МОДА. Задание 1.

Яковлев Д. М.
E-mail: st095998<at>student.spbu.ru

Мелас В. Б.
E-mail: vbmelas<at>yandex.ru

12 декабря 2024 г.

Содержание

1 Теоретический вопрос

Постановка. Рассмотрим регрессионную модель:

$$y_i = f(x_i)^{\mathrm{T}}\theta + \varepsilon_i, \quad i = 1, 2, \ldots, n.$$
$$f(x) = (f_1(x), f_2(x), \ldots, f_m(x))^{\mathrm{T}}, \quad \theta = (\theta_1, \theta_2, \ldots, \theta_m)^{\mathrm{T}}$$

со стандартными предположениями об ошибках ε_i ($\mathbf{E}\varepsilon_i = 0, \mathbf{E}\varepsilon_i^2 = \sigma^2, \mathbf{E}\varepsilon_i\varepsilon_j = 0$) и $x_1 = -1, x_2 = 1$. Перепишем систему (2) в матричном виде

$$\boldsymbol{Y} = \boldsymbol{X}\theta + \varepsilon,$$

где

$$\boldsymbol{Y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad \boldsymbol{X} = \begin{pmatrix} f_1(x_1) & f_2(x_1) & \vdots & f_m(x_1) \\ f_1(x_2) & f_2(x_2) & \vdots & f_m(x_2) \\ \vdots & \vdots & \vdots & \vdots \\ f_1(x_n) & f_2(x_n) & \vdots & f_m(x_n) \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix}.$$

Теорема 1 (Гаусса-Маркова в случае матриц полного ранга)**.** *Если матрица $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ — невырожденная, то оценка по методу наименьших квадратов*

$$\widehat{\theta} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} \tag{1}$$

является наилучшей оценкой в классе линейных несмещённых оценок, или

$$\forall z \in \mathbb{R}^m, \forall \widetilde{\theta} : \mathbf{E}\widetilde{\theta} = \theta,\ \mathbf{D} z^{\mathrm{T}}(\widehat{\theta} - \theta) \leqslant \mathbf{D} z^{\mathrm{T}}(\widetilde{\theta} - \theta).$$

Кроме того, ковариационная матрица оценки имеет вид:

$$\mathbf{D}\widehat{\theta} = \sigma^2(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}.$$

Краткое доказательство. Опишем схему доказательства:

1. Показываем, что оценка по МНК $\widehat{\theta}$ принадлежит классу линейных несмещённых оценок. Линейность
$$\widehat{\theta} = \boldsymbol{S}\boldsymbol{Y}$$
следует из представления оценки, а несмещённость проверяется прямой подстановкой в $\mathbf{E}\widehat{\theta}$

2. Определяем общий вид линейных несмещённых оценок $\widetilde{\theta} = \boldsymbol{A}\boldsymbol{Y}$. Необходимым и достаточным условием несмещённости $\widetilde{\theta}$ является
$$\boldsymbol{A}\boldsymbol{X} = \boldsymbol{I}.$$

3. Определяем вид ковариационной матрицы линейной несмещённой оценки $\widetilde{\theta} = \boldsymbol{A}\boldsymbol{Y}$, пользуясь свойствами линейной несмещённой оценки и прямой подстановкой:
$$\mathbf{D}\widetilde{\theta} = \sigma^2\boldsymbol{A}\boldsymbol{A}^{\mathrm{T}}.$$

4. Положим $\boldsymbol{S} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}$ и покажем, что
$$\mathbf{D}\widehat{\theta} \leqslant \mathbf{D}\widetilde{\theta}$$
для любой линейной несмещённой оценки $\widetilde{\theta} = \boldsymbol{A}\boldsymbol{Y}$. Так как неравенство рассматривается для матриц в смысле положительной определённости, пользуемся тем, что
$$\mathbf{0} \leqslant \boldsymbol{L}\boldsymbol{L}^{\mathrm{T}}$$
для произвольной матрицы $\boldsymbol{L}$ и подставляем $\boldsymbol{L} = \boldsymbol{A} - \boldsymbol{S}$.

□

2 Задача

Постановка. Пусть $f(x) = (1, x, x^2)^{\mathrm{T}}$, модель $f(x)^{\mathrm{T}}\theta$, $\theta = (\theta_1, \theta_2, \theta_3)^{\mathrm{T}}$. Построить несмещённую оценку для θ_2 непосредственно на основе определения по результатам измерений в точках -1 (результат y_1), 1 (результат y_2).

Решение.

2.1 Матричный подход

Рассмотрим регрессионную модель:

$$y_i = f(x_i)^{\mathrm{T}}\theta + \varepsilon_i = x_i^2\theta_3 + x_i\theta_2 + \theta_1 + \varepsilon_i, \quad i = 1, 2, \tag{2}$$

со стандартными предположениями об ошибках ε_i и $x_1 = -1, x_2 = 1$. Перепишем систему (2) в матричном виде

$$\boldsymbol{Y} = \boldsymbol{X}\theta + \varepsilon,$$

где

$$\boldsymbol{Y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}, \quad \boldsymbol{X} = \begin{pmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \end{pmatrix}.$$

Составим параметрическую функцию такую, что

$$\tau = \boldsymbol{T}\theta = \theta_2 \to \boldsymbol{T} = (0 \quad 1 \quad 0).$$

Вычисление матриц $\boldsymbol{X}, \boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y}, \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}, (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-}$:

$$\boldsymbol{X} = \begin{pmatrix} 1 & -1 & 1 \\ 1 & 1 & 1 \end{pmatrix}, \quad \boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} \begin{pmatrix} y_1 + y_2 \\ y_2 - y_1 \\ y_1 + y_2 \end{pmatrix},$$

$$\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} = \begin{pmatrix} 2 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix}.$$

Для матрицы $(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-}$

$$\begin{aligned}
\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} &= \begin{pmatrix} 2 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \begin{pmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{pmatrix} \begin{pmatrix} 2 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} \\
&= 4 \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{11} + x_{13} & x_{12} & x_{11} + x_{13} \\ x_{21} + x_{23} & x_{22} & x_{21} + x_{23} \\ x_{31} + x_{33} & x_{32} & x_{31} + x_{33} \end{pmatrix} \\
&= 4 \begin{pmatrix} x_{11} + x_{13} + x_{31} + x_{33} & x_{12} + x_{32} & x_{11} + x_{13} + x_{31} + x_{33} \\ x_{21} + x_{23} & x_{22} & x_{21} + x_{23} \\ x_{11} + x_{13} + x_{31} + x_{33} & x_{12} + x_{32} & x_{11} + x_{13} + x_{31} + x_{33} \end{pmatrix} \\
&= \begin{pmatrix} 2 & 0 & 2 \\ 0 & 2 & 0 \\ 2 & 0 & 2 \end{pmatrix} = \boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}.
\end{aligned}$$

Отсюда перепишем систему

$$\begin{cases} x_{11} + x_{13} + x_{31} + x_{33} = \dfrac{1}{2} \\ x_{12} + x_{32} = 0 \\ x_{21} + x_{23} = 0 \\ x_{22} = \dfrac{1}{2} \end{cases}$$

Тогда в качестве обобщённо-обратной матрицы можно использовать

$$(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-} = \begin{pmatrix} 1/8 & 0 & 1/8 \\ 0 & 1/2 & 0 \\ 1/8 & 0 & 1/8 \end{pmatrix}.$$

По теореме Гаусса-Маркова в случае матриц неполного ранга, вычислим несмещённую оценку $\widehat{\theta_2}$:

$$\widehat{\theta_2} = \boldsymbol{T}(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} = (y_2 - y_1)/2.$$

2.2 Метод наименьших квадратов

Рассмотрим задачу минимизации

$$\min_{\theta\in\mathbb{R}^3} \sum_{i=1}^{2} (y_i - (x_i^2\theta_3 + x_i\theta_2 + \theta_1))^2.$$

Для оценки параметра θ_2 найдём точку экстремума суммы квадратов как функции $g(\theta)$:

$$\frac{\partial g(\theta)}{\partial\theta_1} = -2\sum_{i=1}^{2} (y_i - (x_i^2\theta_3 + x_i\theta_2 + \theta_1)) = 0$$

$$\frac{\partial g(\theta)}{\partial\theta_2} = -2\sum_{i=1}^{2} x_i(y_i - (x_i^2\theta_3 + x_i\theta_2 + \theta_1)) = 0$$

$$\frac{\partial g(\theta)}{\partial\theta_3} = -2\sum_{i=1}^{2} x_i^2(y_i - (x_i^2\theta_3 + x_i\theta_2 + \theta_1)) = 0$$

В ходе решения этой системы линейных получим оценки $\widehat{\theta}$ и оценку $\widehat{\theta_2}$. Подставив x_i и сократив коэффициент перед суммой, заметим, что $\dfrac{\partial g(\theta)}{\partial\theta_1}$ и $\dfrac{\partial g(\theta)}{\partial\theta_3}$ совпадают, поэтому останется система

$$\begin{cases} y_1 + y_2 = 2(\theta_3 + \theta_1) \\ y_2 - y_1 = 2\theta_2 \end{cases}$$

Отсюда получаем

$$\widehat{\theta_2} = (y_2 - y_1)/2.$$

$\square$