

Уравнение множественной регрессии

$$y_t = a_0 + a_1 x_{1t} + a_2 x_{2t} + a_3 x_{3t} + \dots + a_k x_{kt} + U_t \quad (8.1)$$

Наилучшая линейная процедура получения оценок параметров уравнения (8.1) и условия, при которых эта процедура дает несмещенные и эффективные оценки, сформулирована в теореме Гаусса-Маркова.

Постановка задачи:

Имеем случайную выборку наблюдений за поведением экономического объекта объемом n .

Выборка

$$\begin{array}{cccc} y_1 & x_{11} & x_{21} & x_{31} \dots x_{k1} \\ y_2 & x_{12} & x_{22} & x_{32} \dots x_{k2} \\ y_3 & x_{13} & x_{23} & x_{33} \dots x_{k3} \\ \vdots & \vdots & \vdots & \vdots \\ y_n & x_{1n} & x_{2n} & x_{3n} \dots x_{kn} \end{array}$$

Система уравнений наблюдений

$$\begin{aligned} y_1 &= a_0 + a_1 x_{11} + a_2 x_{21} + a_3 x_{31} + \dots + a_k x_{k1} + u_1 \\ y_2 &= a_0 + a_1 x_{12} + a_2 x_{22} + a_3 x_{32} + \dots + a_k x_{k2} + u_2 \\ y_3 &= a_0 + a_1 x_{13} + a_2 x_{23} + a_3 x_{33} + \dots + a_k x_{k3} + u_3 \\ &\vdots \\ y_n &= a_0 + a_1 x_{1n} + a_2 x_{2n} + a_3 x_{3n} + \dots + a_k x_{kn} + u_n \end{aligned} \quad (8.2)$$

Уравнение множественной регрессии

Вводятся обозначения:

$$X = \begin{pmatrix} x_{11} & x_{21} & \cdots & x_{k1} \\ x_{12} & x_{22} & \cdots & x_{k2} \\ \dots & \dots & \dots & \dots \\ x_{1n} & x_{2n} & \cdots & x_{kn} \end{pmatrix} \quad \vec{A} = \begin{pmatrix} a_0 \\ a_1 \\ \dots \\ a_k \end{pmatrix} \quad \vec{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} \quad \vec{U} = \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix}$$

X – матрица коэффициентов при неизвестных параметрах;

Y – вектор выборочных значений эндогенной переменной;

U – вектор выборочных значений случайного возмущения;

A – вектор неизвестных параметров модели.

Найти: $\tilde{A}$, $\text{Cov}(\tilde{A}\tilde{A})$, σ_u , $\sigma(\tilde{y}(z))$

Уравнение множественной регрессии

Теорема Гаусса-Маркова.

Если матрица X неколлинеарна и вектор случайных возмущений удовлетворяет следующим требованиям:

1. $M(u) = 0$

2. $\sigma^2(u) = \sigma_u^2$

3. $\text{Cov}(u_i, u_j) = 0$
при $i \neq j$

4. $\text{Cov}(x_i, u_i) = 0$

Уравнение множественной регрессии

Теорема Гаусса-Маркова (Продолжение).

Тогда наилучшей линейной процедурой оценки параметров модели (8.1) является:

$$\hat{A} = (X^T X)^{-1} X^T Y$$

При этом:

$$\begin{aligned} \text{Cov}(\hat{A}, \hat{A}) &= \tilde{\sigma}_u^2 (X^T X)^{-1} \\ \tilde{\sigma}_u^2 &= \frac{1}{n-k} \sum (y_i - \tilde{y}_i)^2 = \sum u_i^2 \\ \tilde{Y}(Z) &= \tilde{a}_0 + a_1 z_1 + \dots + \tilde{a}_k z_k \\ \sigma^2(\tilde{y}(Z)) &= \tilde{\sigma}_u^2 (1 + q_0) \\ q_0 &= Z^T (X^T X)^{-1} Z \end{aligned}$$

Теорема Гаусса-Маркова

Примеры применения теоремы Гаусса-Маркова

Пример 1. Пусть имеем выборку из n наблюдений за случайной величиной Y .

Найти наилучшие оценки среднего значения и дисперсии этой переменной.

В терминах теоремы Гаусса –Маркова задача формулируется так: необходимо построить модель типа $Y = a_0 + u$, при этом имеем:

$$X = \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} \quad X^T = (1 \ 1 \ \dots \ 1)$$

Примеры применения теоремы Гаусса-Маркова

Пример 1. (Продолжение)

Решение.

$$\begin{aligned} (X^T X) &= (1 \ 1 \ \dots \ 1) \begin{pmatrix} 1 \\ 1 \\ \dots \\ 1 \end{pmatrix} = n & (X^T X)^{-1} &= \frac{1}{n} \\ (X^T \tilde{Y}) &= (1 \ 1 \ \dots \ 1) \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} = \sum y_i & \tilde{a}_0 &= (X^T X)^{-1} (X^T \tilde{Y}) = \frac{\sum y_i}{n} \\ \sigma^2(Y) &= \sigma_u^2 = \frac{1}{n-1} \sum (y_i - \tilde{a}_0)^2 & \sigma^2(a_0) &= \sigma_u^2 (X^T X)^{-1} = \frac{\sigma_u^2}{n} \end{aligned}$$

Примеры применения теоремы Гаусса-Маркова

Пример 2. Уравнение парной регрессии.

Построить модель типа $Y = a_0 + a_1 x + u$, по данным выборки наблюдений за переменными Y и x объемом n .

В схеме Гаусса-Маркова имеем:

$$X = \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{pmatrix} \quad X^T = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix} \quad Y = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} \quad A = \begin{pmatrix} a_0 \\ a_1 \end{pmatrix} \quad U = \begin{pmatrix} u_1 \\ u_2 \\ \dots \\ u_n \end{pmatrix}$$

$$(X^T X) = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \dots & \dots \\ 1 & x_n \end{pmatrix} = \begin{pmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{pmatrix} \quad (X^T X)^{-1} = \frac{1}{n \sum x_i^2 - \sum x_i \sum x_i} \begin{pmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & n \end{pmatrix}$$

Примеры применения теоремы Гаусса-Маркова

Пример 2. Уравнение парной регрессии.(Продолжение).

Последовательно вычисляем $X^T Y$ и оценку вектора A .

$$\begin{pmatrix} X^T \\ Y \end{pmatrix} = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix} = \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix}$$

Система нормальных уравнений имеет вид :

$$(X^T X) A = (X^T Y)$$

$$\tilde{A} = \frac{1}{n \sum x_i^2 - \sum x_i \sum x_i} \begin{pmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & n \end{pmatrix} \begin{pmatrix} \sum y_i \\ \sum x_i y_i \end{pmatrix}$$

$$\tilde{A} = \frac{1}{n \sum x_i^2 - \sum x_i \sum x_i} \begin{pmatrix} \sum x_i^2 \sum y_i - \sum x_i \sum x_i y_i \\ n \sum x_i y_i - \sum x_i \sum y_i \end{pmatrix}$$

(8.3)

Примеры применения теоремы Гаусса-Маркова

Пример 2. Уравнение парной регрессии. (Продолжение).

Вычислим дисперсии (ковариационную матрицу) параметров модели.

$$\text{Cov} \begin{pmatrix} \tilde{A}_0 \\ \tilde{A}_1 \end{pmatrix} = \sigma_u^2 (X^T X)^{-1} = \sigma_u^2 \frac{1}{n \sum x_i^2 - \sum x_i \sum x_i} \begin{pmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & n \end{pmatrix}$$

Следовательно:

$$\sigma^2(\tilde{a}_0) = \sigma_u^2 \frac{\sum x_i^2}{n \sum x_i^2 - \sum x_i \sum x_i}$$

$$\sigma^2(\tilde{a}_1) = \sigma_u^2 \frac{n}{n \sum x_i^2 - \sum x_i \sum x_i}$$

$$\text{Cov}(\tilde{a}_0, \tilde{a}_1) = \sigma_u^2 \frac{-\sum x_i}{n \sum x_i^2 - \sum x_i \sum x_i}$$

Примеры применения теоремы Гаусса-Маркова

Пример 2. Уравнение парной регрессии.(Продолжение).

Расчет дисперсии прогнозирования.

Прогноз осуществляется в точке $Z=\{1,z\}$

$$\sigma^2(Y(z)) = \sigma_u^2(1 + q_0)$$

$$\begin{aligned} q_0 &= Z^T (X^T X)^{-1} Z = (1 \ z) \frac{1}{n \sum x_i^2 - \sum x_i \sum x_i} \begin{pmatrix} \sum x_i^2 & -\sum x_i \\ -\sum x_i & n \end{pmatrix} \begin{pmatrix} 1 \\ z \end{pmatrix} = \\ &= \frac{1}{n \sum x_i^2 - \sum x_i \sum x_i} \begin{pmatrix} \sum x_i^2 & -z \sum x_i \\ -\sum x_i & nz \end{pmatrix} \begin{pmatrix} 1 \\ z \end{pmatrix} = \frac{\sum x_i^2 - 2z \sum x_i + n z^2}{n \sum x_i^2 - \sum x_i \sum x_i} = \frac{\sum x_i^2 - \frac{1}{n} \sum x_i^2 \sum x_i^2 + (z - \bar{x})^2}{n \sum x_i^2 - \sum x_i \sum x_i} \\ \sigma^2(Y(z)) &= \sigma_u^2 \left(1 + \frac{1}{n} + \frac{(z - \bar{x})^2}{n \sum (x_i - \bar{x})^2} \right) \end{aligned}$$

Примеры применения теоремы Гаусса-Маркова

Пример 2. Уравнение парной регрессии.(Продолжение).