За підтримки Назарія Саворони та Нікіти Сазонова

Пропуски під "самостійні доведення"

Знайдемо оцінку $\hat{\alpha}_w$: (ЗМНК)

$$\|e\|_w^2 = e^T W e = (y - X\alpha)^T W (y - X\alpha) =$$

$$= \alpha^T X^T W X \alpha - 2\alpha^T X^T W y + \|y\|_w^2.$$

• Необхідна умова екстремуму:

$$\{\operatorname{grad}_\alpha \|e\|_w^2\}\big|_{\alpha=\hat{\alpha}_w} = \{2(X^T W X)\alpha - 2X^T W y\}\big|_{\alpha=\hat{\alpha}_w}$$

$$= \Theta_p.$$

• Отримаємо так звану систему нормальних рівнянь для оцінки ЗМНК:

$$(X^T W X)\hat{\alpha}_w = X^T W y.$$

$$\hat{\alpha}_w = (X^T W X)^{-1} \cdot X^T W y$$

• Знайдемо похибку оцінювання $\Delta(\hat{\alpha}_w)$ для оцінки $\hat{\alpha}_w$:

$$\underline{\Delta(\hat{\alpha}_w)} = \hat{\alpha}_w - \alpha = (X^T W X)^{-1} \cdot X^T W y - \alpha =$$

$$= (X^T W X)^{-1} \cdot X^T W X \alpha + (X^T W X)^{-1} \cdot X^T W e - \alpha = \underline{(X^T W X)^{-1} X^T W e}$$

__Наслідок__ Оцінка МНК $\hat{\alpha}$ та її похибка $\Delta(\hat{\alpha})$ набувають вигляду ($W = E$):

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

$$\Delta(\hat{\alpha}) = \hat{\alpha} - \alpha = (X^T X)^{-1} X^T e,$$

система норм. рівнянь для оцінки МНК: $(X^T X)\hat{\alpha} = X^T y.$

• Значення залежної змінної оцінюватимемо:

$\hat{y} = X\hat{\alpha}$ або $\hat{y}(k) = x^T(k)\hat{\alpha}$.

$$\hat{\sigma}^2 = \frac{1}{N-p}\|y - X\hat{\alpha}\|^2$$

__Властивості оцінок $\hat{\alpha}, \hat{y}, \hat{\sigma}^2$:__

I. а) $\hat{\alpha} \sim N(\alpha, \sigma^2 (X^T X)^{-1})$. i-тий діагональний елемент

б) $\hat{\alpha}_i \sim N(\alpha_i, \sigma^2 d_i)$

розмірність вектора α

II. статистика $$\frac{(\hat{\alpha} - \alpha)^T (X^T X)(\hat{\alpha} - \alpha)}{\sigma^2} \sim \chi^2(p)$$

III. а) $\hat{y} \sim N(X\alpha, \sigma^2 X (X^T X)^{-1} X^T)$

б) $\hat{y}(k) \sim N(x^T(k)\alpha, \sigma^2 x^T(k)(X^T X)^{-1} x(k)), \ k = \overline{1,N}$.

IV. а) $\frac{(N-p)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(N-p)$

б) $M\hat{\sigma}^2 = \sigma^2$, $D\sigma^2 = \frac{2\sigma^4}{N-p}$

V. Оцінки $\hat{\sigma}^2$ та $\hat{\alpha}$ є незалежними

VI Оцінки $\hat{\alpha}$, $\hat{g}$, $\hat{\sigma}^2$ є ефективними

VII Теорема Андерсона - Тейлора:

Вкажемо явно у регресійній моделі поточну кількість спостережень N:

$y_N = X_N\alpha + e_N$,

Позначимо через $\hat{\alpha}(N) = (X_N^T X_N)^{-1} X_N^T y_N$ оцінку МНК для α в цій моделі.

- Оцінка $\hat{\alpha}(N)$ є сильно слушною

$$\Updownarrow$$

$$(X_N^T X_N)^{-1} \xrightarrow[N\to\infty]{} \Theta_p$$ - нульова матриця порядку p.

$A>0$ - матриця A - додатньо визначена

Озн. Х $q\times q$ матриці A, B є симетричними.

Тоді вважаємо, що $A \geq B \Leftrightarrow A-B \geq 0$.

Заув.:

1. Якщо I.а $\Rightarrow$ $M\hat{\alpha} = \alpha$,

$$M\left((\hat{\alpha}-\alpha)(\hat{\alpha}-\alpha)^T\right) = \sigma^2 (X^T X)^{-1}.$$

2. VI $\Rightarrow$ $$M(\hat{\sigma}^2-\sigma^2)^2 = \min_{\tilde{\sigma}^2 \in \mathcal{U}_{\sigma^2}} M(\tilde{\sigma}^2-\sigma^2)^2.$$

$$M\left((\hat{\alpha}-\alpha)(\hat{\alpha}-\alpha)^T\right) = \min_{\tilde{\alpha} \in \mathcal{U}_{\alpha}} M\left((\tilde{\alpha}-\alpha)(\tilde{\alpha}-\alpha)^T\right).$$

$$M\left((\hat{y}-X\alpha)(\hat{y}-X\alpha)^T\right) = \min_{\tilde{y} \in \mathcal{U}_{y}} M\left((\tilde{y}-X\alpha)(\tilde{y}-X\alpha)^T\right),$$

де $\mathcal{U}_{\sigma^2}$, $\mathcal{U}_{\alpha}$, $\mathcal{U}_{y}$ – класи незміщених оцінок.

3. Якщо $\vec{\xi} \sim N(\vec{m}, V)$, тоді $\vec{a} + B\vec{\xi} \sim N(\vec{a} + B\vec{m}, BVB^T)$.

Озн. Матриця P наз. проекційною, якщо вона є симетричною та ідемпотентною, тобто $P = P^T$ та $P^2 = P$.

Позн. $\hat{\alpha}(N) = (X_N^T X_N)^{-1} X_N^T y_N$ оц. МНК для α у цій мод. Тоді оц. $\hat{\alpha}(N)$ є сильно слушною $\Leftrightarrow$

$$\Leftrightarrow \quad (X_N^T X_N)^{-1} \xrightarrow[N\to\infty]{} \Theta_p,$$

де Θ_p – нуль-матр. пор. p.

Нех $q \times q$ матр A, B симетр. Тоді $A \geq B \Leftrightarrow$
$\Leftrightarrow A - B \geq 0$.

!1 З Іа $\Rightarrow$ $\hat{\alpha}$ – незм $\Rightarrow$ $M\hat{\alpha} = \alpha$, а її к-ка розс. $M\{(\hat{\alpha}-\alpha)(\hat{\alpha}-\alpha)^T\} = \sigma^2 (X^T X)^{-1}$

!2 З ІІ $\Rightarrow$ $M(\hat{\sigma}^2 - \sigma^2)^2 = \min\limits_{\tilde{\sigma}^2 \in \mathfrak{A}_{\sigma^2}} M(\tilde{\sigma}^2 - \sigma^2)^2$,

$$M\{(\hat{\alpha}-\alpha)(\hat{\alpha}-\alpha)^T\} = \min_{\tilde{\alpha}\in\mathfrak{A}_\alpha} M\{(\tilde{\alpha}-\alpha)(\tilde{\alpha}-\alpha)^T\},$$

$$M\{(\hat{y} - X\alpha)(\hat{y}-X\alpha)^T\} = \min_{\tilde{y}\in\mathfrak{A}_y} M\{(\tilde{y} - X\alpha)(\tilde{y} - X\alpha)^T\},$$

де $\mathfrak{A}_x$ – клас незм. оц. для x.

!3 Нех $\vec{\xi} \sim \mathcal{N}(\vec{m}, V)$, тоді

$$\vec{a} + B\vec{\xi} \sim \mathcal{N}(\vec{a} + B\vec{m}, BVB^T),$$

де $\vec{a}$, B – вект та матр. відпов. розм.

Д. Іа З І прип. клас. регр. ан.:

$e \sim \mathcal{N}(\Theta_N, \sigma^2 E_N)$, $\sigma^2 \in \mathbb{R}_+$ та вигляду регр. моделі $y = X\alpha + e$ випливає:

$y \sim \mathcal{N}(X\alpha, \sigma^2 E_N)$.

А оцінка $\hat{\alpha}$ є лін. перетв. вект. y. Тому з

$$\begin{cases} \hat{\alpha} = (X^T X)^{-1} X^T y \\ y \sim N(X\alpha, \sigma^2 E_N) \end{cases} \Bigg| \Rightarrow$$

$$\Rightarrow \hat{\alpha} \sim N\left(\cancel{(X^T X)^{-1} X^T X} \alpha, (X^T X)^{-1} X^T \sigma^2 E_N X (X^T X)^{-1}\right) \Rightarrow$$

$$\Rightarrow \hat{\alpha} \sim N(\alpha, \sigma^2 (X^T X)^{-1})$$

I б Нех $l_i = (0 \ldots 0 \overset{i}{1} 0 \ldots 1)^T$ є i-ою ортою. Тоді т.я. $\hat{\alpha}_i = l_i^T \hat{\alpha} \Rightarrow$

$$\Rightarrow \hat{\alpha}_i \sim N(l_i^T \alpha, \sigma^2 l_i^T (X^T X)^{-1} l_i) \Rightarrow$$

$$\Rightarrow \hat{\alpha}_i \sim N(\alpha_i, \sigma^2 d_i)$$

Лема 1. Нех. $\vec{\xi} \sim N(\vec{m}, V)$, $V > 0$, $\vec{\xi} \in \mathbb{R}^q$. Тоді

$$(\vec{\xi} - \vec{m})^T V^{-1} (\vec{\xi} - \vec{m}) \sim \chi^2(q)$$

II 3 I a $\Rightarrow$ $\hat{\alpha} \sim N(\alpha, \sigma^2(X^T X)^{-1})$, відомо, що $\alpha \in \mathbb{R}^p$. Застос. Л1 до $\hat{\alpha}$ отрим:

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

$$= \frac{(\hat{\alpha}-\alpha)^T (X^T X)(\hat{\alpha}-\alpha)}{\sigma^2} \sim \chi^2(p)$$

III a $\hat{y} = X\hat{\alpha}$

III б $\hat{y}(k) = l_k^T \hat{y}$, $k = \overline{1, N}$

iii $\hat{b}$ $\hat{y}(k) = l_k^T \hat{y}$, $k = \overline{1, N}$

Матр. P наз проекційною $\Leftrightarrow P = P^T$, $P^2 = P$.

Лема 2 Нех.

- $\vec{\xi} \sim \mathcal{N}(\theta_N, E_N)$, $\vec{\xi} \in \mathbb{R}^N$
- $P = P^T$, $P \in M_N(\mathbb{R})$
- $\operatorname{rank}(P) = r$

Тоді: $\vec{\xi}^T P \vec{\xi} \sim \chi^2(r) \Leftrightarrow P^2 = P$

Лема 3: P - проекц $\Rightarrow \operatorname{rank}(P) = tr(P)$

що

IV a) Перш ніж встанов, що $\frac{(N-p)\hat{\sigma}^2}{\sigma^2} \sim \chi^2(N-p)$, отрим для $\hat{\sigma}^2$ представл. через e та y:

$$\hat{\sigma}^2 = \frac{1}{N-p}\|y - X\hat{\alpha}\|^2 = \frac{1}{N-p}\|y - X(X^T X)^{-1}X^T y\|^2 =$$

$$= \frac{1}{N-p}\|[E_N - X(X^T X)^{-1}X^T]y\|^2 = \frac{1}{N-p}\|Py\|^2 =$$

$$= \left[\begin{array}{l} \frac{1}{N-p} y^T P^T P y = \frac{1}{N-p} y^T P^2 y = \frac{1}{N-p} y^T P y \\ \frac{1}{N-p}\|PX\alpha + Pe\|^2 = \frac{1}{N-p}\|Pe\|^2 = \frac{1}{N-p} e^T P e, \end{array}\right.$$

де $P = E_N - X(X^T X)^{-1}X^T$

Отрим

$$\hat{\sigma}^2 = \left[\begin{array}{ll} \frac{1}{N-p} y^T P y & (*) \\ \frac{1}{N-p} e^T P e & (**) \end{array}\right.$$

Скорист. тим, що $PX = \Theta_{N,p}$, після транспон. отрим таке $X^T P = \Theta_{p,N}$

$P^2 = P$;

Підрах:

$tr(P) = tr(E_N - X(X^TX)^{-1}X^T) =$

$= tr(E_N) - tr(X(X^TX)^{-1}X^T) =$

$= N - tr(E_p) = N - p$

Скорист тим, що:

$A \in M_{p,q}(\mathbb{R}),\ B \in M_{q,p}(\mathbb{R})$, то

$tr(AB) = tr(BA)$

Т.я. P - проєкц, то за 13 та (0.14):

$rank(P) = tr(P) = N - p$

Т.я. $\frac{e}{\sigma} \sim \mathcal{N}(\theta_N, E_N)$, то скорист. Л2, отрим

$$\frac{(N-p)\hat{\sigma}^2}{\sigma^2} \overset{(**)}{=} \frac{(N-p)\frac{1}{N-p}e^TPe}{\sigma^2} = \frac{e^TPe}{\sigma^2} =$$

$$= \left(\frac{e}{\sigma}\right)^T P \left(\frac{e}{\sigma}\right) \overset{Л2}{\sim} \chi^2(N-p)$$

IV б. Т.я. $M\chi^2(q) = q$, $D\chi^2(q) = 2q$, то з IV а випливає:

$$\frac{(N-p)\hat{\sigma}^2}{\sigma^2} = \chi^2(N-p) \Rightarrow \left[\begin{array}{l} \frac{N-p}{\sigma^2} M\hat{\sigma}^2 = M\left\{\frac{(N-p)\hat{\sigma}^2}{\sigma^2}\right\} = N-p \\ \frac{(N-p)^2}{\sigma^4} D\hat{\sigma}^2 = D\left\{\frac{(N-p)\hat{\sigma}^2}{\sigma^2}\right\} = 2(N-p) \end{array}\right.$$

$$\Rightarrow \left\{\begin{array}{l} \frac{\cancel{(N-p)}}{\sigma^2} M\hat{\sigma}^2 = \cancel{N-p}^{\,1} \\ \frac{(N-p)^{\cancel{2}}}{\sigma^4} D\hat{\sigma}^2 = 2\cancel{(N-p)} \end{array}\right. \Rightarrow \left\{\begin{array}{l} M\hat{\sigma}^2 = \sigma^2 \\ D\hat{\sigma}^2 = \frac{2\sigma^4}{N-p} \end{array}\right.$$