Регрессионный анализ. Занятие 1

Выведение оценки константы посредством метода наименьших квадратов

1. **Случай регрессии на контанту**

Запишем исходную спецификацию:

$$y_i = \beta_0 + \varepsilon_i$$

Перепишем в терминах модельных (предсказанных) значений, то есть, отклик (зависимая переменная) в среднем равна константе (некоторому постоянному значению):

$$\hat{y}_i = \hat{\beta}_0$$

Руководствуясь принципом МНК, минимизируем сумму квадратов остатков:

$$\frac{\partial \sum_{i=1}^{n}(y_i - \hat{\beta}_0)^2}{\partial \hat{\beta}_0} = 0$$

$$\frac{\partial \sum_{i=1}^{n}(y_i^2 - 2y_i\hat{\beta}_0 + \hat{\beta}_0^2)}{\partial \hat{\beta}_0} = 0$$

$$\frac{\partial (\sum_{i=1}^{n} y_i^2 - 2\hat{\beta}_0\sum_{i=1}^{n} y_i + n\hat{\beta}_0^2)}{\partial \hat{\beta}_0} = 0$$

$$-2\sum_{i=1}^{n} y_i + 2n\hat{\beta}_0 = 0$$

$$\hat{\beta}_0 = \frac{\sum_{i=1}^{n} y_i}{n}$$

$$\boxed{\hat{\beta}_0 = \bar{y}}$$

2. **Случай парной регрессии (один предиктор)**

Найдем оптимальную оценку константы ($\hat{\beta}_0$) в парной линейной регрессии, при которой сумма квадратов остатков будет минимальна.

Рассмотрим частную производную по $\hat{\beta}_0$:

$$\frac{\partial \sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2}{\partial \hat{\beta}_0} = 0$$

$$(-2)\sum_{i=1}^{n}(y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i) = 0$$

$$\sum_{i=1}^{n} y_i - n\hat{\beta}_0 - \hat{\beta}_1\sum_{i=1}^{n} x_i = 0$$

$$\hat{\beta}_0 = \frac{\sum_{i=1}^{n} y_i}{n} - \hat{\beta}_1 \frac{\sum_{i=1}^{n} x_i}{n}$$

Мы получили оценку константы в парной регрессии:

$$\boxed{\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}}$$

Продолжение следует. На следующем занятии мы получим МНК-оценку коэффициента при предикторе в парной линейной регрессии.