T Var(\underset{\sim}{w})\, a}$ $\left(= \sum_{i=1}^{n} a_i^2 Var(w_i) \sum_{i=1}^{n} \sum_{\substack{j=1 \\ i \neq j}}^{n} a_i a_j Cov(w_i, w_j)\right)$

οπότε

$Var(\hat{Y}_0) = \underset{\sim}{x_0}' Var(\underset{\sim}{\hat{\beta}})\, x_0 = \underset{\sim}{x_0}^T [\sigma^2 (X^T X)^{-1}] \underset{\sim}{x_0} = \sigma^2 \underset{\sim}{x_0}^T (X^T X)^{-1} \underset{\sim}{x_0}$

Άρα έτσι μπορώ να βγάλω κάποιο συμπέρασμα για συγκεκριμένη τιμή

▶ Πίνακας ΑΝΑΛΥΣΗ ΜΟΝΤΕΛΟΥ ΠΟΛΛ/ΗΣ ΓΡΑΜΜΙΚΗΣ ΠΑΛΙΝΔΡΟΜΗΣΗΣ

$SS_{tot} = \sum_{i=1}^{n} (Y_i - \bar{Y})^2 = \sum (Y_i^2 - 2\bar{Y}Y_i + \bar{Y}^2) = \sum (Y_i^2 - 2\bar{Y}\sum Y_i + n\bar{Y}^2)$

$= \sum Y_i^2 - 2\bar{Y} n \bar{Y} + n\bar{Y}^2$

άρα

$$\boxed{SS_{tot} = \underset{\sim}{Y}'\underset{\sim}{Y} - n\bar{Y}^2}$$

(α)

$$SS_{res} = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 = \sum_{i=1}^{n} e_i^2 = \sum_{i=1}^{n} e^T e = (y - \hat{y})^T (y - \hat{y})$$

$$= (y - X\hat{\beta})^T (y - X\hat{\beta}) = (y^T - \hat{\beta}^T X^T)(y - X\hat{\beta})$$

$$= y^T y - y^T X \hat{\beta} - \hat{\beta}^T X^T y + \hat{\beta}^T X^T X \hat{\beta}$$ (αποτελούν αριθ.)

$$= y^T y - 2\hat{\beta}^T X^T y + \hat{\beta}^T X^T X \hat{\beta}$$

- $$\hat{\beta}^T X^T X \hat{\beta} = \hat{\beta}^T X^T X (X^T X)^{-1} X^T y$$
$$= \hat{\beta}^T X^T y$$

Άρα

$$SS_{res} = y^T y - 2\hat{\beta}^T X^T y + \hat{\beta}^T X^T y \Rightarrow$$

$$\boxed{SS_{res} = y^T y - \hat{\beta}^T X^T y}$$

Υπολογισμος SSres

$$SS_{reg} = \sum_{i=1}^{n} (\hat{y}_i - \bar{y})^2 = \sum (\hat{y}_i^2 - 2\hat{y}_i \bar{y} + \bar{y}^2)$$

$$= \sum \hat{y}_i^2 - 2\bar{y} \sum \hat{y}_i + n\bar{y}^2$$

$$= \hat{y}^T \hat{y} + n\bar{y}^2 - 2\bar{y} \sum_{i=1}^{n} \hat{y}_i$$ επειδή $\hat{y} = X\hat{\beta}$

$$= \hat{\beta}^T X^T X \hat{\beta} + n\bar{y}^2 - 2\bar{y} \sum \hat{y}_i$$

$\hat{\beta}^T X^T X \hat{\beta} = \hat{\beta}^T X^T y$

$$= \hat{\beta}^T X^T y + n\bar{y}^2 - 2\bar{y} \underbrace{\sum \hat{y}_i}_{} $$ → απόδειξη στη επόμενη

$$= \hat{\beta}^T X^T y + n\bar{y}^2 - 2\bar{y} n \bar{y} = \hat{\beta}^T X^T y - n\bar{y}^2$$

Ανάλυση $\bar{y}$ ς $\hat{y}_i$

Το $\hat{\underset{\sim}{\beta}}$ ικανοποιεί τις κανονικές εξισώσεις $X^T X \underset{\sim}{\beta} = X^T \underset{\sim}{y}$

$$(X^T X)\hat{\underset{\sim}{\beta}} = X^T \underset{\sim}{y} \Rightarrow X^T \underset{\sim}{y} - X^T X \hat{\underset{\sim}{\beta}} = \underset{\sim}{0} \Rightarrow X^T(\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}}) = \underset{\sim}{0}$$

$$\Rightarrow [X^T(\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}})]^{-1} = \underset{\sim}{0} \Rightarrow (\underset{\sim}{y} - X\hat{\underset{\sim}{\beta}})^T X = \underset{\sim}{0} \Rightarrow$$

$$(\underset{\sim}{y} - \hat{\underset{\sim}{y}})^T X = \underset{\sim}{0} \Rightarrow$$

θα είναι δραστικά

$$(y_1 - \hat{y}_1, y_2 - \hat{y}_2, \dots, y_n - \hat{y}_n) \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ \vdots & & x_{np} \end{pmatrix} = \underset{\sim}{0}$$

$$\Rightarrow \sum_{i=1}^{n} (y_i - \hat{y}_i) = 0$$

$$\Rightarrow \sum y_i - \sum \hat{y}_i = 0 \Rightarrow \sum \hat{y}_i = n\bar{y}$$

Άρα

$$\boxed{SS_{reg} = \hat{B}^T X \underset{\sim}{y} - n\bar{y}^2}$$

Τελικά έχω

$$\boxed{SS_{tot} = SS_{reg} + SS_{res}}$$

ΠΙΝΑΚΑΣ ΑΝΑΛΙΑ

Πηγή Μεταβλ.	SS	Β.Ε	MS	F-πηλίκο
Μοντελ. παλινδ.	SS_{reg}	p	$MS_{reg} = \frac{SS_{reg}}{p}$	
Υπόλοιπο	SS_{res}	$n-p-1$	$MS_{res} = \frac{SS_{res}}{n-p-1}$	
Ολική	SS_{tot}	$n-1$		

- Εκτιμητής της σ^2 : $E(MS_{res}) = \sigma^2$

- Συντελεστής προσδιορισμού ή προσαρμοστικότητας

$$R^2 = \frac{SS_{reg}}{SS_{tot}} = 1 - \frac{SS_{res}}{SS_{tot}}$$

καθ. αρ

$0 \leq R^2 \leq 1$

αν είναι κοντά στο 1 τότε έχω υποσχόμενο μοντέλο

$$R^2\text{-adj} = 1 - \frac{MS_{res}}{MS_{tot}}$$