Δευτέρα 09/11/20

Πολλαπλή γραμμική παλινδρόμηση

$\boxed{Y = b_0 + b_1 X_1 + b_2 X_2 + \dots + b_p X_p + \varepsilon.}$ το μοντέλο της πολλαπλής γραμμικής παλινδρόμησης

Y: εξαρτημένη ή απόκριση
X_i: ανεξάρτητες ή επεξηγηματικές $i = 1, 2, \dots p$
ε: σφάλματα (πιθανοί παράγοντες επηρεασμού της Y οι οποίοι δεν λαμβάνονται στο μοντέλο
b_i: παράμετροι μοντέλου.

Μορφή δεδομένων.

Πειραμ. Μονάδες	Y	X_1	X_2 ...	X_p
1	Y_1	$X_{1,1}$	$X_{1,2}$	$X_{1,p}$
2	Y_2	$X_{2,1}$	$X_{2,2}$	$X_{2,p}$
⋮	⋮	⋮	⋮	⋮
n	Y_n	$X_{n,1}$	$X_{n,2}$	$X_{n,p}$

Έτσι το μοντέλο της π.γ.π γράφεται ισοδύναμα

$$\boxed{Y_i = b_0 + b_1 X_{i,1} + b_2 X_{i,2} + \dots + b_p X_{i,p} + \varepsilon_i, \quad i = 1, \dots n.}$$

Αν θεωρήσουμε
$\underset{\sim}{Y} = (Y_1, Y_2, \dots Y_n)^T$: διάνυσμα παρατηρήσεων της εξαρτημένης μεταβλητής
$\underset{\sim}{\varepsilon} = (\varepsilon_1, \varepsilon_2, \dots \varepsilon_n)^T$: διάνυσμα σφαλμάτων
$\underset{\sim}{b} = (b_0, b_1, \dots b_p)^T$: διάνυσμα παραμέτρων

$$X = \begin{pmatrix} 1 & X_{11} & X_{12} & X_{1p} \\ 1 & X_{21} & X_{22} & X_{2p} \\ \vdots & \vdots & \vdots & \\ 1 & X_{n1} & X_{n2} & X_{np} \end{pmatrix}$$: πίνακας σχεδιασμού

Τότε $\boxed{\underset{\sim}{Y} = X \underset{\sim}{b} + \underset{\sim}{\varepsilon}.}$

Εκτιμητές των παραμέτρων $\underset{\sim}{\beta}$

ΕΕΤ: $S(\underset{\sim}{\beta}) = \sum_{i=1}^{n} \varepsilon_i^2 = \varepsilon^T \cdot \varepsilon = (\underset{\sim}{Y} - X\underset{\sim}{\beta})^T (Y - X\underset{\sim}{\beta}) =$

ανάστροφος.

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

$\underline{\underset{\sim}{Y}^T X \underset{\sim}{\beta} = \underset{\sim}{\beta}^T X^T \underset{\sim}{Y}}$ $\quad \underset{\sim}{Y}^T \underset{\sim}{Y} - 2\underset{\sim}{\beta}^T X^T \underset{\sim}{Y} + \underset{\sim}{\beta}^T X^T X \underset{\sim}{\beta}$

Επιδιώκουμε να ελαχιστοποιήσουμε την S.

$\frac{\partial S(\underset{\sim}{\beta})}{\partial \underset{\sim}{\beta}} = 0 \Rightarrow -2 \frac{\partial}{\partial \underset{\sim}{\beta}}(\underset{\sim}{\beta}^T X^T \underset{\sim}{Y}) + \frac{\partial}{\partial \underset{\sim}{\beta}}(\underset{\sim}{\beta}^T X^T X \underset{\sim}{\beta}) = 0 \quad (1)$

<u>Ισχύουν:</u> Αν a, β διανύσματα και A πίνακας τότε

$\frac{\partial}{\partial \underset{\sim}{\beta}}(\underset{\sim}{\beta}^T \underset{\sim}{a}) = \underset{\sim}{a} \qquad \frac{\partial}{\partial \underset{\sim}{\beta}}(\underset{\sim}{\beta}^T A \underset{\sim}{\beta}) = 2A\underset{\sim}{\beta}$

Οπότε από (1) έχουμε.

$-2X^T \underset{\sim}{Y} + 2X^T X \underset{\sim}{\beta} = 0 \Rightarrow X^T X \underset{\sim}{\beta} = X^T Y.$ (Κανονικές εξισώσεις)

Ο ΕΕΤ $\hat{\underset{\sim}{\beta}}$ θα προκύψει από την επίλυση ως προς $\underset{\sim}{\beta}$ των κανονικών εξισώσεων, αν ο $X^T \cdot X$ είναι αντιστρέψιμος τότε

$\underset{\sim}{\beta} = (X^T X)^{-1} X^T Y$ ~~και επομένως οι ΕΕΤ είναι~~

υπολογίζω ακρότατο και ο Εσσιανός είναι θετικά ορισμ. ⇒ ελάχιστο

Άρα $\hat{\underset{\sim}{\beta}} = (X^T X)^{-1} X^T Y.$ ΕΕΤ.

Παρατήρηση:

Αν ο $X^T X$ είναι αντιστρέψιμος δηλ $\exists (X^T X)^{-1}$ τότε το μοντέλο λέγεται πλήρους βαθμίδας.

Αλλιώς αν $\nexists (X^T X)^{-1}$ τότε λέγεται μη πλήρους βαθμίδας, και οι ΕΕΤ δεν είναι μοναδικοί. Δίνονται όμως από τη σχέση

$\hat{\underset{\sim}{\beta}} = (X^T X)^{-} (X^T \underset{\sim}{Y})$ όπου $(X^T X)^{-}$ είναι ένας γενικευμένος αντίστροφος του $(X^T X)^{-1}$

Υποθέσεις για τα σφάλματα:

① $E(\varepsilon_i)=0$, $i=1,2,\dots n$ $(\Leftrightarrow E(\underset{\sim}{\varepsilon})=\underset{\sim}{0})$

② $Var(\varepsilon_i)=\sigma^2$, $i=1,2,\dots n$ $(\Leftrightarrow Var(\underset{\sim}{\varepsilon})=\sigma^2 I_n)$.

③ $Cov(\varepsilon_i,\varepsilon_j)=0$, $i,j=1,2,\dots n$, $i\neq j$

④ $\varepsilon_i \sim N(0,\sigma^2)$, $i=1,2,\dots n$.

Συνέπειες στα Y_i

$\underset{\sim}{Y}=X\underset{\sim}{\beta}+\underset{\sim}{\varepsilon}$

① $E(Y_i)=\beta_0+\beta_1 X_{i1}+\dots+\beta_p X_{ip}$ $(\Leftrightarrow E(\underset{\sim}{Y})=E(X\underset{\sim}{\beta}+\underset{\sim}{\varepsilon})=X\underset{\sim}{\beta}$

② $Var(Y_i)=\sigma^2$, $i=1,\dots n$ $(\Leftrightarrow Var(\underset{\sim}{Y})=\sigma^2 I_n)$

③ Y_i ασυσχέτιστα, $Cov(Y_i,Y_j)=0$, $i,j=1,\dots n$, $i\neq j$

④ $Y_i \sim N(\beta_0+\beta_1 X_{i1}+\dots+\beta_p X_{ip},\sigma^2)$ $i=1,\dots n$.

Από ΕΠΣ θυμίζω

Έστω τ.δ $\underset{\sim}{w}=(w_1,\dots,w_n)^T$

Η μέση τιμή τ.δ $\underset{\sim}{w}$ ορίζεται $E(\underset{\sim}{w})=(Ew_1,\dots Ew_n)^T$.

Άρα $E(\varepsilon_i)=0$ $\forall i=1,\dots n \Rightarrow E(\underset{\sim}{\varepsilon})=\begin{pmatrix}E(\varepsilon_1)\\ \vdots \\ E(\varepsilon_n)\end{pmatrix}=\begin{pmatrix}0\\ \vdots \\ 0\\ \vdots\\0\end{pmatrix}=\underset{\sim}{0}$.

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

~~Ορισμός~~

Ο πίνακας αυτός είναι συμμετρικός αφού $Cov(w_i,w_j)=E(w_iw_j)-Ew_iEw_j$ και θετικά ορισμένος (όπως στην μονοδιάστατη περίπτωση η διακύμανση είναι μη αρνητική ποσότητα

$$Var(\underset{\sim}{\varepsilon})=\begin{pmatrix}Var\,\varepsilon_1 & & Cov(\varepsilon_i,\varepsilon_j)\\ & Var(\varepsilon_2) & \\ Cov(\varepsilon_j,\varepsilon_i) & & \\ & & Var(\varepsilon_n)\end{pmatrix}=\begin{pmatrix}\sigma^2 & 0 & \cdots & \vdots\\ 0 & \sigma^2 & 0 & \vdots\\ \vdots & 0 & \ddots & 0\\ \cdots & \cdots & & \sigma^2\end{pmatrix}$$

$=\sigma^2 I_n$.

Παρατήρηση

Έστω τ.δ $\underset{\sim}{W}=(W_1,\dots W_n)^T$

Το $\underset{\sim}{W}$ έχει πολυδιάστατη κανονική κατανομή με παραμέτρους $\mu=E(\underset{\sim}{W})$ και πίνακα $\Sigma=Var(\underset{\sim}{W})$ αν η σ.π.π.

$$f_{\underset{\sim}{W}}(w)=\frac{1}{(2\pi)^{-1/2}|\Sigma|^{-1/2}}\,e^{-\frac{1}{2}(\underset{\sim}{w}-\underset{\sim}{\mu})'\cdot\Sigma^{-1}(\underset{\sim}{w}-\underset{\sim}{\mu})}$$

Ισχύει: έστω $\underset{\sim}{W}=(W_1,\dots W_n)$ και $W_i\sim$ Normal και ασυσχέτιστες

$\Rightarrow \underset{\sim}{W}\sim N_n(\underset{\sim}{\mu}=E(\underset{\sim}{W}),\ \Sigma=Var(\underset{\sim}{W}))$

Επειδή από την (4) κάθε $\varepsilon_i\sim$ Normal και από (3) ε_i ασυσχέτιστα συνεπάγεται ότι $\underset{\sim}{\varepsilon}=N_n(\mu=E(\varepsilon),\Sigma=Var(\underset{\sim}{\varepsilon}))$

$\equiv N(0,\sigma^2 I_n)$

Συνεπώς $\boxed{\underset{\sim}{Y}\sim N_n(X\underset{\sim}{\beta},\sigma^2 I_n)}$

Υποθέσεις για τα σφάλματα

1) $E(\underset{\sim}{\varepsilon})=\underset{\sim}{0}$ $\longrightarrow$ $E(\underset{\sim}{Y})=X\underset{\sim}{\beta}$

2) $Var(\underset{\sim}{\varepsilon})=\sigma^2 I_n$ $\longrightarrow$ $Var(\underset{\sim}{Y})=\sigma^2 I_n$

3) $\underset{\sim}{\varepsilon}\sim N_n(0,\sigma^2 I_n)$ $\longrightarrow$ $\underset{\sim}{Y}\sim N_n(X\underset{\sim}{\beta},\sigma^2 I_n)$

Υπό τις υποθέσεις 1 και 2 για τα σφάλματα του μοντέλου:

$\boxed{E(\hat{\underset{\sim}{\beta}})=\underset{\sim}{\beta}}$ Αμεροληψία

Απόδ

$E(\hat{\underset{\sim}{\beta}})=E((X^TX)^{-1}X^T\underset{\sim}{Y})=(X^TX)^{-1}X^TE(\underset{\sim}{Y})=(X^TX)^{-1}X^TX\underset{\sim}{\beta}=$
$=I_{p+1}\underset{\sim}{\beta}=\underset{\sim}{\beta}.$

$\boxed{Var(\hat{\underset{\sim}{\beta}})=\sigma^2(X^TX)^{-1}}$

Ισχύει $Var(A\underset{\sim}{W})=A\,Var(\underset{\sim}{W})A^T$

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

$$Var(\hat{\beta}) = Var(\underbrace{(X^T X)^{-1} X^T}_{A} \cdot Y) = (X^T X)^{-1} X^T Var(Y) ((X^T X)^{-1} X^T)^T$$

$$= (X^T X)^{-1} X^T \cdot \sigma^2 I_n X (X^T X)^{-1} =$$

$$= \sigma^2 (X^T X)^{-1} X^T X (X^T X)^{-1} = \sigma^2 (X^T X)^{-1}$$

Διακύμανση της Πρόβλεψης:

Έστω $x_{01}, x_{02}, \dots x_{0p}$ τιμές των ανεξάρτητων $X_1, X_2, \dots X_p$

Τότε η προβλεπόμενη τιμή Y είναι $\hat{Y}_0 = \hat{\beta}_0 + \hat{\beta}_1 x_{01} + \dots + \hat{\beta}_p x_{0p}$.

Τότε $\boxed{Var(\hat{Y}_0) = \sigma^2 \, x_0^T (X^T X)^{-1} x_0.}$ όπου $x = (1, x_{01}, \dots x_{0p})^T$.

Απόδ.

$$Var(\hat{Y}_0) = Var(\hat{\beta}_0 + \hat{\beta}_1 x_{01} + \dots + \hat{\beta}_p x_{0p}) = Var(\hat{\beta}^T \cdot x_0) =$$

$$= Var(x_0^T \cdot \hat{\beta})$$

$$Var(\hat{Y}_0) = Var(x_0^T \cdot \beta) = x_0^T Var(\beta) x_0 = x_0^T (\sigma^2 I (X^T X)^{-1}) x_0$$

$$= \sigma^2 x_0^T (X^T X)^{-1} x_0$$

$$\Rightarrow Var(\hat{Y}_0) = \sigma^2 x_0^T (X^T X)^{-1} x_0 \quad \text{αριθμός}$$

SStot, SSreg, SSres (κατ' αναλογία με α.γ.π)

Εκτιμώμενο μοντέλο: $\hat{Y} = X \hat{\beta}$

Υπόλοιπα: $e = Y - \hat{Y}$

Υπολογισμός των SS

$$SStot = \sum_{i=1}^{n} (Y_i - \bar{Y})^2 = Y^T Y - n \bar{Y}^2$$

$$SSres = \sum_{i=1}^{n} e_i^2 = e^T e = (Y - \hat{Y})^T (Y - \hat{Y}) =$$

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

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

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

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

άρα $SS_{res.} = \underset{\sim}{y}^T \underset{\sim}{y} - \hat{\underset{\sim}{\beta}}^T X^T \underset{\sim}{y}$