Матричний підхід до лінійної багатофакторної моделі

План

1. Лінійна багатофакторна економетрична модель і основні її припущення в матричній формі.
2. МНК в матричній формі.
3. Дисперсійно-коваріаційна матриця .
4. Прогнозування за економетричною моделлю.

1.Лінійна багатофакторна економетрична модель і основні її припущення в матричній формі

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} \ldots + \beta_m x_{mi} + u_i, \qquad (1)$$

де y_i – i-е значення залежної змінної; x_{ji} – i-е значення j-ої незалежної змінної; $\beta_0,\ \beta_1, \ldots,\ \beta_m$ – невідомі детерміновані параметри; u_i – i-е значення випадкової змінної.

$$
\begin{cases}
y_1 = \beta_0 + \beta_1 x_{11} + \beta_2 x_{21} \ldots + \beta_m x_{m1} + u_1; \\
y_2 = \beta_0 + \beta_1 x_{12} + \beta_2 x_{22} \ldots + \beta_m x_{m2} + u_2; \\
\vdots \\
y_n = \beta_0 + \beta_1 x_{1n} + \beta_2 x_{2n} \ldots + \beta_m x_{mn} + u_n.
\end{cases}
\tag{2}
$$

$$
\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{U}, \tag{3}
$$

$$
\boldsymbol{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}; \boldsymbol{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{m1} \\ 1 & x_{12} & x_{22} & \cdots & x_{m2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{mn} \end{bmatrix}; \boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_m \end{bmatrix}; \boldsymbol{U} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix}.
$$

Припущення 1.

$$M\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} M(u_1) \\ M(u_2) \\ \vdots \\ M(u_n) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Припущення 2.

$$M(\mathbf{UU}') = M\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \cdot \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix},$$

$$M(\mathbf{UU}') = \begin{bmatrix} \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix} =$$

$$= \sigma^2 \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = \sigma^2 E$$

Дисперсійно–коваріаційна матриця випадкових величин u_i. По діагоналі її стоять дисперсії, поза діагоналлю – коваріації.

Припущення 4. Вектор випадкових величин u_i має нормальний закон розподілу

Припущення 5. Матриця $\boldsymbol{X}$ утворюється з фіксованих елементів, тобто не випадкова.

Припущення 6. Між факторами x_i, x_j відсутня мультиколінеарність, тобто фактори незалежні між собою. Це означає, що стовпці матриці $\boldsymbol{X}$ лінійно незалежні, тобто не знайдеться таких чисел γ_0, γ_1, …, γ_m, серед яких не всі дорівнюють нулю, щоб виразити один стовпець як лінійну комбінацію інших:

$$\gamma_0 \cdot 1 + \gamma_1 x_{1i} + \ldots + \gamma_m x_{mi} = 0 \text{ або } \boldsymbol{\gamma}' \cdot \boldsymbol{X} = \boldsymbol{0}.$$

2. МНК в матричній формі

$$\hat{\mathbf{Y}} = \begin{bmatrix} \hat{y}_1 \\ \hat{y}_2 \\ \vdots \\ \hat{y}_n \end{bmatrix}; \ \hat{\boldsymbol{\beta}} = \begin{bmatrix} \hat{\beta}_0 \\ \hat{\beta}_1 \\ \vdots \\ \hat{\beta}_m \end{bmatrix}; \ \mathbf{e} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}; \ \mathbf{Y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix};$$

$$\mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{21} & \cdots & x_{m1} \\ 1 & x_{12} & x_{22} & \cdots & x_{m2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{mn} \end{bmatrix}.$$

$$\mathbf{Y} = \mathbf{X}\hat{\boldsymbol{\beta}} + \mathbf{e}$$

$$e(\hat{\beta}) = \sum_{i=1}^{n} e_i^2 = \mathbf{e}'\mathbf{e} = (\boldsymbol{Y} - \boldsymbol{X}\hat{\beta})'(\boldsymbol{Y} - \boldsymbol{X}\hat{\beta}) =$$

$$= \boldsymbol{Y}'\boldsymbol{Y} - 2\hat{\beta}'\boldsymbol{X}'\boldsymbol{Y} + \hat{\beta}'\boldsymbol{X}'\boldsymbol{X}\hat{\beta} \to \min$$

$$[(\boldsymbol{X}\hat{\beta})' = \hat{\beta}'\boldsymbol{X}'; \;\; \hat{\beta}'\boldsymbol{X}'\boldsymbol{Y} = \boldsymbol{Y}'\boldsymbol{X}\hat{\beta}].$$

$$\frac{\partial e(\hat{\beta})}{\partial \hat{\beta}} = -2\boldsymbol{X}'\boldsymbol{Y} + 2\boldsymbol{X}'\boldsymbol{X}\hat{\beta} = 0.$$

$$\boldsymbol{X}'\boldsymbol{X}\hat{\beta} = \boldsymbol{X}'\boldsymbol{Y}$$

$$\hat{\beta} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{Y}$$

(Символ штрих (′) означає операцію транспонування)

3. Дисперсійно-коваріаційна матриця

$$M[(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})\cdot(\hat{\boldsymbol{\beta}}-\boldsymbol{\beta})']=$$

$$=\begin{bmatrix} M(\hat{\beta}_0-\beta_0)^2 & M(\hat{\beta}_0-\beta_0)(\hat{\beta}_1-\beta_1) & \cdots & M(\hat{\beta}_0-\beta_0)(\hat{\beta}_m-\beta_m) \\ M(\hat{\beta}_1-\beta_1)(\hat{\beta}_0-\beta_0) & M(\hat{\beta}_1-\beta_1)^2 & \cdots & M(\hat{\beta}_1-\beta_1)(\hat{\beta}_m-\beta_m) \\ \vdots & \vdots & \ddots & \vdots \\ M(\hat{\beta}_m-\beta_m)(\hat{\beta}_0-\beta_0) & M(\hat{\beta}_m-\beta_m)(\hat{\beta}_1-\beta_1) & \cdots & M(\hat{\beta}_m-\beta_m)^2 \end{bmatrix}$$

$$= \begin{bmatrix} \sigma^2_{\hat\beta_0} & \sigma_{\hat\beta_0\hat\beta_1} & \cdots & \sigma_{\hat\beta_0\hat\beta_m} \\ \sigma_{\hat\beta_1\hat\beta_0} & \sigma^2_{\hat\beta_1} & \cdots & \sigma_{\hat\beta_1\hat\beta_m} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{\hat\beta_m\hat\beta_0} & \sigma_{\hat\beta_m\hat\beta_1} & \cdots & \sigma^2_{\hat\beta_m} \end{bmatrix}.$$

$$\mathrm{var}(\hat{\boldsymbol\beta}) = \hat\sigma_u^2 (\boldsymbol{X'X})^{-1}$$

де $\hat\sigma_u^2$ – оцінена дисперсія випадкової величини:

$$\hat\sigma_u^2 = \frac{\sum_{i=1}^{n} e_i^2}{n-m-1}$$

$$\hat\sigma_u^2 = \frac{\boldsymbol{Y'Y} - \hat\beta \boldsymbol{X'Y}}{n-m-1} = \frac{\mathbf{e'e}}{n-m-1}.$$

4. Прогнозування за економетричною моделлю

$$H_0 : \beta_0 = \beta_1 = \dots = \beta_m = 0.$$

H_1 : не всі β_j $(j = \overline{0, m})$ дорівнюють нулю.

F–відношення з m та $n - m - 1$ ступенями вільності:

$$F=\frac{\dfrac{\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2}{m}}{\dfrac{\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}{n-m-1}}=\frac{(n-m-1)\sum_{i=1}^{n}(\hat{y}_i-\bar{y})^2}{m\sum_{i=1}^{n}(y_i-\hat{y}_i)^2},$$

де m – кількість факторів, що ввійшли в модель, n – кількість спостережень у вибірці.

$$F=\frac{(n-m-1)R^2}{m(1-R^2)}$$

$$F_{кр}(m;n-m-1;\gamma)$$

Якщо $F > F_{кр}$, то ми відкидаємо H_0 з ризиком помилитися не більше ніж в $\gamma\%$ випадків, і приймаємо, що побудоване рівняння економетричної моделі адекватне реальній дійсності.

Якщо $F \leq F_{кр}$ – H_0 приймаємо і вважаємо, що побудована модель неадекватна. Тоді необхідно, можливо, будувати нелінійну модель або ввести додаткові фактори.

Точковий прогноз:

$$\hat{y}_{n+k} = \hat{\beta}_0 + \hat{\beta}_1 x_{1,n+k} + \ldots + \hat{\beta}_m x_{m,n+k}$$

або в матричній формі:

$$\hat{\boldsymbol{Y}}_{n+k} = \boldsymbol{X}_{n+k}\hat{\beta},$$

де $\boldsymbol{X}_{n+k}$ – матриця очікуваних пояснювальних змінних.

Інтервал, у який з певною заданою імовірністю $p = 1 - \gamma$ потрапляє дійсне значення залежної змінної:

$$\hat{\boldsymbol{Y}}_{n+k} \pm t_{\gamma}\hat{\sigma}_u \sqrt{1 + \boldsymbol{X}'_{n+k}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}_{n+k}},$$

де $\boldsymbol{X}'_{n+k}$ вектор значень з m факторів у період $n+k$.

$$\hat{\boldsymbol{Y}}_{n+k} \pm t_{\gamma}\hat{\sigma}_u \sqrt{\boldsymbol{X}'_{n+k}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}_{n+k}}.$$