Эконометрика

ОБОБЩЕННЫЙ МЕТОД НАИМЕНЬШИХ КВАДРАТОВ

Обобщенный метод наименьших квадратов

Пусть функция регрессии представлена функцией, линейным образом зависящей от параметров - нелинейной по объясняющим переменным, но линейной по параметрам или, нелинейной по параметрам, но внутренне линейной (в этом случае необходимо произвести преобразование переменных):

$$y = a_0 + a_1\phi_1(x_1, x_2, \ldots, x_m) + a_2\phi_2(x_1, x_2, \ldots, x_m) + \ldots + a_n\phi_n(x_1, x_2, \ldots, x_m).$$

Модель данных в этом случае будет

$$y_i = a_0 + a_1\phi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) + a_2\phi_2(x_{1i}, x_{2i}, \ldots, x_{mi}) + \ldots + a_n\phi_n(x_{1i}, x_{2i}, \ldots, x_{mi}) + \varepsilon_i.$$

Предпосылки обобщенного метода наименьших квадратов

1. $(x_{1i}, x_{2i}, \ldots, x_{mi})$–детерминированные (нестохастические) переменные ;
2. Каждое измерение случайной погрешности характеризуется нулевым средним, не зависящим от значений наблюдаемых переменных;
3. Теоретическая дисперсия случайной составляющей постоянна для всех наблюдений, а их величины независимы от значений наблюдаемых переменных (**гомоскедастичность**);
4. Отсутствует **автокорреляция ошибок,** то есть отсутствие систематической связи между значениями случайной составляющей в любых двух наблюдениях;
5. Случайные погрешности имеют нормальное распределение.

Принцип наименьших квадратов

Утверждает, что выбор параметров функции регрессии является оптимальным в случае, когда сумма квадратов отклонений эмпирических значений результирующей переменной от теоретических значений этой переменной, рассчитанной по функции регрессии, является минимальной.

В этом случае он записывается,

$$Q(a_0, a_1, ..., a_n) = \sum_{i=1}^{N} \left(y_i - \hat{y}_i(a_0, a_1, ..., a_n; x_{1i}, x_{2i}, ..., x_{mi}) \right)^2 =$$

$$= \sum_{i=1}^{N} \left(y_i - (a_0 + a_1\varphi_1(x_{1i}, x_{2i}, ..., x_{mi}) + ... + a_n\varphi_n(x_{1i}, x_{2i}, ..., x_{mi})) \right)^2 \to \min .$$

В этом случае *N*–число наблюдений.

По принципу Лагранжа

$$Q(a_0, a_1, \ldots, a_n) = \sum_{i=1}^{N}\left(y_i - \hat{y}_i(a_0, a_1, \ldots, a_n; x_{1i}, x_{2i}, \ldots, x_{mi})\right)^2 =$$

$$= \sum_{i=1}^{N}\left(y_i - \left(a_0 + a_1\varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) + \ldots + a_n\varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi})\right)\right)^2 \to \min.$$

$$\Downarrow$$

$$\begin{cases} \dfrac{\partial Q(a_0, a_1, \ldots, a_n)}{\partial a_0} = 0 \\ \dfrac{\partial Q(a_0, a_1, \ldots, a_n)}{\partial a_1} = 0 \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \dfrac{\partial Q(a_0, a_1, \ldots, a_n)}{\partial a_n} = 0 \end{cases}$$

Проведем преобразования

$$
\begin{cases}
\dfrac{\partial Q(a_0, a_1, \ldots, a_n)}{\partial a_0} = \displaystyle\sum_{i=1}^{N} \dfrac{\partial (y_i - (a_0 + a_1\varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) + \ldots a_n\varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi})))^2}{\partial a_0} \\
\dfrac{\partial Q(a_0, a_1, \ldots, a_n)}{\partial a_1} = \displaystyle\sum_{i=1}^{N} \dfrac{\partial (y_i - (a_0 + a_1\varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) + \ldots a_n\varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi})))^2}{\partial a_1} \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\dfrac{\partial Q(a_0, a_1, \ldots, a_n)}{\partial a_n} = \displaystyle\sum_{i=1}^{N} \dfrac{\partial (y_i - (a_0 + a_1\varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) + \ldots a_n\varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi})))^2}{\partial a_n}
\end{cases}
$$

Проведем преобразования

$$\begin{cases} 2\sum_{i=1}^{N}(y_i-(a_0+a_1\varphi_1(x_{1i},x_{2i},\ldots,x_{mi})+\ldots a_n\varphi_n(x_{1i},x_{2i},\ldots,x_{mi})))\cdot(-1)=0 \\ 2\sum_{i=1}^{N}(y_i-(a_0+a_1\varphi_1(x_{1i},x_{2i},\ldots,x_{mi})+\ldots a_n\varphi_n(x_{1i},x_{2i},\ldots,x_{mi})))\cdot(-\varphi_1(x_{1i},x_{2i},\ldots,x_{mi}))=0 \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\ 2\sum_{i=1}^{N}(y_i-(a_0+a_1\varphi_1(x_{1i},x_{2i},\ldots,x_{mi})+\ldots a_n\varphi_n(x_{1i},x_{2i},\ldots,x_{mi})))\cdot(-\varphi_n(x_{1i},x_{2i},\ldots,x_{mi}))=0 \end{cases}$$

Умножим каждое уравнение на ½.

Проведем преобразование

$$\begin{cases} \sum_{i=1}^{N}(y_i-(a_0+a_1\varphi_1(x_{1i},x_{2i},\dots,x_{mi})+\dots a_n\varphi_n(x_{1i},x_{2i},\dots,x_{mi})))\cdot(-1)=0 \\ \sum_{i=1}^{N}(y_i-(a_0+a_1\varphi_1(x_{1i},x_{2i},\dots,x_{mi})+\dots a_n\varphi_n(x_{1i},x_{2i},\dots,x_{mi})))\cdot(-\varphi_1(x_{1i},x_{2i},\dots,x_{mi}))=0 \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ \sum_{i=1}^{N}(y_i-(a_0+a_1\varphi_1(x_{1i},x_{2i},\dots,x_{mi})+\dots a_n\varphi_n(x_{1i},x_{2i},\dots,x_{mi})))\cdot(-\varphi_n(x_{1i},x_{2i},\dots,x_{mi}))=0 \end{cases}$$

Разобьем суммы на слагаемые

$$
\begin{cases}
-\sum_{i=1}^{N} y_i + \sum_{i=1}^{N} a_0 + \sum_{i=1}^{N} a_1\varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) + \ldots + \sum_{i=1}^{N} a_n\varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi}) = 0 \\
-\sum_{i=1}^{N} y_i \cdot \varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) + \sum_{i=1}^{N} a_0 \cdot \varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) + \sum_{i=1}^{N} a_1\varphi_1{}^2(x_{1i}, x_{2i}, \ldots, x_{mi}) + \\
\ldots + \sum_{i=1}^{N} a_n\varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi}) \cdot \varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) = 0 \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
-\sum_{i=1}^{N} y_i \cdot \varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi}) + \sum_{i=1}^{N} a_0 \cdot \varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi}) + \\
\sum_{i=1}^{N} a_1\varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) \cdot \varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi}) + \ldots + \sum_{i=1}^{N} a_n\varphi_n{}^2(x_{1i}, x_{2i}, \ldots, x_{mi}) = 0
\end{cases}
$$

Перенесем независимые переменные в право

$$
\begin{cases}
\sum_{i=1}^{N} a_0 + \sum_{i=1}^{N} a_1\varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) + \ldots + \sum_{i=1}^{N} a_n\varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi}) = \sum_{i=1}^{N} y_i \\
\sum_{i=1}^{N} a_0 \cdot \varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) + \sum_{i=1}^{N} a_1\varphi_1^{\,2}(x_{1i}, x_{2i}, \ldots, x_{mi}) + \\
\ldots + \sum_{i=1}^{N} a_n\varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi}) \cdot \varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) = \sum_{i=1}^{N} y_i \cdot \varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) \\
\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \\
\sum_{i=1}^{N} a_0 \cdot \varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi}) + \sum_{i=1}^{N} a_1\varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) \cdot \varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi}) + \\
\ldots + \sum_{i=1}^{N} a_n\varphi_n^{\,2}(x_{1i}, x_{2i}, \ldots, x_{mi}) = \sum_{i=1}^{N} y_i \cdot \varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi})
\end{cases}
$$

Учитывая, что

$$\sum_{i=1}^{N} a_0 = a_0 \sum_{i=1}^{N} 1 = a_0 (1 + \ldots + 1) = N a_0$$

$$\begin{cases}
N a_0 + a_1 \sum\limits_{i=1}^{N} \varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) + \ldots + a_n \sum\limits_{i=1}^{N} \varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi}) = \sum\limits_{i=1}^{N} y_i \\
a_0 \sum\limits_{i=1}^{N} \varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) + a_1 \sum\limits_{i=1}^{N} \varphi_1{}^2(x_{1i}, x_{2i}, \ldots, x_{mi}) + \\
\ldots + a_n \sum\limits_{i=1}^{N} \varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi}) \cdot \varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) = \sum\limits_{i=1}^{N} y_i \cdot \varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\
a_0 \sum\limits_{i=1}^{N} \varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi}) + a_1 \sum\limits_{i=1}^{N} \varphi_1(x_{1i}, x_{2i}, \ldots, x_{mi}) \cdot \varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi}) + \\
\ldots + a_n \sum\limits_{i=1}^{N} \varphi_n{}^2(x_{1i}, x_{2i}, \ldots, x_{mi}) = \sum\limits_{i=1}^{N} y_i \cdot \varphi_n(x_{1i}, x_{2i}, \ldots, x_{mi})
\end{cases}$$

Запишем систему линейных уравнений в матричной форме

$$\begin{pmatrix} N & \sum_{i=1}^{N}\varphi_1(x_{1i},..,x_{mi}) & \dots & \sum_{i=1}^{N}\varphi_n(x_{1i},..,x_{mi}) \\ \sum_{i=1}^{N}\varphi_1(x_{1i},..,x_{mi}) & \sum_{i=1}^{N}\varphi_1^{\ 2}(x_{1i},..,x_{mi}) & \dots & \sum_{i=1}^{N}\varphi_1(x_{1i},..,x_{mi})\cdot\varphi_n(x_{1i},..,x_{mi}) \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^{N}\varphi_n(x_{1i},..,x_{mi}) & \sum_{i=1}^{N}\varphi_1(x_{1i},..,x_{mi})\cdot\varphi_n(x_{1i},..,x_{mi}) & \dots & \sum_{i=1}^{N}\varphi_n^{\ 2}(x_{1i},..,x_{mi}) \end{pmatrix} \cdot \begin{pmatrix} a_0 \\ a_1 \\ \dots \\ a_n \end{pmatrix} =$$

$$= \begin{pmatrix} \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} y_i\varphi_1(x_{1i},..,x_{mi}) \\ \dots \\ \sum_{i=1}^{N} y_i\varphi_n(x_{1i},..,x_{mi}) \end{pmatrix}$$

Решим систему методом Крамера

Для этого необходимо, чтобы

$$D = \begin{vmatrix} N & \sum\limits_{i=1}^{N} \varphi_1(x_{1i},..,x_{mi}) & \ldots & \sum\limits_{i=1}^{N} \varphi_n(x_{1i},..,x_{mi}) \\ \sum\limits_{i=1}^{N} \varphi_1(x_{1i},..,x_{mi}) & \sum\limits_{i=1}^{N} \varphi_1{}^2(x_{1i},..,x_{mi}) & \ldots & \sum\limits_{i=1}^{N} \varphi_1(x_{1i},..,x_{mi}) \cdot \varphi_n(x_{1i},..,x_{mi}) \\ \ldots & \ldots & \ldots & \ldots \\ \sum\limits_{i=1}^{N} \varphi_n(x_{1i},..,x_{mi}) & \sum\limits_{i=1}^{N} \varphi_1(x_{1i},..,x_{mi}) \cdot \varphi_n(x_{1i},..,x_{mi}) & \ldots & \sum\limits_{i=1}^{N} \varphi_n{}^2(x_{1i},..,x_{mi}) \end{vmatrix} \neq 0$$

Тогда

$$D_0 = \begin{vmatrix} \sum_{i=1}^{N} y_i & \sum_{i=1}^{N} \varphi_1(x_{1i},..,x_{mi}) & \dots & \sum_{i=1}^{N} \varphi_n(x_{1i},..,x_{mi}) \\ \sum_{i=1}^{N} y_i \varphi_1(x_{1i},..,x_{mi}) & \sum_{i=1}^{N} \varphi_1{}^2(x_{1i},..,x_{mi}) & \dots & \sum_{i=1}^{N} \varphi_1(x_{1i},..,x_{mi}) \cdot \varphi_n(x_{1i},..,x_{mi}) \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^{N} y_i \varphi_n(x_{1i},..,x_{mi}) & \sum_{i=1}^{N} \varphi_1(x_{1i},..,x_{mi}) \cdot \varphi_n(x_{1i},..,x_{mi}) & \dots & \sum_{i=1}^{N} \varphi_n{}^2(x_{1i},..,x_{mi}) \end{vmatrix}$$

$$a_0 = \frac{D_0}{D}$$

$$D_1 = \begin{vmatrix} N & \sum_{i=1}^{N} y_i & \dots & \sum_{i=1}^{N} \varphi_n(x_{1i},..,x_{mi}) \\ \sum_{i=1}^{N} \varphi_1(x_{1i},..,x_{mi}) & \sum_{i=1}^{N} y_i \varphi_1(x_{1i},..,x_{mi}) & \dots & \sum_{i=1}^{N} \varphi_1(x_{1i},..,x_{mi}) \cdot \varphi_n(x_{1i},..,x_{mi}) \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^{N} \varphi_n(x_{1i},..,x_{mi}) & \sum_{i=1}^{N} y_i \varphi_n(x_{1i},..,x_{mi}) & \dots & \sum_{i=1}^{N} \varphi_n{}^2(x_{1i},..,x_{mi}) \end{vmatrix}$$

$$a_1 = \frac{D_1}{D}$$

Последний определитель будет

$$D_n = \begin{vmatrix} N & \sum_{i=1}^{N} \varphi_1(x_{1i},..,x_{mi}) & \dots & \sum_{i=1}^{N} y_i \\ \sum_{i=1}^{N} \varphi_1(x_{1i},..,x_{mi}) & \sum_{i=1}^{N} \varphi_1{}^2(x_{1i},..,x_{mi}) & \dots & \sum_{i=1}^{N} y_i\varphi_1(x_{1i},..,x_{mi}) \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^{N} \varphi_n(x_{1i},..,x_{mi}) & \sum_{i=1}^{N} \varphi_1(x_{1i},..,x_{mi})\cdot\varphi_n(x_{1i},..,x_{mi}) & \dots & \sum_{i=1}^{N} y_i\varphi_n(x_{1i},..,x_{mi}) \end{vmatrix}$$

$$a_n = \frac{D_n}{D}$$

Определители раскрываются через миноры

$$\begin{vmatrix} d_{11} & d_{12} & \dots & d_{1n} \\ d_{21} & d_{22} & \dots & d_{2n} \\ \dots & \dots & \dots & \dots \\ d_{n1} & d_{n2} & \dots & d_{nn} \end{vmatrix} = d_{11} \cdot (-1)^{1+1} \cdot \begin{vmatrix} d_{22} & \dots & d_{2n} \\ \dots & \dots & \dots \\ d_{n2} & \dots & d_{nn} \end{vmatrix} + d_{12} \cdot (-1)^{1+2} \cdot \begin{vmatrix} d_{21} & d_{13} \dots & d_{2n} \\ \dots & \dots & \dots \\ d_{n1} & d_{n3} \dots & d_{nn} \end{vmatrix} +$$

$$+ d_{13} \cdot (-1)^{1+3} \cdot \begin{vmatrix} d_{21} & d_{22} & d_{24} \dots & d_{2n} \\ \dots & & \dots & \dots \\ d_{n1} & d_{n2} & d_{n4} \dots & d_{nn} \end{vmatrix} + \dots + d_{1n} \cdot (-1)^{1+n} \cdot \begin{vmatrix} d_{21} & \dots & d_{2n-1} \\ \dots & \dots & \dots \\ d_{n1} & \dots & d_{nn-1} \end{vmatrix}$$

В Excel определители можно посчитать с помощью функции МОПРЕД(определитель).

Эконометрика

КОРРЕЛЯЦИЯ ДЛЯ НЕЛИНЕЙНОЙ РЕГРЕССИИ. КОЭФФИЦИЕНТЫ ЭЛАСТИЧНОСТИ

Корреляция в случае нелинейной регрессии

Уравнение нелинейной регрессии дополняется показателями корреляции:

$$\boldsymbol{R}^2 = \frac{\delta^2}{\sigma^2} = \left(1 - \frac{\varepsilon^2}{\sigma^2}\right) = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2},$$

где δ^2 – **объясненная уравнением регрессии дисперсия результирующего признака,** а ε^2– **остаточная** (необъясненная уравнением регрессии) **дисперсия результирующего признака,** σ^2 – **полная дисперсия результирующего признака.** Величину R^2 называют **показателем (индексом) корреляции.** Она изменяется в границах от 0 до 1 и чем она ближе к 1**,** тем теснее связь рассматриваемых признаков, тем более надежно уравнение регрессии.

Коэффициенты эластичности

- характеристика силы связи фактора с результатом, показывающая, на сколько процентов изменится значение результата при изменении каждого фактора на 1%. Коэффициент эластичности (в случае парной регрессии) рассчитывается как:

$$Э = \frac{dy}{dx} : \frac{y}{x} = \frac{dy}{dx} \cdot \frac{x}{y}$$

Различают

Средние коэффициенты эластичности

$$Э_{\bar{x}} = \frac{dy}{dx} \cdot \frac{\bar{x}}{y(\bar{x})}$$

и точечные коэффициенты эластичности

$$Э_{x_0} = \frac{dy}{dx} \cdot \frac{x_0}{y(x_0)}$$

Они показывают на сколько процентов изменится значение *y* при росте *x* на 1% относительно среднего уровня или уровня x_0.

Введем коэффициенты эластичности для различных функций регрессии

1. Для линейной функции $y = \alpha + \beta x$

Коэффициент эластичности будет:

$$Э_{(x_0)} = \frac{\beta x_0}{\alpha + \beta x_0}, \qquad Э_{\bar{x}} = \frac{\beta \bar{x}}{\alpha + \beta \bar{x}}$$

2. Для параболы $y = \alpha + \beta x + \gamma x^2$

Коэффициент эластичности будет:

$$Э_{(x_0)} = \frac{(\beta + 2\gamma x_0)x_0}{\alpha + \beta x_0 + \gamma x_0^2}, \qquad Э_{\bar{x}} = \frac{(\beta + 2\gamma \bar{x})\bar{x}}{\alpha + \beta \bar{x} + \gamma \bar{x}^2}$$

Введем коэффициенты эластичности для различных функций регрессии

3. Для равносторонней гиперболы $y = \alpha + \dfrac{\beta}{x}$

Коэффициент эластичности будет:

$$Э_{(x_0)} = \frac{-\beta}{x_0^{\ 2}} \cdot \frac{1}{\alpha + \dfrac{\beta}{x_0}} \cdot x_0 = \frac{-\beta x_0}{\dfrac{x_0^2 \cdot (\alpha x_0 + \beta)}{x_0}} =$$

$$\frac{-\beta}{\alpha x_0 + \beta}, \qquad Э_{\bar{x}} = \frac{-\beta}{\alpha \bar{x} + \beta}$$

4. Для степенной функции $y = \alpha + x^{\beta}$

Коэффициент эластичности будет:

$$Э_{(x_0)} = \frac{\alpha \cdot \beta x_0^{\beta}}{\alpha \cdot x_0^{\beta}} = \beta, \qquad Э_{\bar{x}} = \beta$$

Введем коэффициенты эластичности для различных функций регрессии

5. Для показательной функции $y = \alpha \cdot \beta^{x}$

Коэффициент эластичности будет:

$$Э_{(x_0)} = \frac{\alpha \ln \beta \cdot \beta^{x_0} \cdot x_0}{\alpha \beta^{x_0}} = x_0 \ln \beta, \qquad Э_{\bar{x}} = \bar{x} \cdot \ln \beta$$

Частные коэффициенты

В случае многомерной функции регрессии можно рассчитать **частные коэффициенты эластичности** показывают на сколько процентов в среднем изменяется результат с увеличением конкретного фактора x_j на 1% от своего среднего уровня при фиксированном положении других факторов модели. Частный коэффициент эластичности рассчитываются по формуле

$$Э_j = \frac{\partial y}{\partial x_j} \cdot \frac{\overline{x_j}}{\overline{\overline{y(x_1, \ldots, x_n)}}}, \qquad j = 1, \ldots, n$$

β–коэффициенты

Для характеристики степени связи между результирующей переменной и факторными признаками в случае многомерной регрессии используются еще и **стандартизованные частные коэффициенты регрессии – *β*–коэффициенты**. Они показывают, на какую часть своего среднего квадратического отклонения σ_{ij} изменится результат y с увеличением соответствующего фактора x_j на величину своего среднего квадратического отклонения σ_{xj} при неизменном влиянии прочих факторов модели.

Частные коэффициенты эластичности и *β*–коэффициенты можно использовать для ранжирования факторов по силе их влияния на результат. Чем они больше для соответствующего фактора, тем сильнее влияние этого фактора на результат.

Получение β-коэффициентов

Расчет β–коэффициентов осуществляется с помощью нахождения коэффициентов регрессии стандартизованной системы линейных уравнений. Вводится стандартизованная переменная:

$$t_{x_{ji}} = \frac{x_{ji} - \overline{x_j}}{\sigma_{x_j}}, \qquad j = 1,...,m, \qquad t_{y_i} = \frac{y_i - \overline{y}}{\sigma_y}$$

Таким образом, начало отсчета каждой стандартизованной переменной совмещается с ее средним значением, а в качестве единицы изменения принимается ее среднее квадратическое отклонение (σ). Если связь между переменными в естественном масштабе линейная, то изменение начала отсчета и единицы изменения этого свойства не нарушат.

Получение β-коэффициентов

Тогда система линейных уравнений для нахождения β–коэффициентов будет:

$$\begin{cases} r_{x_1 y} = \beta_1 + r_{x_1 x_2}\beta_2 + \ldots + r_{x_1 x_m}\beta_m \\ r_{x_2 y} = r_{x_2 x_1}\beta_1 + \beta_2 + \ldots + r_{x_2 x_m}\beta_m \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ r_{x_m y} = r_{x_m x_1}\beta_1 + r_{x_m x_2}\beta_2 + \ldots + \beta_m \end{cases},$$

где

$$r_{x_j x_k} = \sum_{i=1}^{n} t_{x_{ji}} \cdot t_{x_{ki}}$$

Если коэффициенты этой системы найдены, то решение системы уравнений в естественном масштабе будут:

$$b_j = \beta_j \cdot \frac{\sigma_y}{\sigma_x}, \qquad j = 1,\ldots,m, \qquad a = \overline{y} - \sum_{j=1}^{m} b_j \cdot \overline{x_j}$$

$$\text{для} \qquad \text{функции} \qquad y = a + b_1 x_1 + \ldots.. + b_m x_m$$

Частные коэффициенты эластичности

- Эмпирические частные коэффициенты эластичности;
- Частные коэффициенты эластичности или оценки частных коэффициентов эластичности.

Эмпирические частные коэффициенты эластичности

Рассчитываются по каждому фактору модели, для *j* фактора он будет равен:

$$Э_{ji} = \frac{\Delta y_i}{\Delta x_{ji}} \cdot \frac{x_{ji}}{y_i},$$

где $\Delta y_i = y_i - y_{i-1}, \qquad \Delta x_{ji} = x_{ji} - x_{ji-1}$

Итоговый коэффициент эластичности равен

$$\overline{Э_j} = \frac{1}{m} \cdot \sum_{i=1}^{m} Э_{ji},$$

Оценки частных коэффициентов эластичности

Рассчитываются для каждого фактора модели. Для *j* фактора они равны

$$\hat{Э}_{ji} = \hat{a}_j \cdot \frac{x_{ji}}{\hat{y}_i}$$

Итоговый коэффициент для каждого фактора равен:

$$\overline{\hat{Э}_{ji}} = \frac{1}{n} \sum_{i=1}^{n} \hat{Э}_{ji}$$

Частная корреляция

Показатели парной корреляции – r_{yx} характеризуют тесноту связи результата и фактора, не принимая во внимание возможного влияния на результат других факторов модели. Поэтому в множетсвенном регрессионном анализе возникает проблема определения тесноты связи между двумя разными факторами.

$$r_{yx_i / x_1 \ldots x_{i-1}, x_{i+1} \ldots x_m} = \sqrt{1 - \frac{1 - R^2_{y(x_1 \ldots . x_m)}}{1 - R^2_{y(x_1 \ldots x_{i-1}, x_{i+1} \ldots x_m)}}}$$

где $R^2_{y(x_1 \ldots . x_m)}$ - коэффициент множественной детерминации у с комплексом факторов $x_1, \ldots, x_m$;

а $R^2_{y(x_1 \ldots x_{i-1}, x_{i+1} \ldots x_m)}$ - коэффициент множественной детерминации

у с комплексом факторов $x_1 \ldots x_{i-1}, x_{i+1} \ldots x_m$

Частная корреляция

Частные коэффициенты корреляции используются для ранжирования факторов в модели по степени влияния на результат.

Они изменяются на промежутке от 0 до 1, и чем ближе они к 1, тем сильнее влияет этот фактор на результат, а чем ближе к 0, тем слабее.

Их также используют для отсева факторов.

Эконометрика

ПРОБЛЕМА МУЛЬТИКОЛЛИНИАРНОСТИ. ФИКТИВНЫЕ ПЕРЕМЕННЫЕ

Мультиколлинеарность

Мультиколлинеарность – это нестрогая линейная зависимость между факторными признаками (что противоречит предпосылкам применения МНК для поиска параметров функции регрессии). **Мультиколлинеарность может привести к следующим неприятным последствиям**:

- Оценки параметров станут ненадежными (большие статистические ошибки, малая значимость), при этом сама модель может быть в целом значима (завышенное значение множественного коэффициента корреляции).
- Небольшое изменение исходных данных приведет к существенному изменению оценок параметров регрессии.
- Оценки параметров модели будут иметь неправильные с точки зрения теории знаки или чрезмерно большие значения (модель будет непригодна для прогнозирования).
- Делает невозможным определение изолированного влияния факторов на результат.

Снижение мультиколлинеарности

Мультиколлинеарность не всегда оказывает неблагоприятное влияние, если другие условия благоприятны:

- Число наблюдений значительно.
- Выборочные дисперсии факторов велики, а дисперсия случайной составляющей мала.
- При условии влияния этих благоприятных факторов, оценки параметров могут оказаться вполне приемлемы.

Обнаружение мультиколлинеарности

На практике **о наличии мультиколлинеарности** судят:

По матрице парных коэффициентов корреляции (корреляционной матрице):

$$
\begin{pmatrix}
1 & r_{01} & \dots & r_{0j} & \dots & r_{0n} \\
r_{10} & 1 & \dots & r_{1j} & \dots & r_{1n} \\
\dots & \dots & \dots & \dots & \dots & \dots \\
r_{j0} & r_{j1} & \dots & 1 & \dots & r_{jn} \\
\dots & \dots & \dots & \dots & \dots & \dots \\
r_{n0} & r_{n1} & \dots & r_{nj} & \dots & 1
\end{pmatrix},
$$

где r_{jk} – коэффициенты парной линейной корреляции между j-м и k-м факторами (j, k=1,…, n), а r_{0j} – коэффициент парной линейной корреляции между результатом и j-м фактором (j=1,…, n). На главной диагонали стоят единицы, так как там стоят коэффициенты, показывающие степень связи признаков самих с собой. Матрица является симметричной относительно главной диагонали ($r_{jk}=r_{kj}$).

На практике можно воспользоваться

Парными коэффициентами корреляции:

$$r_{x_j y} = \frac{\sum_{i=1}^{N} (x_{ji} - \overline{x}_j)(y_i - \overline{y})}{\sqrt{\sum_{i=1}^{N} (x_{ji} - \overline{x_j})^2} \cdot \sqrt{\sum_{i=1}^{N} (y_i - \overline{y})^2}},$$

$$r_{x_j x_k} = \frac{\sum_{i=1}^{N} (x_{ji} - \overline{x}_j)(x_{ki} - \overline{x}_k)}{\sqrt{\sum_{i=1}^{N} (x_{ji} - \overline{x_j})^2} \cdot \sqrt{\sum_{i=1}^{N} (x_{ki} - \overline{x}_k)^2}}$$

Если имеет место мультиколлинеарность, то в модель следует включать не все факторы, а только те факторы, которые менее ответственны за нее (имеют меньшие по модулю значения коэффициентов корреляции), при условии, что качество модели снижается несущественно.

Обнаружение мультиколлинеарности

2. По величине коэффициентов множественной детерминации

$$R^2_{x_j(x_1,\dots,x_{j-1},x_{j+1},\dots,x_n)} = \sum_{j=1}^{m} \beta_j \cdot r_{yx_i} =$$

$$= 1 - \frac{\sum_{i=1}^{N}\left(y_i - \hat{y}_{x_j(x_1,\dots,x_{j-1},x_{j+1},\dots,x_n)}\right)^2}{\sum_{i=1}^{N}(y_i - \bar{y})^2}$$

которые показывают зависимость фактора x_j других факторов. Чем он ближе этот коэффициент к 1, тем больше ответственность за мультиколлинеарность конкретного фактора. Сравнивая коэффициенты множественной детерминации для различных факторов, можно проранжировать их по степени ответственности за мультиколлинеарность.

Фиктивные переменные

В некоторых случаях в модель необходимо ввести некоторую качественную переменную, изменение которой может существенно влиять на результат.

В данном случае качественная переменная может быть введена в уравнение в форме фиктивной переменной. Для этого вводится система цифровых обозначений и в модель включается фактор, принимающий одно из значений в рамках заданной системы.

Например, разный уровень образования работников дает разный прирост уровня их заработной платы и может быть введен в модель на равне с возрастом и стажем.