Моделирование и анализ данных
2024. Том 14. № 1. С. 121–134
DOI: https://doi.org/10.17759/mda.2024140108
ISSN: 2219-3758 (печатный)
ISSN: 2311-9454 (online)

Modelling and Data Analysis
2024. Vol. 14, no. 1, pp. 121–134
DOI: https://doi.org/10.17759/mda.2024140108
ISSN: 2219-3758 (print)
ISSN: 2311-9454 (online)

УДК 519.862.6

Решение оптимизационной задачи оценивания моделей полносвязной линейной регрессии

Базилевский М.П.*
Иркутский государственный университет путей сообщения
(ФГБОУ ВО ИрГУПС), г. Иркутск, Российская Федерация
ORCID: https://orcid.org/0000-0002-3253-5697
e-mail: mik2178@yandex.ru

Статья посвящена проблеме оценивания моделей полносвязной линейной регрессии методом максимального правдоподобия. Ранее для этого был разработан специальный численный метод, основанный на решении нелинейной системы методом простых итераций. При этом не исследовались вопросы выбора начальных приближений и выполнения достаточных условий сходимости. В данной статье предложен новый способ решения оптимизационной задачи оценивания полносвязных регрессий, схожий с методом оценивания ортогональных регрессий. Доказано, что при равных дисперсиях ошибок взаимосвязанных переменных оценки b-параметров полносвязной регрессии равны компонентам собственного вектора, соответствующего наименьшему собственному числу обратной ковариационной матрицы. А если отношения дисперсий ошибок переменных равны отношениям дисперсий переменных, то оценки b-параметров равны компонентам собственного вектора, соответствующего наименьшему собственному числу обратной корреляционной матрицы, умноженным на конкретные отношения среднеквадратических отклонений переменных. Проведен численный эксперимент, подтверждающий корректность разработанного математического аппарата. Предложенный способ решения оптимизационной задачи оценивания полносвязных регрессий может эффективно применяться при решении задач построения множественно-полносвязных линейных регрессий.

Ключевые слова: модель полносвязной линейной регрессии, метод максимального правдоподобия, численный метод, оптимизация, ортогональная регрессия, корреляционная матрица, собственный вектор, собственное значение.

Для цитаты:
Базилевский М.П. Решение оптимизационной задачи оценивания моделей полносвязной линейной регрессии // Моделирование и анализ данных. 2024. Том 14. № 1. С. 121–134. DOI: https://doi.org/10.17759/mda.2024140108

***Базилевский Михаил Павлович,** кандидат технических наук, доцент кафедры математики, Иркутский государственный университет путей сообщения (ФГБОУ ВО ИрГУПС), г. Иркутск, Российская Федерация, ORCID: https://orcid.org/0000-0002-3253-5697, e-mail: mik2178@yandex.ru

1. ВВЕДЕНИЕ

В регрессионном анализе [1] простейшей принято считать модель множественной линейной регрессии:

$$y_i = \alpha_0 + \sum_{j=1}^{m} \alpha_j x_{ij} + \varepsilon_i \,, \; i = \overline{1,n} \,, \tag{1}$$

где y_i, $i=\overline{1,n}$ – значения зависимой (объясняемой) переменной y, содержащие ошибки ε_i, $i=\overline{1,n}$; x_{ij}, $i=\overline{1,n}$, $j=\overline{1,m}$ – значения m независимых (объясняющих) переменных, не содержащие ошибок; α_0, α_1, …, α_m – неизвестные параметры; n – объем выборки.

Технология оценивания и построения линейных регрессий (1) на сегодняшний день в высшей степени развита и подробно описана в многочисленной литературе [1–3]. Но значения независимых переменных даже при использовании современных технических средств также часто оказываются не вполне правильно измеренными. Поэтому возникает задача оценивания моделей с ошибками и в объясняющих переменных. В зарубежной литературе такие модели известны как «errors-in-variables models» (EIV-модели) или «measurement error models». Обзоры методов их оценивания можно найти в работах [4, 5].

Среди EIV-моделей высокое прикладное значение имеет регрессия Деминга [6], содержащая всего две переменных y и x, которая в клинических лабораториях служит прекрасным инструментом для численного сопоставления новых измерительных методов с существующими (см., например, [7,8]). Описание методов оценивания EIV-моделей с несколькими переменными можно найти в работе [9]. Такие модели имеют следующую спецификацию:

$$y_i = y_i^* + \varepsilon_i^{(y)} \,, \; i = \overline{1,n} \,, \tag{2}$$

$$x_{ij} = x_{ij}^* + \varepsilon_i^{(x_j)} \,, \; i = \overline{1,n} \,, \; j = \overline{1,m} \,, \tag{3}$$

$$y_i^* = a_1 x_{i1}^* + a_2 x_{i2}^* + \ldots + a_m x_{im}^* \,, \; i = \overline{1,n} \,, \tag{4}$$

где y_i^*, x_{i1}^*, …, x_{im}^*, $i=\overline{1,n}$ – неизвестные истинные значения переменных; $\varepsilon_i^{(y)}$, $\varepsilon_i^{(x_1)}$, …, $\varepsilon_i^{(x_m)}$, $i=\overline{1,n}$ – случайные ошибки переменных; a_1, …, a_m – неизвестные параметры. Если дисперсии ошибок переменных одинаковы, то задача оценивания модели (2) – (4) состоит в минимизации суммы квадратов расстояний от точек выборки до плоскости. Такая регрессия была названа ортогональной [9]. Её оценивание осуществляется по следующему алгоритму.

Все наблюдения объединяются в одну матрицу U. В ней осуществляется центрирование всех столбцов.

Для матрицы $U^T U$ находятся характеристические векторы и числа.

Выбирается минимальное характеристическое число и соответствующий минимальный характеристический вектор.

Как отмечено в [10], первой работой, в которой была рассмотрена ортогональная регрессия, была работа К. Пирсона [11].

Существуют и другие известные методы оценивания EIV-моделей (2) – (4) при равных дисперсиях ошибок переменных, например, метод наименьших полных квадратов (Total Least Squares, TLS), предложенный в [12].

В работах [13, 14] автором была предложена модель полносвязной линейной регрессии (МПЛР), обобщающая регрессию Деминга. В МПЛР все истинные переменные связаны между собой линейными функциональными зависимостями:

$$x_{ij}^* = a_j + b_j x_{im}^* ,\ i = \overline{1, n} ,\ j = \overline{1, m-1} , \tag{5}$$

где a_j, b_j, $j = \overline{1, m-1}$ – неизвестные параметры.

Совокупность уравнений (3) и (5) представляет собой МПЛР.

Для приближенного оценивания МПЛР (3), (5) с помощью минимизации сумм квадратов ошибок в [14, 15] был предложен специальный численный метод, основанный на решении системы нелинейных уравнений методом простых итераций. При этом не исследовались проблемы, связанные с выбором начального приближения и выполнением достаточного признака сходимости метода. Поэтому в некоторых случаях он не обеспечивает требуемую точность, а может и вовсе расходиться.

Целью данной работы является разработка на основе метода Пирсона для ортогональной регрессии (2) – (4) метода решения оптимизационной задачи оценивания неизвестных параметров МПЛР (3), (5).

2. ПОСТАНОВКА ЗАДАЧИ

Пусть в МПЛР (3), (5) $\varepsilon_i^{(x_j)} \sim N\left(0, \sigma^2_{\varepsilon^{(x_j)}}\right)$, $i = \overline{1, n}$, $j = \overline{1, m}$, т.е. ошибки имеют нормальные законы распределения с нулевыми математическими ожиданиями и постоянными дисперсиями. В [16] показано, что нахождение оценок МПЛР методом максимального правдоподобия требует решения следующей оптимизационной задачи:

$$\sum_{i=1}^{n} \sum_{j=1}^{m-1} \lambda_j \left(x_{ij} - a_j - b_j x_{im}^*\right)^2 + \sum_{i=1}^{n} \left(x_{im} - x_{im}^*\right)^2 \to \min , \tag{6}$$

где $\lambda_j = \dfrac{\sigma^2_{\varepsilon^{(x_m)}}}{\sigma^2_{\varepsilon^{(x_j)}}}$, $j = \overline{1, m-1}$ – отношения дисперсий ошибок переменных.

Применение необходимых условий экстремума функции (6) приводит к нелинейной системе [17]:

$$b_p \cdot D_{x_m^*} = K_{x_p x_m^*} ,\ p = \overline{1, m-1} , \tag{7}$$

где, $D_{x_m^*} = \left(D_{x_m} + \sum_{j=1}^{m-1} \lambda_j^2 b_j^2 D_{x_j} + 2\sum_{j_1=1}^{m-2} \sum_{j_2=j_1+1}^{m-1} \lambda_{j_1} \lambda_{j_2} b_{j_1} b_{j_2} K_{x_{j_1} x_{j_2}} + 2\sum_{j=1}^{m-1} \lambda_j b_j K_{x_j x_m} \right)\left(1+\sum_{j=1}^{m-1} \lambda_j b_j^2\right)^{-2}$

$K_{x_p x_m^*} = \left(1+\sum_{j=1}^{m-1} \lambda_j b_j^2\right)^{-1} \left(\sum_{j=1}^{m-1} \lambda_j b_j K_{x_j x_p} + K_{x_m x_p}\right)$, символом D обозначены дисперсии переменных, K – ковариации.

Для нахождения решения нелинейной системы (7), при котором функция (6) достигает глобального минимума, в [15] был предложен специальный численный метод. К сожалению, при решении некоторых прикладных задач этот метод не обеспечивал сходимость.

В [15] для оценки общего качества аппроксимации МПЛР предложен аддитивный коэффициент детерминации:

$$R_{\text{add}}^2 = \sum_{j=1}^{m} R_{x_j}^2 ,$$

где $R_{x_j}^2$ – коэффициент детерминации для взаимосвязанной переменной x_j.

И в [15] показано, что оптимизационная задача $R_{\text{add}}^2 \to \max$, эквивалентна задаче (6) при условиях

$$\lambda_j = \frac{D_{x_m}}{D_{x_j}}, \; j = \overline{1, m-1} . \tag{8}$$

Решение задачи (6) при условиях (8) имеет высокое прикладное значение, поскольку на её основе строятся множественно-полносвязные линейные регрессии [18]. Поэтому была поставлена следующая *задача*: для оценивания МПЛР (3), (5) разработать на основе метода Пирсона для ортогональной регрессии метод решения оптимизационной задачи (6) в условиях (8).

3. РЕШЕНИЕ ЗАДАЧИ

Введем ковариационную матрицу V и корреляционную матрицу R_{xx}:

$$V = \begin{pmatrix} D_{x_1} & K_{x_1x_2} & \dots & K_{x_1x_m} \\ K_{x_1x_2} & D_{x_2} & \dots & K_{x_2x_m} \\ \dots & \dots & \dots & \dots \\ K_{x_1x_m} & K_{x_2x_m} & \dots & D_{x_m} \end{pmatrix}, \; R_{xx} = \begin{pmatrix} 1 & r_{x_1x_2} & \dots & r_{x_1x_m} \\ r_{x_1x_2} & 1 & \dots & r_{x_2x_m} \\ \dots & \dots & \dots & \dots \\ r_{x_1x_m} & r_{x_2x_m} & \dots & 1 \end{pmatrix}.$$

Введем нормированные (стандартизованные) переменные, значения которых находятся по формулам:

$$x_{i1}^{\bullet} = \frac{x_{i1} - \overline{x_1}}{\sigma_{x_1}}, \; \dots, \; x_{im}^{\bullet} = \frac{x_{im} - \overline{x_m}}{\sigma_{x_m}},$$

где $\overline{x_1}$, …, $\overline{x_m}$ – средние значения переменных;
σ_{x_1}, …, σ_{x_m} – среднеквадратические отклонения переменных;
$x_1^{\bullet}$, …, $x_m^{\bullet}$ – стандартизованные переменные, для которых среднее значение равно 0, а среднеквадратическое отклонение равно 1.

Теорема. Если МПЛР (3), (5) оценивается по критерию (6), то вектор оптимальных оценок её параметров $\begin{pmatrix} b_1 & \dots & b_{m-1} & 1 \end{pmatrix}$:

- при $\lambda_j = 1$, $j = \overline{1, m-1}$, равен собственному вектору $\begin{pmatrix} \tilde{b}_1 & \dots & \tilde{b}_{m-1} & 1 \end{pmatrix}$, соответствующему наименьшему собственному числу матрицы V^{-1};
- при условиях (8), равен вектору $\begin{pmatrix} \tilde{\beta}_1 \dfrac{\sigma_{x_1}}{\sigma_{x_m}} & \dots & \tilde{\beta}_{m-1} \dfrac{\sigma_{x_{m-1}}}{\sigma_{x_m}} & 1 \end{pmatrix}$, где $\begin{pmatrix} \tilde{\beta}_1 & \dots & \tilde{\beta}_{m-1} & 1 \end{pmatrix}$ – собственный вектор, соответствующий наименьшему собственному числу матрицы R_{xx}^{-1}.

Доказательство.

В работе [17] установлено, из нелинейной системы (7) вытекают соотношения

$$\lambda_p = \frac{\Delta_p}{b_p \cdot \Delta_m}, \quad p = \overline{1, m-1}, \tag{9}$$

где Δ_j – определитель, полученный из определителя матрицы V путем замены в нём j-го столбца на столбец $\begin{pmatrix} b_1 & \dots & b_{m-1} & 1 \end{pmatrix}^T$.

Составим систему с неизвестными z_1, z_2, …, z_m:

$$\begin{pmatrix} D_{x_1} & K_{x_1 x_2} & \dots & K_{x_1 x_m} \\ K_{x_1 x_2} & D_{x_2} & \dots & K_{x_2 x_m} \\ \dots & \dots & \dots & \dots \\ K_{x_1 x_m} & K_{x_2 x_m} & \dots & D_{x_m} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \dots \\ z_m \end{pmatrix} = \begin{pmatrix} b_1 \\ \dots \\ b_{m-1} \\ 1 \end{pmatrix}. \tag{10}$$

Пусть $z_1 = \chi \cdot \lambda_1 b_1$, $z_2 = \chi \cdot \lambda_2 b_2$, …, $z_{m-1} = \chi \cdot \lambda_{m-1} b_{m-1}$, $z_m = \chi$, где χ – отличное от нуля неизвестное число.

Тогда система (10) примет вид:

$$\begin{pmatrix} D_{x_1} & K_{x_1 x_2} & \dots & K_{x_1 x_m} \\ K_{x_1 x_2} & D_{x_2} & \dots & K_{x_2 x_m} \\ \dots & \dots & \dots & \dots \\ K_{x_1 x_m} & K_{x_2 x_m} & \dots & D_{x_m} \end{pmatrix} \begin{pmatrix} \chi \cdot \lambda_1 b_1 \\ \dots \\ \chi \cdot \lambda_{m-1} b_{m-1} \\ \chi \end{pmatrix} = \begin{pmatrix} b_1 \\ \dots \\ b_{m-1} \\ 1 \end{pmatrix}. \tag{11}$$

Запишем для системы (11) формулы Крамера:

$$\chi \cdot \lambda_1 b_1 = \frac{\Delta_1}{|V|}, \dots, \chi \cdot \lambda_{m-1} b_{m-1} = \frac{\Delta_{m-1}}{|V|}, \ \chi = \frac{\Delta_m}{|V|}. \tag{12}$$

Обозначим любое решение системы (11) $\chi = \tilde{\chi}$, $b_j = \tilde{b}_j$, $j = \overline{1, m-1}$. Тогда формулы (12) примут вид:

$$\tilde{\chi} \cdot \lambda_1 \tilde{b}_1 = \frac{\Delta_1}{|V|}, \ \ldots, \ \tilde{\chi} \cdot \lambda_{m-1} \tilde{b}_{m-1} = \frac{\Delta_{m-1}}{|V|}, \ \tilde{\chi} = \frac{\Delta_m}{|V|}.$$

Поделив первые $(m-1)$ уравнений на последнее, получим

$$\lambda_1 \tilde{b}_1 = \frac{\Delta_1}{\Delta_m}, \ \ldots, \ \lambda_{m-1} \tilde{b}_{m-1} = \frac{\Delta_{m-1}}{\Delta_m},$$

т.е. все решения системы (11) также являются решениями систем (7) и (9). Далее будем искать решения системы (11).

Домножим выражение (11) на V^{-1} слева, учитывая, что $|V| \neq 0$. Получим

$$\begin{pmatrix} D_{x_1} & K_{x_1 x_2} & \ldots & K_{x_1 x_m} \\ K_{x_1 x_2} & D_{x_2} & \ldots & K_{x_2 x_m} \\ \ldots & \ldots & \ldots & \ldots \\ K_{x_1 x_m} & K_{x_2 x_m} & \ldots & D_{x_m} \end{pmatrix}^{-1} \begin{pmatrix} b_1 \\ \ldots \\ b_{m-1} \\ 1 \end{pmatrix} = \chi \begin{pmatrix} \lambda_1 b_1 \\ \ldots \\ \lambda_{m-1} b_{m-1} \\ 1 \end{pmatrix}. \tag{13}$$

Выражение (13) означает, что матрица линейного оператора V^{-1} преобразует вектор $(b_1 \quad \ldots \quad b_{m-1} \quad 1)^T$ в вектор $\chi(\lambda_1 b_1 \quad \ldots \quad \lambda_{m-1} b_{m-1} \quad 1)^T$.

Если $\lambda_1 = \lambda_2 = \ldots = \lambda_{m-1} = 1$ (дисперсии ошибок одинаковы), то выражение (13) принимает вид:

$$\begin{pmatrix} D_{x_1} & K_{x_1 x_2} & \ldots & K_{x_1 x_m} \\ K_{x_1 x_2} & D_{x_2} & \ldots & K_{x_2 x_m} \\ \ldots & \ldots & \ldots & \ldots \\ K_{x_1 x_m} & K_{x_2 x_m} & \ldots & D_{x_m} \end{pmatrix}^{-1} \begin{pmatrix} b_1 \\ \ldots \\ b_{m-1} \\ 1 \end{pmatrix} = \chi \begin{pmatrix} b_1 \\ \ldots \\ b_{m-1} \\ 1 \end{pmatrix}. \tag{14}$$

Из (14) следует, что при $\lambda_1 = \lambda_2 = \ldots = \lambda_{m-1} = 1$ вектор $(b_1 \quad \ldots \quad b_{m-1} \quad 1)^T$ – собственный вектор матрицы V^{-1}, а χ – соответствующее ему собственное значение.

Матрица V^{-1} может иметь до m собственных значений. Определим, какое из них обеспечивает глобальный минимум задачи (6).

Известно, что, с одной стороны, коэффициенты детерминации взаимосвязанных переменных находятся по формулам $R_{x_j}^2 = 1 - \frac{\sum_{i=1}^{n} \varepsilon_i^{(x_j)}}{D_{x_j}}$, $j = \overline{1, m}$. Тогда задачу (6) можно представить в виде

$$\sum_{j=1}^{m-1} \lambda_j D_{x_j} \left(1 - R_{x_j}^2\right) + D_{x_m} \left(1 - R_{x_m}^2\right) \to \min,$$

что равносильно следующей постановке:

$$\sum_{j=1}^{m-1} \lambda_j D_{x_j} R_{x_j}^2 + D_{x_m} R_{x_m}^2 \to \max . \tag{15}$$

С другой стороны, коэффициенты детерминации взаимосвязанных переменных можно найти по формулам $R_{x_j}^2 = \dfrac{K_{x_j x_m^*}^2}{D_{x_j} D_{x_m^*}}$, $j = \overline{1, m}$, или, с использованием (7), по формулам

$$R_{x_j}^2 = \frac{b_j^2 D_{x_m^*}}{D_{x_j}},\ j = \overline{1, m-1},\ R_{x_m}^2 = \frac{D_{x_m^*}}{D_{x_m}} . \tag{16}$$

Подставляя (16) в (15), получим:

$$D_{x_m^*} \left(1 + \sum_{j=1}^{m-1} \lambda_j b_j^2 \right) \to \max . \tag{17}$$

В [17] установлено, что $D_{x_m^*} = \left(1 + \sum_{j=1}^{m-1} \lambda_j b_j^2 \right)^{-1} \cdot \dfrac{|V|}{\Delta_m}$. Тогда задача (17) примет вид $\dfrac{|V|}{\Delta_m} \to \max$. Как следует из (12), $\chi = \dfrac{\Delta_m}{|V|}$, поэтому задаче (17) равносильна задача

$$\chi^{-1} \to \max . \tag{18}$$

Выражение (18) означает, что в точке глобального минимума функции (6) величина χ должна принимать наименьшее значение. Поэтому при $\lambda_1 = \lambda_2 = \ldots = \lambda_{m-1} = 1$ нужно выбрать тот собственный вектор матрицы V^{-1}, который соответствует наименьшему собственному значению χ.

Вторая часть теоремы доказывается аналогично. Для этого запишем МПЛР (3), (5) в стандартизованном виде [19]

$$x_{ij}^{\bullet} = x_{ij}^{\bullet *} + \varepsilon_i^{\left(x_j^{\bullet}\right)},\ i = \overline{1, n},\ j = \overline{1, m}, \tag{19}$$

$$x_{ij}^{\bullet *} = \beta_j x_{im}^{\bullet *},\ i = \overline{1, n},\ j = \overline{1, m-1}, \tag{20}$$

где $x_{i1}^{\bullet *}$, …, $x_{im}^{\bullet *}$, $i = \overline{1, n}$ – неизвестные истинные значения стандартизованных переменных; β_1, …, β_{m-1} – стандартизованные коэффициенты (бета-коэффициенты); $\varepsilon_i^{\left(x_j^{\bullet}\right)}$, $i = \overline{1, n}$, $j = \overline{1, m}$ – случайные ошибки переменных.

Тогда нахождение оценок МПЛР (19), (20) методом максимального правдоподобия требует решения следующей оптимизационной задачи:

$$\sum_{i=1}^{n} \sum_{j=1}^{m-1} \lambda_j^{\bullet} \left(x_{ij}^{\bullet} - \beta_j x_{im}^{\bullet *} \right)^2 + \sum_{i=1}^{n} \left(x_{im}^{\bullet} - x_{im}^{\bullet *} \right)^2 \to \min , \tag{21}$$

где $\lambda_j^\bullet = \dfrac{\sigma^2_{\varepsilon\left(x_m^\bullet\right)}}{\sigma^2_{\varepsilon\left(x_j^\bullet\right)}}$, $j=\overline{1,m-1}$. Причем, поскольку $\sigma^2_{\varepsilon\left(x_j^\bullet\right)} = \sigma^2_{\varepsilon(x_j)} / D_{x_j}$, $j=\overline{1,m}$, то имеют место следующие соотношения:

$$\lambda_j^\bullet = \frac{\sigma^2_{\varepsilon\left(x_m^\bullet\right)}}{\sigma^2_{\varepsilon\left(x_j^\bullet\right)}} = \frac{\sigma^2_{\varepsilon(x_m)}}{\sigma^2_{\varepsilon(x_j)}} \cdot \frac{D_{x_j}}{D_{x_m}} = \lambda_j \cdot \frac{D_{x_j}}{D_{x_m}}, \quad j=\overline{1,m}. \tag{22}$$

Для задачи (21) система (13) имеет вид

$$\begin{pmatrix} 1 & r_{x_1x_2} & \dots & r_{x_1x_m} \\ r_{x_1x_2} & 1 & \dots & r_{x_2x_m} \\ \dots & \dots & \dots & \dots \\ r_{x_1x_m} & r_{x_2x_m} & \dots & 1 \end{pmatrix}^{-1} \begin{pmatrix} \beta_1 \\ \dots \\ \beta_{m-1} \\ 1 \end{pmatrix} = \chi \begin{pmatrix} \lambda_1^\bullet \beta_1 \\ \dots \\ \lambda_{m-1}^\bullet \beta_{m-1} \\ 1 \end{pmatrix}. \tag{23}$$

Из (23) следует, что при $\lambda_1^\bullet = \lambda_2^\bullet = \ldots = \lambda_{m-1}^\bullet = 1$ вектор $\begin{pmatrix} \beta_1 & \dots & \beta_{m-1} & 1 \end{pmatrix}^T$ – собственный вектор матрицы R_{xx}^{-1}, а χ – соответствующее ему собственное значение. Аналогично, глобальный минимум функции (21) обеспечивает наименьшее собственное значение.

Из соотношений (22) следует, что задачи поиска глобальных минимумов функции (21) для МПЛР (19), (20) при $\lambda_1^\bullet = \lambda_2^\bullet = \ldots = \lambda_{m-1}^\bullet = 1$ и функции (6) для МПЛР (3), (5) при условиях (8) эквивалентны.

Оценки МПЛР (3), (5) связаны с оценками МПЛР (19), (20) соотношениями $b_j = \beta_j \dfrac{\sigma_{x_j}}{\sigma_{x_m}}$, $j=\overline{1,m-1}$. С помощью этих формул, зная собственный вектор матрицы R_{xx}^{-1}, соответствующий наименьшему собственному значению, находятся оценки МПЛР (3), (5) в условиях (8).

Теорема доказана.

Алгоритмы поиска собственных значений и собственных векторов симметричных матриц широко известны (см., например, [20]).

4. ЧИСЛЕННЫЙ ЭКСПЕРИМЕНТ

Для демонстрации справедливости доказанной теоремы на практике была взята задача оценивания МПЛР, решенная в работе [17] при условиях (8). В качестве взаимосвязанных переменных выступают: x_1 – население в трудоспособном возрасте (%), x_2 – численность рабочей силы (тыс. человек), x_3 – численность пенсионеров (тыс. человек). Ковариационная и корреляционная матрицы для этих переменных имеют вид:

$$V=\begin{pmatrix} 9,096 & 121,234 & -93,284 \\ 121,234 & 2403,759 & -1451,256 \\ -93,284 & -1451,256 & 1176,521 \end{pmatrix},\ R_{xx}=\begin{pmatrix} 1 & 0,8198 & -0,9017 \\ 0,8198 & 1 & -0,8629 \\ -0,9017 & -0,8629 & 1 \end{pmatrix}.$$

Сначала оптимизационная задача (6) решалась с помощью предложенного в работе [15] численного метода. Для этого был разработан специальный скрипт на языке hansl эконометрического пакета Gretl. В качестве начальных приближений для параметров b_1 и b_2 МПЛР были взяты МНК-оценки парных регрессий x_1 от x_3 и x_2 от x_3:

$$b_1^0=-0,079288,\ b_2^0=-1,233514.$$

Точность сходимости была установлена 0,0000001.

В результате численного оценивания МПЛР при условиях $\lambda_1=\lambda_2=1$ за 7 итераций были получены оценки

$$\tilde{b}_1=-0,082087,\ \tilde{b}_2=-1,508027,$$

а при условиях (8) $\lambda_1=129,3403$, $\lambda_2=0,4894$ за 31 итерацию – оценки

$$\tilde{b}_1=-0,08653069,\ \tilde{b}_2=-1,38485469.$$

Затем решение оптимизационной задачи (6) осуществлялось предложенным в данной статье методом. Обратная матрица для матрицы V имеет вид:

$$V^{-1}=\begin{pmatrix} 0,6105 & -0,0061 & 0,0408 \\ -0,0061 & 0,0016 & 0,0015 \\ 0,0408 & 0,0016 & 0,0060 \end{pmatrix}.$$

Для этой матрицы были найдены следующие собственные числа и соответствующие им собственные векторы:

- для числа $\chi_1=0,000296$ вектор $\begin{pmatrix}-0,082087 & -1,508027 & 1\end{pmatrix}^T$;
- для числа $\chi_2=0,00465$ вектор $\begin{pmatrix}-0,060655 & 0,66642 & 1\end{pmatrix}^T$;
- для числа $\chi_3=0,61331$ вектор $\begin{pmatrix}14,875743 & -0,14662 & 1\end{pmatrix}^T$.

Тогда при $\lambda_1=\lambda_2=1$ оптимальные оценки параметров b_1 и b_2 МПЛР равны компонентам собственного вектора, соответствующего наименьшему собственному числу $\chi_1=0,000296$ матрицы V^{-1}, т.е. $\tilde{b}_1=-0,082087$, $\tilde{b}_2=-1,508027$.

Обратная матрица для матрицы R_{xx} имеет вид:

$$R_{xx}^{-1}=\begin{pmatrix} 5,5533 & -0,9073 & 4,2246 \\ -0,9073 & 4,0656 & 2,6903 \\ 4,2246 & 2,6903 & 7,1312 \end{pmatrix}.$$

Найденные для этой матрицы собственные числа и соответствующие им собственные векторы оказались следующими:

- для числа $\chi_1 = 0,367176$ вектор $\left(-0,984094 \quad -0,968857 \quad 1\right)^T$;
- для числа $\chi_2 = 5,389291$ вектор $\left(-3,028804 \quad 4,108581 \quad 1\right)^T$;
- для числа $\chi_3 = 10,993702$ вектор $\left(0,727665 \quad 0,293035 \quad 1\right)^T$.

Наименьшее собственное число $\chi_1 = 0,367176$, поэтому оптимальные оценки параметров β_1 и β_2 равны компонентам соответствующего собственного вектора матрицы R_{xx}^{-1}, т.е. $\tilde{\beta}_1 = -0,984094$, $\tilde{\beta}_2 = -0,968857$. Тогда оптимальные оценки параметров b_1 и b_2 МПЛР при условиях (8) $\lambda_1 = 129,3403$, $\lambda_2 = 0,4894$ находятся по формулам:

$$\tilde{b}_1 = \tilde{\beta}_1 \frac{\sigma_{x_1}}{\sigma_{x_3}} = -0,08653,$$

$$\tilde{b}_2 = \tilde{\beta}_2 \frac{\sigma_{x_2}}{\sigma_{x_3}} = -1,38485.$$

Как видно, решения оптимизационной задачи (6) оценивания МПЛР (3), (5) разработанным ранее численным методом и решения предложенным в данной статье методом совпадают.

5. ЗАКЛЮЧЕНИЕ

В работе найдено решение оптимизационной задачи оценивания МПЛР (3), (5) по критерию (6). Доказано, что при $\lambda_j = 1$, $j = \overline{1, m-1}$, оценки параметров b_1, b_2, …, b_{m-1} равны компонентам собственного вектора, соответствующего наименьшему собственному числу матрицы V^{-1}, а при условиях (8) – компонентам собственного вектора, соответствующего наименьшему собственному числу матрицы R_{xx}^{-1}, умноженным на числа $\dfrac{\sigma_{x_j}}{\sigma_{x_m}}$, $j = \overline{1, m-1}$. Корректность доказанной теоремы продемонстрирована на ранее решенном примере оценивания МПЛР методом простых итераций. Новым способом были получены те же самые оценки. Достоинство нового способа решения задачи оценивания МПЛР состоит в том, что нет необходимости в выборе начального приближения и в проверке выполнения достаточного признака сходимости численного метода. К тому же на сегодняшний день существуют эффективные алгоритмы поиска собственных значений и собственных векторов симметричных матриц. К недостатку можно отнести то, что новый метод справедлив только тогда, когда определители матриц V и R_{xx} отличны от нуля, т.е. при отсутствии полной коллинеарности между взаимосвязанными переменными.

Литература

1. *Montgomery D.C., Peck E.A., Vining G.G.* Introduction to linear regression analysis. John Wiley & Sons, 2021.

2. *Тырсин А.Н.* Алгоритмы спуска по узловым прямым в задаче оценивания регрессионных уравнений методом наименьших модулей // Заводская лаборатория. Диагностика материалов. 2021. Т. 87. № 5. С. 68–75.
3. *Liu Z., Yang Y.* Least absolute deviations estimation for uncertain regression with imprecise observations // Fuzzy Optimization and Decision Making. 2020. Vol. 19. P. 33–52.
4. *Кендалл М., Стьюарт А.* Статистические выводы и связи. М.: Наука, 1973. 899 с.
5. *Gillard J.* An overview of linear structural models in errors in variables regression // REVSTAT: Statistical Journal. 2010. Vol. 8. № 1. P. 57–80. DOI: 10.57805/revstat.v8i1.90
6. *Deming W.E.* Statistical adjustment of data. New York, Wiley, 2011. 288 p.
7. *Wickremsinhe E., et al.* Standard venipuncture vs a capillary blood collection device for the prospective determination of abnormal liver chemistry // The Journal of Applied Laboratory Medicine. 2023. Vol. 8. № 3. P. 535–550. DOI: 10.1093/jalm/jfac127
8. *Tantisaranon P., Dumkengkhachornwong K., Hnoonual A.* Influence of reduced centrifugation time on clinical chemistry analytes and literature review // Turkish Journal of Biochemistry. 2023. Vol. 48. № 4. P. 376–387. DOI: 10.1515/tjb-2022-0211
9. *Демиденко Е.З.* Линейная и нелинейная регрессии. М.: Финансы и статистика, 1981. 304 с.
10. *Ломов А.А.* Идентифицируемость и вариационные оценки параметров дискретных стационарных линейных динамических систем: дис. … д-ра физ.-мат. наук: 05.13.01. Новосибирск, 2011. 356 с.
11. *Pearson K.* On lines and planes of closet fit to systems of points in space // Phil. Mag. 1901. Vol. 6. № 2. P. 559–572.
12. *Golub G.H., Van Loan C.F.* An analysis of the total least squares problem // SIAM Journal on Numerical Analysis. 1980. Vol. 17. № 6. P. 883–893.
13. *Базилевский М.П.* Исследование двухфакторной модели полносвязной линейной регрессии // Моделирование, оптимизация и информационные технологии. 2019. Т. 7. № 2 (25). С. 80–96.
14. *Базилевский М.П.* Методы построения регрессионных моделей с ошибками во всех переменных. Иркутск: ИрГУПС, 2019. 208 с.
15. *Базилевский М.П.* Метод выпрямления искаженных из-за мультиколлинеарности коэффициентов в регрессионных моделях // Информатика и её применения. 2021. Т. 15. № 2. С. 60–65. DOI:10.14357/19922264210209
16. *Базилевский М.П.* Интерпретация оценок параметров моделей полносвязной линейной регрессии // International Journal of Open Information Technologies. 2023. Т. 11. № 10. С. 21–25.
17. *Базилевский М.П.* Идентификация областей возможных оценок параметров моделей полносвязной линейной регрессии // Моделирование и анализ данных. 2023. Т. 13. № 3. С. 52–65. DOI:10.17759/mda.2023130304
18. *Базилевский М.П.* Математическое моделирование с помощью множественно-полносвязных линейных регрессий // System Analysis & Mathematical Modeling. 2023. Т. 5. № 4. С. 457–475.
19. *Фёрстер Э., Рёнц Б.* Методы корреляционного и регрессионного анализа. М.: Финансы и статистика, 1983. 303 с.
20. *Имомов А.* Алгоритмы методов Крылова и Данилевского для собственных значений и векторов матриц // PEDAGOG. 2023. Т. 6. № 2. С. 104–111.

Solving an Optimization Problem for Estimating Fully Connected Linear Regression Models

Mikhail P. Bazilevskiy*
Irkutsk State Transport University (ISTU), Irkutsk, Russia
ORCID: https://orcid.org/0000-0002-3253-5697
e-mail: mik2178@yandex.ru

This article is devoted to the problem of estimating fully connected linear regression models using the maximum likelihood method. Previously, a special numerical method was developed for this purpose, based on solving a nonlinear system using the method of simple iterations. At the same time, the issues of choosing initial approximations and fulfilling sufficient conditions for convergence were not studied. This article proposes a new method for solving the optimization problem of estimating fully connected regressions, similar to the method of estimating orthogonal regressions. It has been proven that, with equal error variances of interconnected variables, estimates of b-parameters of fully connected regression are equal to the components of the eigenvector corresponding to the smallest eigenvalue of the inverse covariance matrix. And if the ratios of the error variances of the variables are equal to the ratios of the variances of the variables, then the b-parameter estimates are equal to the components of the eigenvector corresponding to the smallest eigenvalue of the inverse correlation matrix, multiplied by the specific ratios of the standard deviations of the variables. A numerical experiment was carried out to confirm the correctness of the developed mathematical apparatus. The proposed method for solving the optimization problem of estimating fully connected regressions can be effectively used when solving problems of constructing multiple fully connected linear regressions.

Keywords: fully connected linear regression model, maximum likelihood estimation, numerical method, optimization, orthogonal regression, correlation matrix, eigenvector, eigenvalue.

For citation:
Bazilevskiy M.P. Solving an Optimization Problem for Estimating Fully Connected Linear Regression Models. *Modelirovanie i analiz dannykh = Modelling and Data Analysis,* 2024. Vol. 14, no. 1, pp. 121–134. DOI: https://doi.org/10.17759/mda.2024140108 (In Russ., abstr. in Engl.).

****Mikhail P. Bazilevskiy,*** PhD (Engineering), Associate Professor, Department of Mathematics, Irkutsk State Transport University (ISTU), Irkutsk, Russia, ORCID: https://orcid.org/0000-0002-3253-5697, e-mail: mik2178@yandex.ru

References

1. Montgomery D.C., Peck E.A., Vining G.G. *Introduction to linear regression analysis*. John Wiley & Sons, 2021.
2. Tyrsin A.N. Algoritmy spuska po uzlovym pryamym v zadache otsenivaniya regressionnykh uravneniy metodom naimen'shikh moduley [Algorithms for descent along nodal straight lines in the problem of estimating regression equations using the least absolute deviations method], *Industrial Laboratory. Diagnostics of Materials*, 2021, vol. 87, no. 5, pp. 68–75.
3. Liu Z., Yang Y. Least absolute deviations estimation for uncertain regression with imprecise observations, *Fuzzy Optimization and Decision Making*, 2020, vol. 19, pp. 33–52.
4. Kendall M., St'yuart A. *Statisticheskie vyvody i svyazi* [Statistical findings and connections]. Moscow, Science, 1973. 899 p.
5. Gillard J. An overview of linear structural models in errors in variables regression, *REVSTAT: Statistical Journal*, 2010, vol. 8, no. 1, pp. 57–80. DOI: 10.57805/revstat.v8i1.90
6. Deming W.E. *Statistical adjustment of data*. New York, Wiley, 2011. 288 p.
7. Wickremsinhe E., et al. Standard venipuncture vs a capillary blood collection device for the prospective determination of abnormal liver chemistry, *The Journal of Applied Laboratory Medicine*, 2023, vol. 8, no. 3, pp. 535–550. DOI: 10.1093/jalm/jfac127
8. Tantisaranon P., Dumkengkhachornwong K., Hnoonual A. Influence of reduced centrifugation time on clinical chemistry analytes and literature review, *Turkish Journal of Biochemistry*, 2023, vol. 48, no. 4, pp. 376–387. DOI: 10.1515/tjb-2022-0211
9. Demidenko E.Z. *Lineynaya i nelineynaya regressii* [Linear and nonlinear regressions]. Moscow, Finansy i statistika, 1981. 304 p.
10. Lomov A.A. *Identifitsiruemost' i variatsionnye otsenki parametrov diskretnykh statsionarnykh lineynykh dinamicheskikh sistem* [Identifiability and variational estimates of parameters of discrete stationary linear dynamic systems]. Novosibirsk, 2011. 356 p.
11. Pearson K. On lines and planes of closet fit to systems of points in space, *Phil. Mag*, 1901, vol. 6, no. 2, pp. 559–572.
12. Golub G.H., Van Loan C.F. An analysis of the total least squares problem, *SIAM Journal on Numerical Analysis*, 1980, vol. 17, no. 6, pp. 883–893.
13. Bazilevskiy M.P. Issledovanie dvukhfaktornoy modeli polnosvyaznoy lineynoy regressii [Investigation of a two-factor fully connected linear regression model], *Modeling, Optimization and Information Technology*, 2019, vol. 7, no. 2 (25), pp. 80–96.
14. Bazilevskiy M.P. *Metody postroeniya regressionnykh modeley s oshibkami vo vsekh peremennykh* [Methods for constructing errors-in-variables regression models]. Irkutsk, IrGUPS, 2019. 208 p.
15. Bazilevskiy M.P. Metod vypryamleniya iskazhennykh iz-za mul'tikollinearnosti koeffitsientov v regressionnykh modelyakh [Method of straightening distorted due to multicollinearity coefficients in regression models], *Informatics and Applications,* 2021, vol. 15, no. 2, pp. 60–65. DOI:10.14357/19922264210209
16. Bazilevskiy M.P. Interpretatsiya otsenok parametrov modeley polnosvyaznoy lineynoy regressii [Interpretation of Parameter Estimates for Fully Connected Linear Regression Models], *International Journal of Open Information Technologies*, 2023, vol. 11, no. 10, pp. 21–25.
17. Bazilevskiy M.P. Identifikatsiya oblastey vozmozhnykh otsenok parametrov modeley polnosvyaznoy lineynoy regressii [Identification of possible estimates areas for parameters of fully connected linear regression models], *Modeling and Data Analysis*, 2023, vol. 13, no. 3, pp. 52–65. DOI:10.17759/mda.2023130304
18. Bazilevskiy M.P. Matematicheskoe modelirovanie s pomoshch'yu mnozhestvenno-polnosvyaznykh lineynykh regressiy [Mathematical modeling using multiple and fully-connected linear regressions], *System Analysis & Mathematical Modeling*, 2023, vol. 5, no. 4, pp. 457–475.

19. Ferster E., Rents B. *Metody korrelyatsionnogo i regressionnogo analiza* [Methods of correlation and regression analysis]. Moscow, Finansy i statistika, 1983. 303 p.
20. Imomov A. Algoritmy metodov Krylova i Danilevskogo dlya sobstvennykh znacheniy i vektorov matrits [Algorithms of Krylov and Danilevsky methods for eigenvalues and matrix vectors], *PEDAGOG*, 2023, vol. 6, no. 2, pp. 104–111.

Получена 15.01.2024
Принята в печать 30.01.2024

Received 15.01.2024
Accepted 30.01.2024