Методы линейной множественной регрессии в матричной форме

к.э.н. Ивашнев Л.И.
Университет машиностроения
8(985) 284-26-98

Аннотация. Статья содержит краткое изложение трех базовых и трех взвешенных методов линейной множественной регрессии в матричной форме, которые вместе с методом наименьших квадратов К. Гаусса составляют новый инструментарий регрессионного анализа. Статья содержит матричные формулы, которые могут использоваться для получения уравнений линейной множественной регрессии базовым и взвешенным методом наименьших квадратов, методом получения уравнений регрессии без свободного члена и методом получения уравнений регрессии общего вида. В статье дан пример применения матричных методов для получения коэффициентов уравнения регрессии общего вида, т.е. уравнения от равноправных показателей.

Ключевые слова: *матричная формула, базовые и взвешенные методы, линейная множественная регрессия, метод наименьших квадратов, метод получения уравнений регрессии без свободного члена, метод получения уравнений регрессии общего вида*

В учебном пособии [3] изложены 6 методов линейной множественной регрессии, которые можно разделить на:

три *базовых метода*:

- метод наименьших квадратов [1],
- метод получения вырожденных уравнений регрессии [2],
- метод получения уравнений регрессии общего вида [3],

и три *метода взвешенной линейной множественной регрессии* [3]:

- взвешенный метод наименьших квадратов,
- взвешенный метод получения вырожденных уравнений регрессии,
- взвешенный метод получения уравнений регрессии общего вида.

Также в этом пособии дана матричная форма каждого из этих методов.

Из этих методов широко известен только метод наименьших квадратов, разработанный К. Гауссом, остальные методы пока не получили широкого признания. Эти методы не включены в учебные программы ВУЗов и почти не знакомы научным работникам и специалистам. Тем не менее, применяя новые методы можно значительно повысить точность и качество работ в области моделирования и в прогнозных и аналитических работах. Более того, эти методы могут найти применение практически в любых отраслях науки и техники.

Изложение методов регрессии в матричной форме дано в разных главах книги [3] и не создает целостной картины матричного регрессионного анализа. Поэтому, в настоящей статье делается попытка такого однообразного изложения матричных способов представления методов линейной множественной регрессии.

При расчете коэффициентов линейного уравнения множественной регрессии методом наименьших квадратов (базовым или взвешенным) используется уравнение регрессии вида:

$$y = a_0 + a_1x_1 + a_2x_2 + ... + a_mx_m . \qquad (1)$$

Для получения вырожденного уравнения регрессии, т.е. уравнения регрессии без свободного члена (базового или взвешенного) используется уравнение:

$$y = a_1x_1 + a_2x_2 + ... + a_mx_m . \qquad (2)$$

Определение коэффициентов уравнения регрессии общего вида (базового или взвешенного) выполняется для уравнения регрессии вида:

$$a_1x_1 + a_2x_2 + ... + a_mx_m = 1 . \qquad (3)$$

При этом используются следующие матрицы:

$$X=\begin{Bmatrix} x_{11} & x_{21} & \dots & x_{m1}\\ x_{12} & x_{22} & \dots & x_{m2}\\ \dots & \dots & \dots & \dots\\ x_{1n} & x_{2n} & \dots & x_{mn}\end{Bmatrix},\ X^*=\begin{Bmatrix} x_{11} & x_{12} & \dots & x_{1n}\\ x_{21} & x_{21} & \dots & x_{2n}\\ \dots & \dots & \dots & \dots\\ x_{m1} & x_{m2} & \dots & x_{mn}\end{Bmatrix},\ Y=\begin{Bmatrix} y_1\\ y_2\\ \dots\\ y_n\end{Bmatrix},\ P=\begin{Bmatrix} p_1\\ p_2\\ \dots\\ p_n\end{Bmatrix},\ \bar{\mathbf{1}}=\begin{Bmatrix} 1\\ 1\\ \dots\\ 1\end{Bmatrix},\ A=\begin{Bmatrix} a_0\\ a_1\\ \dots\\ a_m\end{Bmatrix},$$

причем матрицы Y и X состоят из элементов исходной выборки, матрица X^* – транспонированная матрица X, P – вектор, т.е. матрица-столбец весовых коэффициентов строк исходной выборки, $\bar{\mathbf{1}}$ – единичный вектор, т.е. матрица-столбец, элементами которой являются единицы, а матрица-столбец A состоит из неизвестных коэффициентов искомого уравнения регрессии.

При применении метода наименьших квадратов осуществляется расчет коэффициентов уравнения регрессии (1), причем используется матрица $\overline{X}$, которая получается добавлением в матрицу X слева единичного столбца, т.е. столбца, состоящего из единиц. Используется также матрица $\overline{X}^*$, получающаяся транспонированием матрицы $\overline{X}$, т.е. используются матрицы:

$$\overline{X}^*=\begin{Bmatrix} 1 & 1 & \dots & 1\\ x_{11} & x_{12} & \dots & x_{1n}\\ x_{21} & x_{22} & \dots & x_{2n}\\ \dots & \dots & \dots & \dots\\ x_{m1} & x_{m2} & \dots & x_{mn}\end{Bmatrix},\ \overline{X}=\begin{Bmatrix} 1 & x_{11} & x_{21} & \dots & x_{m1}\\ 1 & x_{12} & x_{22} & \dots & x_{m2}\\ \dots & \dots & \dots & \dots & \dots\\ 1 & x_{1n} & x_{2n} & \dots & x_{mn}\end{Bmatrix}.$$

Перемножая эти матрицы получаем матрицу:

$$\overline{X}^*\overline{X}=\begin{Bmatrix} 1 & 1 & \dots & 1\\ x_{11} & x_{12} & \dots & x_{1n}\\ x_{21} & x_{22} & \dots & x_{2n}\\ \dots & \dots & \dots & \dots\\ x_{m1} & x_{m2} & \dots & x_{mn}\end{Bmatrix}\begin{Bmatrix} 1 & x_{11} & x_{21} & \dots & x_{m1}\\ 1 & x_{12} & x_{22} & \dots & x_{m2}\\ \dots & \dots & \dots & \dots & \dots\\ 1 & x_{1n} & x_{2n} & \dots & x_{mn}\end{Bmatrix}=\begin{Bmatrix} n & \sum\limits_{i=1}^{n} x_{1i} & \sum\limits_{i=1}^{n} x_{2i} & \dots & \sum\limits_{i=1}^{n} x_{mi}\\ \sum\limits_{i=1}^{n} x_{1i} & \sum\limits_{i=1}^{n} x_{1i}x_{1l} & \sum\limits_{i=1}^{n} x_{1i}x_{2i} & \dots & \sum\limits_{i=1}^{n} x_{1i}x_{mi}\\ \dots & \dots & \dots & \dots & \dots\\ \sum\limits_{i=1}^{n} x_{mi} & \sum\limits_{i=1}^{n} x_{mi}x_{1i} & \sum\limits_{i=1}^{n} x_{mi}x_{2i} & \dots & \sum\limits_{i=1}^{n} x_{mi}x_{mj}\end{Bmatrix}. \quad (4)$$

Эта матрица состоит из коэффициентов левой части (до знаков равенства) системы нормальных уравнений, использующейся для расчета коэффициентов линейного уравнения множественной регрессии методом наименьших квадратов, которая имеет вид:

$$\begin{cases} na_0+a_1\sum\limits_{i=1}^{n} x_{1i}+a_2\sum\limits_{i=1}^{n} x_{2i}+\dots+a_m\sum\limits_{i=1}^{n} x_{mi}=\sum\limits_{i=1}^{n} y_i\\ a_0\sum\limits_{i=1}^{n} x_{1i}+a_1\sum\limits_{i=1}^{n} x_{1i}x_{1i}+a_2\sum\limits_{i=1}^{n} x_{1i}x_{2i}+\dots+a_m\sum\limits_{i=1}^{n} x_{1i}x_{mi}=\sum\limits_{i=1}^{n} x_{1i}y_i\\ a_0\sum\limits_{i=1}^{n} x_{2i}+a_1\sum\limits_{i=1}^{n} x_{2i}x_{1i}+a_2\sum\limits_{i=1}^{n} x_{2i}x_{2i}+\dots+a_m\sum\limits_{i=1}^{n} x_{2i}x_{mi}=\sum\limits_{i=1}^{n} x_{2i}y_i\\ \dots\dots\dots\dots\dots\dots\dots\dots\\ a_0\sum\limits_{i=1}^{n} x_{mi}+a_1\sum\limits_{i=1}^{n} x_{mi}x_{1i}+a_2\sum\limits_{i=1}^{n} x_{mi}x_{2i}+\dots+a_m\sum\limits_{i=1}^{n} x_{mi}x_{mi}=\sum\limits_{i=1}^{n} x_{mi}y_i \end{cases} \quad (5)$$

Обратная матрица $\left(\overline{X}^*\overline{X}\right)^{-1}$ от матрицы (4) может использоваться для расчета коэффициентов уравнения регрессии матричным методом наименьших квадратов.

Те же операции выполним для получения вырожденного уравнения лилейной множественной регрессии, т.е. уравнения без свободного члена. Вырожденным такое уравнение называется потому, что нулевые значения аргументов и функции являются дополнительным решением соответствующего уравнения. К числу таких функций относится, например, выручка, которая равна нулю при нулевых количествах проданного товара.

В этом случае, перемножая матрицы X^*X получаем матрицу:

$$X^*X = \begin{Bmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{21} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{m1} & x_{m2} & \dots & x_{mn} \end{Bmatrix} \begin{Bmatrix} x_{11} & x_{21} & \dots & x_{m1} \\ x_{12} & x_{22} & \dots & x_{m2} \\ \dots & \dots & \dots & \dots \\ x_{1n} & x_{2n} & \dots & x_{mn} \end{Bmatrix} = \begin{Bmatrix} \sum_{i=1}^{n} x_{1i}x_{1i} & \sum_{i=1}^{n} x_{1i}x_{2i} & \dots & \sum_{i=1}^{n} x_{1i}x_{mi} \\ \sum_{i=1}^{n} x_{2i}x_{1i} & \sum_{i=1}^{n} x_{2i}x_{2i} & \dots & \sum_{i=1}^{n} x_{2i}x_{mi} \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^{n} x_{mi}x_{1i} & \sum_{i=1}^{n} x_{mi}x_{2i} & \dots & \sum_{i=1}^{n} x_{mi}x_{mi} \end{Bmatrix}. \quad (6)$$

Эта матрица состоит из коэффициентов левой части системы нормальных уравнений, использующихся для расчета коэффициентов вырожденного уравнения регрессии, т.е. уравнения регрессии без свободного члена (2), причем эта система нормальных уравнений имеет вид:

$$\begin{cases} a_0\sum_{i=1}^{n} x_{1i}x_{1i} + a_1\sum_{i=1}^{n} x_{1i}x_{2i} + \dots + a_m\sum_{i=1}^{n} x_{1i}x_{mi} = \sum_{i=1}^{n} x_{1i}y_i \\ a_0\sum_{i=1}^{n} x_{2i}x_{1i} + a_1\sum_{i=1}^{n} x_{2i}x_{2i} + \dots + a_m\sum_{i=1}^{n} x_{2i}x_{mi} = \sum_{i=1}^{n} x_{2i}y_i \\ \dots\dots\dots\dots\dots\dots\dots\dots\dots\dots \\ a_0\sum_{i=1}^{n} x_{mi}x_{1i} + a_1\sum_{i=1}^{n} x_{mi}x_{2i} + \dots + a_m\sum_{i=1}^{n} x_{mi}x_{mi} = \sum_{i=1}^{n} x_{mi}y_i \end{cases} \quad (7)$$

Вычисляя обратную матрицу $(X^*X)^{-1}$ от матрицы (6) получаем промежуточную матрицу для расчета искомых коэффициентов вырожденного уравнения регрессии или уравнения регрессии общего вида.

Далее рассмотрим взвешенные методы регрессии, причем наиболее просто получить взвешенное вырожденное уравнение линейной множественной регрессии. В этом случае исходная выборка дополняется столбцом весовых коэффициентов, которые могут означать кратность повторения строки исходной выборки или ее важность. Хотя эта операция и не признается в качестве матричной операции, однако выполняя операцию поэлементного умножения каждого элемента вектора весовых коэффициентов P на элементы соответствующей строки матрицы X получаем матрицу:

$$\begin{Bmatrix} p_1 \\ p_2 \\ \dots \\ p_n \end{Bmatrix} * \begin{Bmatrix} x_{11} & x_{21} & \dots & x_{m1} \\ x_{12} & x_{22} & \dots & x_{m2} \\ \dots & \dots & \dots & \dots \\ x_{1n} & x_{2n} & \dots & x_{mn} \end{Bmatrix} = \begin{Bmatrix} p_1x_{11} & p_1x_{21} & \dots & p_1x_{m1} \\ p_2x_{12} & p_2x_{22} & \dots & p_2x_{m2} \\ \dots & \dots & \dots & \dots \\ p_nx_{1n} & p_nx_{2n} & \dots & p_nx_{mn} \end{Bmatrix}.$$

Если ее транспонировать, то будет получена матрица $\dddot{X}^*$:

$$\dddot{X}^* = \begin{Bmatrix} p_1x_{11} & p_1x_{12} & \dots & p_1x_{1n} \\ p_2x_{21} & p_2x_{21} & \dots & p_2x_{2n} \\ \dots & \dots & \dots & \dots \\ p_nx_{m1} & p_nx_{m2} & \dots & p_nx_{mn} \end{Bmatrix}.$$

Произведение матриц $\dddot{X}^*X$ дает матрицу:

$$\ddot{X}^* X = \begin{Bmatrix} p_1 x_{11} & p_1 x_{12} & \dots & p_1 x_{1n} \\ p_2 x_{21} & p_2 x_{21} & \dots & p_2 x_{2n} \\ \dots & \dots & \dots & \dots \\ p_n x_{m1} & p_n x_{m2} & \dots & p_n x_{mn} \end{Bmatrix} \begin{Bmatrix} x_{11} & x_{21} & \dots & x_{m1} \\ x_{12} & x_{22} & \dots & x_{m2} \\ \dots & \dots & \dots & \dots \\ x_{1n} & x_{2n} & \dots & x_{mn} \end{Bmatrix} = \begin{Bmatrix} \sum_{i=1}^{n} p_i x_{1i} x_{1i} & \sum_{i=1}^{n} p_i x_{1i} x_{2i} & \dots & \sum_{i=1}^{n} p_i x_{1i} x_{mi} \\ \sum_{i=1}^{n} p_i x_{2i} x_{1i} & \sum_{i=1}^{n} p_i x_{2i} x_{2i} & \dots & \sum_{i=1}^{n} p_i x_{2i} x_{mi} \\ \dots & \dots & \dots & \dots \\ \sum_{i=1}^{n} p_i x_{mi} x_{1i} & \sum_{i=1}^{n} p_i x_{mi} x_{2i} & \dots & \sum_{i=1}^{n} p_i x_{mi} x_{mi} \end{Bmatrix}. \quad (8)$$

Элементы матрицы (8) являются коэффициентами левой части системы нормальных уравнений, используемой для расчета коэффициентов взвешенного вырожденного уравнения линейной множественной регрессии и взвешенного уравнения общего вида. Обратная матрица $\left(\ddot{X}^* X\right)^{-1}$ от матрицы (8) используется в таком расчете.

Теперь получим вспомогательную матрицу для расчета коэффициентов взвешенного уравнения линейной множественной регрессии методом наименьших квадратов. Для этого выполним поэлементное умножение элементов единичного вектора на строки матрицы $\overline{X}$:

$$\begin{Bmatrix} p_1 \\ p_2 \\ \dots \\ p_n \end{Bmatrix} * \begin{Bmatrix} 1 & x_{11} & x_{21} & \dots & x_{m1} \\ 1 & x_{12} & x_{22} & \dots & x_{m2} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1n} & x_{2n} & \dots & x_{mn} \end{Bmatrix} = \begin{Bmatrix} p_1 & p_1 x_{11} & p_1 x_{21} & \dots & p_1 x_{m1} \\ p_2 & p_2 x_{12} & p_2 x_{22} & \dots & p_2 x_{m2} \\ \dots & \dots & \dots & \dots & \dots \\ p_n & p_n x_{1n} & p_n x_{2n} & \dots & p_n x_{mn} \end{Bmatrix}.$$

Выполнив транспонирование этой матрицы получаем матрицу $\overline{\ddot{X}}^*$:

$$\overline{\ddot{X}}^* = \begin{Bmatrix} p_1 & p_2 & \dots & p_n \\ p_1 x_{11} & p_2 x_{12} & \dots & p_n x_{1n} \\ p_1 x_{21} & p_2 x_{22} & \dots & p_n x_{2n} \\ \dots & \dots & \dots & \dots \\ p_1 x_{m1} & p_2 x_{m2} & \dots & p_n x_{mn} \end{Bmatrix}$$

и умножив ее на матрицу $\overline{X}$ получаем матрицу $\overline{\ddot{X}}^* \overline{X}$:

$$\overline{\ddot{X}}^* \overline{X} = \begin{Bmatrix} p_1 & p_2 & \dots & p_n \\ p_1 x_{11} & p_2 x_{12} & \dots & p_n x_{1n} \\ p_1 x_{21} & p_2 x_{22} & \dots & p_n x_{2n} \\ \dots & \dots & \dots & \dots \\ p_1 x_{m1} & p_2 x_{m2} & \dots & p_n x_{mn} \end{Bmatrix} \begin{Bmatrix} 1 & x_{11} & x_{21} & \dots & x_{m1} \\ 1 & x_{12} & x_{22} & \dots & x_{m2} \\ \dots & \dots & \dots & \dots & \dots \\ 1 & x_{1n} & x_{2n} & \dots & x_{mn} \end{Bmatrix} = \begin{Bmatrix} \sum_{i=1}^{n} p_i & \sum_{i=1}^{n} p_i x_{1i} & \sum_{i=1}^{n} p_i x_{2i} & \dots & \sum_{i=1}^{n} p_i x_{mi} \\ \sum_{i=1}^{n} p_i x_{2i} & \sum_{i=1}^{n} p_i x_{2i} x_{1i} & \sum_{i=1}^{n} p_i x_{2i} x_{2i} & \dots & \sum_{i=1}^{n} p_i x_{2i} x_{mi} \\ \dots & \dots & \dots & \dots & \dots \\ \sum_{i=1}^{n} p_i x_{mi} & \sum_{i=1}^{n} p_i x_{mi} x_{1u} & \sum_{i=1}^{n} p_i x_{mi} x_{2i} & \dots & \sum_{i=1}^{n} p_i x_{mi} x_{mi} \end{Bmatrix}. \quad (9)$$

Элементы матрицы (9) являются коэффициентами левой части системы нормальных уравнений, используемой для расчета коэффициентов взвешенного уравнения линейной множественной регрессии методом наименьших квадратов. Вычислив для матрицы (9) обратную матрицу получаем промежуточную матрицу $\left(\overline{\ddot{X}}^* \overline{X}\right)^{-1}$, которая может быть использована для расчета коэффициентов уравнения регрессии взвешенным методом наименьших квадратов.

Кроме расчета обратной матрицы надо рассчитать вторую часть расчетной формулы для вектора неизвестных коэффициентов уравнения регрессии, которая для базового метода наименьших квадратов имеет вид:

$$\overline{X}^{*}Y=\begin{Bmatrix}1 & 1 & \dots & 1\\ x_{11} & x_{12} & \dots & x_{1n}\\ x_{21} & x_{22} & \dots & x_{2n}\\ \dots & \dots & \dots & \dots\\ x_{m1} & x_{m2} & \dots & x_{mn}\end{Bmatrix}\begin{Bmatrix}y_1\\ y_2\\ \dots\\ y_m\end{Bmatrix}=\begin{Bmatrix}\sum_{i=1}^{n}y_i\\ \sum_{i=1}^{n}x_{1i}y_i\\ \dots\\ \sum_{i=1}^{n}x_{mi}y_i\end{Bmatrix}, \tag{10}$$

для базового метода получения вырожденного уравнения регрессии:

$$X^{*}Y=\begin{Bmatrix}x_{11} & x_{12} & \dots & x_{1n}\\ x_{21} & x_{21} & \dots & x_{2n}\\ \dots & \dots & \dots & \dots\\ x_{m1} & x_{m2} & \dots & x_{mn}\end{Bmatrix}\begin{Bmatrix}y_1\\ y_2\\ \dots\\ y_m\end{Bmatrix}=\begin{Bmatrix}\sum_{i=1}^{n}x_{1i}y_i\\ \sum_{i=1}^{n}x_{2i}y_i\\ \dots\\ \sum_{i=1}^{n}x_{mi}y_i\end{Bmatrix}, \tag{11}$$

для базового метода получения уравнения регрессии общего вида:

$$X^{*}\overline{1}=\begin{Bmatrix}x_{11} & x_{12} & \dots & x_{1n}\\ x_{21} & x_{21} & \dots & x_{2n}\\ \dots & \dots & \dots & \dots\\ x_{m1} & x_{m2} & \dots & x_{mn}\end{Bmatrix}\begin{Bmatrix}1\\ 1\\ \dots\\ 1\end{Bmatrix}=\begin{Bmatrix}\sum_{i=1}^{n}x_{1i}\\ \sum_{i=1}^{n}x_{2i}\\ \dots\\ \sum_{i=1}^{n}x_{mi}\end{Bmatrix}, \tag{12}$$

для взвешенного метода наименьших квадратов:

$$\overline{\overline{\dddot{X}}}^{*}Y=\begin{Bmatrix}p_1 & p_2 & \dots & p_n\\ p_1x_{11} & p_2x_{12} & \dots & p_nx_{1n}\\ p_1x_{21} & p_2x_{22} & \dots & p_nx_{2n}\\ \dots & \dots & \dots & \dots\\ p_1x_{m1} & p_2x_{m2} & \dots & p_nx_{mn}\end{Bmatrix}\begin{Bmatrix}y_1\\ y_2\\ \dots\\ y_n\end{Bmatrix}=\begin{Bmatrix}\sum_{i=1}^{n}p_iy_i\\ \sum_{i=1}^{n}p_ix_{1i}y_i\\ \dots\\ \sum_{i=1}^{n}p_ix_{mi}y_i\end{Bmatrix}, \tag{13}$$

для взвешенного метода получения вырожденного уравнения регрессии:

$$\dddot{X}^{*}Y=\begin{Bmatrix}p_1x_{11} & p_2x_{12} & \dots & p_nx_{1n}\\ p_1x_{21} & p_2x_{21} & \dots & p_nx_{2n}\\ \dots & \dots & \dots & \dots\\ p_1x_{m1} & p_2x_{m2} & \dots & p_nx_{mn}\end{Bmatrix}\begin{Bmatrix}y_1\\ y_2\\ \dots\\ y_n\end{Bmatrix}=\begin{Bmatrix}\sum_{i=1}^{n}p_ix_{1i}y_i\\ \sum_{i=1}^{n}p_ix_{2i}y_i\\ \dots\\ \sum_{i=1}^{n}p_ix_{mi}y_i\end{Bmatrix}, \tag{14}$$

для взвешенного метода получения уравнения регрессии общего вида:

$$\dddot{X}^*Y=\begin{Bmatrix} p_1x_{11} & p_2x_{12} & \dots & p_nx_{1n} \\ p_1x_{21} & p_2x_{21} & \dots & p_nx_{2n} \\ \dots & \dots & \dots & \dots \\ p_1x_{m1} & p_2x_{m2} & \dots & p_nx_{mn} \end{Bmatrix}\begin{Bmatrix} 1 \\ 1 \\ \dots \\ 1 \end{Bmatrix}=\begin{Bmatrix} \sum_{i=1}^{n} p_ix_{1i} \\ \sum_{i=1}^{n} p_ix_{2i} \\ \dots \\ \sum_{i=1}^{n} p_ix_{mi} \end{Bmatrix}. \tag{15}$$

Итак, для использования базовых методов получения линейных уравнений множественной регрессии могут использоваться следующие матричные формулы:

- базовый метод наименьших квадратов

$$\boldsymbol{A}=\left(\overline{\boldsymbol{X}}^*\,\overline{\boldsymbol{X}}\right)^{-1}\overline{\boldsymbol{X}}^*\boldsymbol{Y}, \tag{16}$$

- базовый метод получения вырожденного уравнения регрессии

$$\boldsymbol{A}=\left(\boldsymbol{X}^*\boldsymbol{X}\right)^{-1}\boldsymbol{X}^*\boldsymbol{Y}, \tag{17}$$

- базовый метод получения уравнения регрессии общего вида

$$\boldsymbol{A}=\left(\boldsymbol{X}^*\boldsymbol{X}\right)^{-1}\boldsymbol{X}^*\bar{1}, \tag{18}$$

- взвешенный метод наименьших квадратов

$$A=\left(\overline{\dddot{X}}^*\,\overline{X}\right)^{-1}\overline{X}^*\,Y, \tag{19}$$

- взвешенный метод получения вырожденного уравнения регрессии

$$A=\left(\dddot{X}^*X\right)^{-1}\overline{X}^*Y. \tag{20}$$

- взвешенный метод получения уравнения регрессии общего вида

$$A=\left(\dddot{X}^*X\right)^{-1}\overline{X}^*\bar{1}. \tag{21}$$

каждая из них представляет собой матрицу-столбец A, элементами которого являются значения коэффициентов a_0, a_1, a_2,…, a_m или a_1, a_2,…, a_m искомого уравнения линейной множественной регрессии.

В качестве примера получим уравнение регрессии общего вида по выборке, представленной в виде матрицы $\overline{X}$, которая имеет следующий вид:

$$X=\begin{Bmatrix} 9{,}526302 & 2{,}5797 & 1{,}505637 \\ 19{,}1572 & 0{,}3678 & 3{,}367738 \\ 20{,}12396 & 1{,}0706 & 3{,}76242 \\ 13{,}36933 & 2{,}9002 & 3{,}112084 \\ 12{,}12286 & 0{,}2893 & 1{,}145446 \\ 5{,}655305 & 2{,}9431 & 0{,}835847 \\ 11{,}32897 & 1{,}861 & 1{,}739339 \end{Bmatrix}.$$

После транспонирования она принимает вид:

$$X^*=\begin{Bmatrix} 9{,}526302 & 19{,}1572 & 20{,}12396 & 13{,}36933 & 12{,}12286 & 5{,}655305 & 11{,}32897 \\ 2{,}579704 & 0{,}367839 & 1{,}070597 & 2{,}900175 & 0{,}289254 & 2{,}943124 & 1{,}860973 \\ 1{,}505637 & 3{,}367738 & 3{,}76242 & 3{,}112084 & 1{,}145446 & 0{,}835847 & 1{,}739339 \end{Bmatrix}.$$

Произведение матриц X^* и X:

$$X^*X = \begin{Bmatrix} 1348,753 & 133,1736 & 234,4988 \\ 133,1736 & 28,55623 & 24,20469 \\ 234,4988 & 24,20469 & 42,48546 \end{Bmatrix},$$

обратная матрица произведения:

$$\left(X^*X\right)^{-1} = \begin{Bmatrix} 0,018375 & 0,000528 & -0,10172 \\ 0,000528 & 0,067736 & -0,04151 \\ -0,10172 & -0,04151 & 0,608626 \end{Bmatrix}$$

и, наконец, произведение транспонированной матрицы на единицу-вектор, то есть единичную матрицу-столбец:

$$X^*\bar{1} = \begin{Bmatrix} 91,28393 \\ 2,01167 \\ 5,46851 \end{Bmatrix}.$$

Произведением обратной матрицы на матрицу-столбец $X^*\bar{1}$ является матрица-столбец:

$$A = (X^*X)^{-1}X^*\bar{1} = \begin{Bmatrix} 0,110198 \\ 0,219807 \\ -0,36938 \end{Bmatrix},$$

Итак, получена матрица-столбец, элементами которого являются коэффициенты a_1, a_2 и a_3 искомого уравнения регрессии общего вида. В итоге получено искомое линейное уравнение регрессии общего вида:

$$0,110198\,X_1 + 0,219807\,X_2 - 0,36938\,X_3 = 1.$$

Важно заметить, что все рассмотренные преобразования можно выполнить, пользуясь встроенными функциями EXCEL, в частности, функциями МУМНОЖ, МОБР, ТРАНСП.

В заключение отмечаем, что использованная выборка представляет собой фрагмент генеральной выборки, элементы которой связаны между собой случайной зависимостью

$$X_1 = 10 - 2\,X_2 + 3\,X_3.$$

Если из полученного уравнения регрессии общего вида выразить X_1, то оно примет вид:

$$X_1 = 9,074581 - 1,994655\,X_2 + 3,35194\,X_3.$$

С другой стороны, базовым методом наименьших квадратов получено уравнение регрессии

$$X_1 = 8,884244 - 1,953421\,X_2 + 3,397745\,X_3.$$

Легко видеть, что погрешности оценок всех коэффициентов уравнения регрессии общего вида (-9,25%; 0,267%; 11,73%) меньше соответствующих погрешностей (-11,16%; 2,33%; 13,26%) уравнения регрессии, полученного методом наименьших квадратов. Это подтверждает пригодность метода получения уравнений регрессии общего вида для применения в экономических исследованиях и для моделирования экономических объектов и процессов. Так же могут использоваться и другие матричные методы линейной множественной регрессии.

Литература

1. Кремер Н.Ш., Путко Б.А. Эконометрика: Учебник для вузов / Под ред. проф. Н.Ш. Кремера. – М.: ЮНИТИ-ДАНА, 2002. – 311 с.
2. Ивашнев Л.И. Методы регрессии в экономической математике: Монография. – М.: Изд-во МГОУ, 2005.
3. Ивашнев Л.И. Методы и модели в экономике: Учеб. пособие. – М.: Издательский дом «Лидер-М», 2011. – 328 с.