Домашнее задание

Математика — Домашнее задание 12

А. Д. Егоров

2024 г.

Оглавление

1. Задача № 1

Условие

По определению

$$R^2 = \frac{ESS}{TSS}, \quad TSS = ESS + RSS,$$

$$TSS = \sum_{i=1}^{n}(y_i - \overline{y})^2, \quad ESS = \sum_{i=1}^{n}(\widehat{y}_i - \overline{y})^2, \quad RSS = \sum_{i=1}^{n}(y_i - \widehat{y}_i)^2,$$

а также $corr(y, \widehat{y}) = \dfrac{cov(y, \widehat{y})}{\sigma_y \sigma_{\widehat{y}}}$. Доказать, что $R^2 = corr^2(y, \widehat{y})$.

Решение

Известно, что $\boldsymbol{y} = \widehat{\boldsymbol{y}} + \widehat{\boldsymbol{\varepsilon}}$, выборочный смешанный момент второго порядка $S^2_{y\widehat{y}}$ такой, что

$$S^2_{y\widehat{y}} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \overline{y})(\widehat{y}_i - \overline{y}), \quad \sigma_y = S_{yy}.$$

С учетом этих формул:

$$corr(y, \widehat{y}) = \frac{cov(y, \widehat{y})}{\sigma_y \sigma_{\widehat{y}}} = \frac{cov(\widehat{y} + \widehat{\varepsilon}, \widehat{y})}{S_{yy} S_{\widehat{y}\widehat{y}}} =$$

$$= \frac{cov(\widehat{y}, \widehat{y}) + cov(\widehat{\varepsilon}, \widehat{y})}{S_{yy} S_{\widehat{y}\widehat{y}}} = \frac{cov(\widehat{y}, \widehat{y})}{S_{yy} S_{\widehat{y}\widehat{y}}} = \frac{S^2_{\widehat{y}\widehat{y}}}{S_{yy} S_{\widehat{y}\widehat{y}}} = \frac{S_{\widehat{y}\widehat{y}}}{S_{yy}}.$$

Тогда

$$R^2 = \frac{ESS}{TSS} = \frac{nS^2_{\widehat{y}\widehat{y}}}{nS^2_{yy}} = \frac{S^2_{\widehat{y}\widehat{y}}}{S^2_{yy}} = corr^2(y, \widehat{y}),$$

что и требовалось доказать.

2. Задача № 2

Условие

Оценки коэффициентов в матричной форме можно вычислить по следующей формуле:

$$\widehat{\boldsymbol{\beta}} = \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \boldsymbol{y}.$$

Оценки МНК несмещенные: $E\left[\widehat{\beta}_j | x\right] = \beta_j$.

Требуется:

1. Выразить оценки коэффициентов $\widehat{\boldsymbol{\beta}}$ через $\boldsymbol{\beta}$ в матричной форме.
2. Доказать свойство несмещенности коэффициентов $\widehat{\boldsymbol{\beta}}$ в матричной форме.

Решение

Известно, что $\boldsymbol{y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$, тогда подставим данное выражение в формулу для $\widehat{\boldsymbol{\beta}}$ и проведем преобразования:

$$\begin{aligned}
\widehat{\boldsymbol{\beta}} &= \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \boldsymbol{y} = \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \left(\boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}\right) = \\
&= \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \left(\boldsymbol{X}^T \boldsymbol{X}\right) \boldsymbol{\beta} + \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \boldsymbol{\varepsilon} = \\
&= \boldsymbol{I}\boldsymbol{\beta} + \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \boldsymbol{\varepsilon} = \boldsymbol{\beta} + \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \boldsymbol{\varepsilon},
\end{aligned}$$

где $\boldsymbol{I}$ — единичная матрица. Окончательно получим

$$\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta} + \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \boldsymbol{\varepsilon}.$$

Математическое ожидание данной величины будет следующим:

$$E\widehat{\boldsymbol{\beta}} = E\boldsymbol{\beta} + E\left[\left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \varepsilon\right] = \boldsymbol{\beta} + \mathbf{0} = \boldsymbol{\beta}.$$

Следовательно, $\widehat{\boldsymbol{\beta}}$ — оценка несмещенная.

3. Задача № 3

Условие

Доказать, что

$$\sum_{i=1}^{n}(y_i-\overline{y})^2 = RSS + ESS + 2\sum_{i=1}^{n}\widehat{\varepsilon}_i(\widehat{y}_i-\overline{y})$$

Решение

Известно, что $y_i=\widehat{y}_i+\widehat{\varepsilon}_i$. Подставим данное выражение в левую часть равенства, раскроем скобки и перегруппируем слагаемые:

$$\sum_{i=1}^{n}(y_i-\overline{y})^2 = \sum_{i=1}^{n}(\widehat{y}_i+\widehat{\varepsilon}_i-\overline{y})^2 = \sum_{i=1}^{n}\left[\widehat{y}_i^2-2\widehat{y}_i\overline{y}+\overline{y}^2+\widehat{\varepsilon}_i^2+2\widehat{y}_i\widehat{\varepsilon}_i-2\widehat{\varepsilon}_i\overline{y}\right] =$$

$$= \sum_{i=1}^{n}\left[(\widehat{y}_i-\overline{y})^2+\widehat{\varepsilon}_i^2+2\widehat{\varepsilon}_i(\widehat{y}_i-\overline{y})\right] = \sum_{i=1}^{n}(\widehat{y}_i-\overline{y})^2+\sum_{i=1}^{n}\widehat{\varepsilon}_i^2+2\sum_{i=1}^{n}\widehat{\varepsilon}_i(\widehat{y}_i-\overline{y}).$$

С учетом того, что $ESS=\sum_{i=1}^{n}(\widehat{y}_i-\overline{y})^2$, $RSS=\sum_{i=1}^{n}(y_i-\widehat{y}_i)^2=\sum_{i=1}^{n}\widehat{\varepsilon}_i^2$, получим

$$\sum_{i=1}^{n}(y_i-\overline{y})^2 = RSS + ESS + 2\sum_{i=1}^{n}\widehat{\varepsilon}_i(\widehat{y}_i-\overline{y}),$$

что и требовалось доказать.

4. Задача № 4

Условие

Используя предпосылку Гаусса-Маркова № 5 (гомоскедастичность стандартного отклонения ошибки) в матричной форме:

$$Var(\boldsymbol{\varepsilon}|\boldsymbol{X}) = \sigma^2 \boldsymbol{I},$$

где $\boldsymbol{I}$ — единичная матрица. Доказать, что

$$Var(\widehat{\boldsymbol{\beta}}|\boldsymbol{X}) = \sigma^2 (\boldsymbol{X}^T \boldsymbol{X})^{-1}.$$

Решение

Воспользуемся равенством $\widehat{\boldsymbol{\beta}} = \boldsymbol{\beta} + \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \boldsymbol{\varepsilon}$, полученным в задаче № 2: пусть $\boldsymbol{A} = \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T$, тогда

$$\begin{aligned} Var\left[\widehat{\boldsymbol{\beta}}|\boldsymbol{X}\right] &= Var\left[(\boldsymbol{\beta} + \boldsymbol{A}\boldsymbol{\varepsilon})\,|\boldsymbol{X}\right] = Var\big[\boldsymbol{\beta}|\boldsymbol{X}\big] + Var\left[\boldsymbol{A}\boldsymbol{\varepsilon}|\boldsymbol{X}\right] = Var\left[\boldsymbol{A}\boldsymbol{\varepsilon}|\boldsymbol{X}\right] = \\ &= Var\left[\boldsymbol{\varepsilon}|\boldsymbol{X}\right] \boldsymbol{A}\boldsymbol{A}^T = \sigma^2 \boldsymbol{A}\boldsymbol{A}^T = \sigma^2 \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \left[\left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T\right]^T = \\ &= \sigma^2 \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} \boldsymbol{X}^T \boldsymbol{X} \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} = \sigma^2 \boldsymbol{I} \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1} = \sigma^2 \left(\boldsymbol{X}^T \boldsymbol{X}\right)^{-1}. \end{aligned}$$

Окончательно получим

$$Var(\widehat{\boldsymbol{\beta}}|\boldsymbol{X}) = \sigma^2 (\boldsymbol{X}^T \boldsymbol{X})^{-1},$$

что и требовалось доказать.