

Цель этой заметки — объяснить, что находится в обратной матрице кросс-произведений  $(X^T X)^{-1}$ .  
Поехали!

Обозначим нашу матрицу  $(X^T X)^{-1}$  как  $V$ . Она симметрична,  $V^T = V$ . По определению,  $X^T X \cdot V = I$ . Разобьём произведение на две части по-другому,  $X^T \cdot X V = I$ .

Заметим, что  $X_* = X V$  — это матрица такого же размера, как  $X$ ,  $n \times k$ . Что же собой представляют эти звезданутые иксы,  $X_*$ ?

Во-первых,  $X_* = X V$ , поэтому новые  $X_*$  — это линейные комбинации исходных  $X$  с весами из матрицы  $V = (X^T X)^{-1}$ . Например,  $x_{1*} = v_{11}x_1 + v_{21}x_2 + \dots + v_{k1}x_k$ .

Во-вторых,  $X_*^T X_* = (X V)^T X V = V^T X^T X V = V = (X^T X)^{-1}$ . Матрица кросс-произведений новых переменных является обратной к матрице кросс-произведений исходных переменных.

В частности,  $v_{11} = x_{1*}^T x_{1*}$ , а  $v_{12} = v_{21} = x_{1*}^T x_{2*}$ .

В-третьих,  $X^T X_* = I$ , то есть новая переменная  $x_{i*}$  перпендикулярна всем старым переменным кроме  $x_i$ . Ортогональность позволяет нам взглянуть на первую звезданутую переменную как на остаток в регрессии. Первая ничем не выделяется, просто индексов меньше будет.

$$x_{1*} = v_{11}x_1 + v_{21}x_2 + \dots + v_{k1}x_k$$

Изолируем  $x_1$  слева, а всё остальное — перенесём вправо:

$$x_1 = -\frac{v_{21}}{v_{11}}x_2 - \frac{v_{31}}{v_{11}}x_3 - \dots - \frac{v_{k1}}{v_{11}}x_k + \frac{1}{v_{11}}x_{1*}$$

Мы только что заметили, что переменная  $x_{1*}$  перпендикулярна  $x_2, x_3, \dots, x_k$ , поэтому перед нами регрессия  $x_1$  на все остальные переменные  $X_{-1}$ :

$$x_1 = \hat{\beta}_{12}x_2 + \hat{\beta}_{13}x_3 + \dots + \hat{\beta}_{1k}x_k + e_1$$

Здесь  $e_1 = x_{1*}/v_{11}$  — это остаток от регрессии  $x_1$  на  $x_2, \dots, x_k$ . И  $\hat{\beta}_{12} = -v_{21}/v_{11} = -v_{12}/v_{11}$  — это оценка коэффициента в регрессии  $x_1$  перед переменной  $x_2$ .

Отсюда мы сразу видим, что  $V = (X^T X)^{-1}$  — это **матрица-мать всех регрессий!** Хочешь узнать коэффициент в регрессии  $x_1$  на все остальные переменные перед  $x_7$ ? В уме,  $\hat{\beta}_{17} = -v_{17}/v_{11}$ .

Найдём, чему равен  $RSS_1$  в этой регрессии.

$$RSS_1 = e_1^T e_1 = x_{1*}^T x_{1*} / v_{11}^2 = v_{11} / v_{11}^2 = 1 / v_{11}.$$

Бинго! Ура,  $v_{11} = 1/RSS_1$ . Заметим, что  $x_{1*} = e_1 v_{11} = e_1 / RSS_1$ . Типичный недиагональный элемент равен  $v_{12} = -\hat{\beta}_{12} v_{11} = -\hat{\beta}_{12} / RSS_1$ . И, наконец, матрицу-мать всех регрессий для  $k = 3$  в студию:

$$V = (X^T X)^{-1} = \begin{pmatrix} 1/RSS_1 & -\hat{\beta}_{12}/RSS_1 & -\hat{\beta}_{13}/RSS_1 \\ -\hat{\beta}_{21}/RSS_2 & 1/RSS_2 & -\hat{\beta}_{23}/RSS_2 \\ -\hat{\beta}_{31}/RSS_3 & -\hat{\beta}_{32}/RSS_3 & 1/RSS_3 \end{pmatrix}$$

Например, первая строка этой матрицы несёт всю информацию о регрессии  $x_1$  на  $x_2$  и  $x_3$ .

В случае пары центрированных регрессоров вспомним, что  $R^2$  одинаков в регрессии  $x_1$  на  $x_2$  и в обратной и равен  $R^2 = \hat{\beta}_{12}\hat{\beta}_{21}$ . Поэтому,  $v_{12}^2 = v_{21}^2 = \hat{\beta}_{12}\hat{\beta}_{21}/(RSS_1 RSS_2) = R^2/(RSS_1 RSS_2)$  и  $V = (X^T X)^{-1}$  равна

$$\begin{pmatrix} 1/RSS_1 & -\hat{\beta}_{12}/RSS_1 \\ -\hat{\beta}_{21}/RSS_2 & 1/RSS_2 \end{pmatrix} = \begin{pmatrix} \frac{1}{RSS_1} & -\frac{R}{\sqrt{RSS_1 RSS_2}} \\ -\frac{R}{\sqrt{RSS_1 RSS_2}} & \frac{1}{RSS_2} \end{pmatrix} = \frac{1}{1 - R^2} \begin{pmatrix} \frac{1}{TSS_1} & -\frac{R}{\sqrt{TSS_1 TSS_2}} \\ -\frac{R}{\sqrt{TSS_1 TSS_2}} & \frac{1}{TSS_2} \end{pmatrix}.$$