

# Эконометрика

Полковников Александр Александрович

Волжский политехнический институт (филиал)  
ФГБОУ ВПО "Волгоградский государственный технический университет"

Конспект лекций для студентов направления  
"Экономика"

## Нормальная линейная модель

Будем предполагать следующее:

- 1 модель наблюдений результата  $y$  с  $(m + 1)$  фактором  $x_0, \dots, x_m$  имеет вид

$$y_i = \theta_0 x_{i0} + \dots + \theta_m x_{im} + e_i, \quad k = 1, \dots, n,$$

где

- $y_i$  — значение результата в  $i$ -м наблюдении,
- $x_{ij}$  — известное фиксированное значение  $j$ -го фактора в  $i$ -м наблюдении,
- $\theta_j$  — неизвестный коэффициент при  $j$ -м факторе,
- $e_i$  — случайная ошибка в  $i$ -м наблюдении.

В матричном виде модель имеет вид

$$y = X\theta + e.$$

- ② Случайные величины  $e_1, \dots, e_n$  независимы в совокупности, имеют нормальное распределение  $\mathcal{N}(0; \sigma^2)$  с нулевым математическим ожиданием и дисперсией  $\sigma^2 > 0$ . Или иначе вектор  $(e_1, \dots, e_n)$  имеет  $n$ -мерное нормальное распределение с математическим ожиданием равным нулевому вектору  $(0, \dots, 0)^T$  и диагональной ковариационной матрицей  $\text{Cov}(e) = \sigma^2 I_n$ , где  $I_n$  — единичная матрица размера  $n \times n$ .

В дальнейшем на предположения этого пункта будем ссылаться как на **стандартные предположения об ошибках** в линейной модели наблюдений.

- ③ Если не оговорено противное, то в число объясняющих переменных включается переменная, тождественно равная единице

$$x_{i0} = 1, \quad i = 1, \dots, n.$$

- ④ Определитель матрицы  $X^T X$  отличен от нуля:

$$\det(X^T X) \neq 0,$$

что можно заменить условием: столбцы матрицы  $X$  линейно независимы.

При сделанных стандартных предположениях об ошибках модели величины  $y_1, \dots, y_n$  являются независимыми в совокупности нормально распределенными случайными величинами

$$y_i \sim \mathcal{N}(\theta_0 x_{i0} + \dots + \theta_m x_{im}; \sigma^2),$$

для которых математические ожидания и дисперсии равны соответственно:

$$\mathbb{E}(y_i) = \theta_0 x_{i0} + \dots + \theta_m x_{im}, \quad \mathbb{D}(y_i) = \sigma^2.$$

Или в матричном виде

$$\mathbb{E}(y) = X\theta, \quad \text{Cov}(y) = \sigma^2 I_n.$$

Обозначим матрицу  $(X^T X)^{-1} X^T$  из представления случайного вектора  $\hat{\theta}$  через  $C$ :

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

Тогда величина  $\hat{\theta}$  является линейным преобразованием нормально распределенного случайного вектора  $y$  и, следовательно, имеет нормальное распределение.

Математическое ожидание этого случайного вектора равно:

$$\mathbb{E}\hat{\theta} = \mathbb{E}\left((X^T X)^{-1} X^T y\right) = (X^T X)^{-1} X^T \mathbb{E}y = (X^T X)^{-1} X^T X\theta = \theta.$$

Ковариационная матрица вектора  $\hat{\theta}$  равна

$$\begin{aligned}\text{Cov}(\hat{\theta}) &= \text{Cov}(Cy) = C \text{Cov}(y) C^T = \\ &= (X^T X)^{-1} X^T \sigma^2 I_n \left( (X^T X)^{-1} X^T \right)^T = \\ &= \sigma^2 (X^T X)^{-1} X^T X (X^T X)^{-1} = \\ &= \sigma^2 (X^T X)^{-1}.\end{aligned}$$

Отсюда, в частности, получаем выражение для дисперсии

$$\mathbb{D}(\hat{\theta}_j) = \sigma^2 (X^T X)^{-1}_{jj},$$

где  $(X^T X)^{-1}_{jj}$  —  $j$ -й диагональный элемент матрицы  $(X^T X)^{-1}$ .  
Для линейных моделей справедлива следующая важная теорема.



## Теорема (Гаусса–Маркова)

*Пусть модель наблюдений имеет вид:*

$$y = X\theta + e, \quad \mathbb{E}(e) = 0, \quad \text{Cov}(e) = \sigma^2 I_n,$$

*где матрица  $X$  имеет линейно независимые столбцы, т. е.*

$$\det(X^T X) \neq 0.$$

*Тогда оценка наименьших квадратов  $\hat{\theta} = (X^T X)^{-1} X^T y$  неизвестного вектора коэффициентов  $\theta$  является **наилучшей линейной несмещенной оценкой**, т. е. является **эффективной**.*

Если к условия теоремы добавить предположение о нормальности случайных ошибок  $e$ , то оценка  $\hat{\theta}$  является наилучшей среди всех несмещенных оценок, а не только в классе линейных.

Рассматривая линейную модель

$$y = X\theta + e$$

с  $e \sim \mathcal{N}(0; \sigma^2)$  установили, что оценка наименьших квадратов  $\hat{\theta}_j$  коэффициента  $\theta_j$  имеет нормальное распределение с параметрами:

$$\mathbb{E}\hat{\theta}_j = \theta_j, \quad \mathbb{D}\hat{\theta}_j = \sigma^2 \left( X^T X \right)^{-1}_{jj}.$$

Поэтому случайная величина

$$\frac{\hat{\theta}_j - \theta_j}{\sqrt{\sigma^2 \left( X^T X \right)^{-1}_{jj}}}$$

имеет стандартное нормальное распределение.

Воспользоваться последней статистикой для построения доверительного интервала невозможно, т. к. дисперсия  $\sigma^2$  не известна. Заменим дисперсию  $\sigma^2$  ее оценкой

$$RMS = \frac{RSS}{n - m - 1}.$$

При выполнении стандартных предположений о модели величина

$$\frac{RSS}{\sigma^2}$$

имеет распределение  $\chi^2$  с  $(n - m - 1)$  степенями свободы. Поэтому  $RMS$  несмещенная оценка параметра  $\sigma$ , а случайная величина

$$\frac{\hat{\theta}_j - \theta_j}{\sqrt{\frac{RSS}{n - m - 1} (X^T X)^{-1}_{jj}}}$$

имеет распределение Стьюдента с  $(n - m - 1)$  степенями свободы.

С доверительной вероятностью  $\beta = 1 - \alpha$  интервальной оценкой коэффициента  $\theta_j$  является

$$\begin{cases} \theta_j \geq \hat{\theta}_j - t_{1-\alpha/2}[n-m-1] \sqrt{\frac{RSS}{n-m-1} (X^T X)_{jj}^{-1}}, \\ \theta_j \leq \hat{\theta}_j + t_{1-\alpha/2}[n-m-1] \sqrt{\frac{RSS}{n-m-1} (X^T X)_{jj}^{-1}} \end{cases}$$

или

$$\begin{cases} \theta_j \geq \hat{\theta}_j - t_{(1+\beta)/2}[n-m-1] \sqrt{\frac{RSS}{n-m-1} (X^T X)_{jj}^{-1}}, \\ \theta_j \leq \hat{\theta}_j + t_{(1+\beta)/2}[n-m-1] \sqrt{\frac{RSS}{n-m-1} (X^T X)_{jj}^{-1}}. \end{cases}$$

В качестве качества множественной регрессии может выступать **множественный индекс детерминации  $R^2$**  и **скорректированный индекс детерминации  $R_{adj}^2$**  (adjusted  $R^2$ ):

$$R^2 = 1 - \frac{RSS}{TSS}, \quad R_{adj}^2 = 1 - \frac{RSS/(n - m - 1)}{TSS/(n - 1)},$$

$$R_{adj}^2 = 1 - (1 - R^2) \frac{n - 1}{n - m - 1}.$$

## Теорема

*Для линейной модели множественный индекс детерминации  $R^2$  выражается через скорректированные коэффициенты регрессии и парные коэффициенты корреляции:*

$$R^2 = \sum_{k=1}^m \beta_k r_{yx_k}.$$

Уравнение регрессии в стандартизированной виде имеет вид:

$$\hat{y}^* = \sum_{k=1}^m \beta_k x_k^*.$$

Запишем индекс детерминации в стандартизированном виде:

$$\begin{aligned} R^2 &= \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} = \frac{\sum_{i=1}^n (\hat{y}_i^* - \bar{y}^*)^2}{\sum_{i=1}^n (y_i^* - \bar{y}^*)^2} = \\ &= \frac{\sum_{i=1}^n (\hat{y}_i^* - 0)^2}{\sum_{i=1}^n (y_i^* - 0)^2} = \frac{\sum_{i=1}^n (\hat{y}_i^*)^2}{n} = \\ &= \frac{1}{n} \sum_{i=1}^n \left( \hat{y}_i^* \cdot \sum_{k=1}^m \beta_k x_{ik}^* \right) = \sum_{k=1}^m \beta_k \left( \frac{1}{n} \sum_{i=1}^n \hat{y}_i^* x_{ik}^* \right) = \sum_{k=1}^m \beta_k r_{\hat{y}x_k}. \end{aligned}$$

$$\begin{aligned} R^2 &= \sum_{k=1}^m \beta_k r_{\hat{y}x_k} = \sum_{k=1}^m \beta_k \langle \hat{y}; x_k \rangle = \sum_{k=1}^m \beta_k \langle y - e; x_k \rangle = \\ &= \sum_{k=1}^m \beta_k (\langle y; x_k \rangle - \langle e; x_k \rangle) = \sum_{k=1}^m \beta_k (\langle y; x_k \rangle - 0) = \sum_{k=1}^m \beta_k r_{yx_k}. \end{aligned}$$

Теорема доказана.



### Теорема

*Для линейной модели множественный индекс детерминации  $R^2$  выражается через парные коэффициенты корреляции:*

$$R^2 = 1 - \frac{\det(r_{yxx})}{\det(r_{xx})},$$

*где  $r_{xx}$  — корреляционная матрица факторов  $x_1, \dots, x_m$ ,  $r_{yxx}$  — корреляционная матрица величин  $y, x_1, \dots, x_m$ .*

Ранее показали, что скорректированные коэффициенты  $\beta_k$  являются решениями системы

$$\begin{cases} \beta_1 + \beta_2 r_{x_1 x_2} + \dots + \beta_m r_{x_1 x_m} = r_{y x_1}, \\ \beta_1 r_{x_1 x_2} + \beta_2 + \dots + \beta_m r_{x_2 x_m} = r_{y x_2}, \\ \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \ddots \quad \quad \quad \vdots \\ \beta_1 r_{x_1 x_m} + \beta_2 r_{x_2 x_m} + \dots + \beta_m = r_{y x_m}. \end{cases}$$

Решая систему методом Крамера, получаем:

$$\beta_k = \frac{\Delta_k}{\det(r_{xx})},$$

где определитель  $\Delta_k$  получается из определителя  $\det(r_{xx})$  заменой  $k$ -го столбца на столбец  $(r_{y x_1}, \dots, r_{y x_m})^T$ .

Преобразуем величину из условия:

$$1 - \frac{\det(r_{yxx})}{\det(r_{xx})} = \frac{\det(r_{xx}) - \det(r_{yxx})}{\det(r_{xx})} =$$

(разлагая  $\det(r_{yxx})$  по первой строке)

$$\begin{aligned} &= \frac{1}{\det(r_{xx})} (\det(r_{xx}) - (\det(r_{xx}) - r_{yx_1} \Delta_1 - \dots - r_{yx_m} \Delta_m)) = \\ &= \sum_{k=1}^m r_{yx_k} \frac{\Delta_k}{\det(r_{xx})} = \sum_{k=1}^m r_{yx_k} \beta_k = R^2. \end{aligned}$$

Теорема доказана.

В частности для двух нетривиальных факторов  $x_1, x_2$  имеем

$$R^2 = 1 - \frac{\begin{vmatrix} 1 & r_{yx_1} & r_{yx_2} \\ r_{yx_1} & 1 & r_{x_1x_2} \\ r_{yx_2} & r_{x_1x_2} & 1 \end{vmatrix}}{\begin{vmatrix} 1 & r_{x_1x_2} \\ r_{x_1x_2} & 1 \end{vmatrix}} =$$
$$= \frac{r_{yx_1}^2 + r_{yx_2}^2 - 2r_{yx_1}r_{yx_2}r_{x_1x_2}}{1 - r_{x_1x_2}^2}.$$

СПАСИБО ЗА ВНИМАНИЕ!