Least squares method

Linear dependence of two quantities. In science, and in particular in physics and statistics we use formulas expressing the dependence of two quantities which were derived based on experimental data. For example, based on measurements of the length of a copper rod in different temperatures we find a formula expressing the dependence of the copper rod length on temperature.

These formulas we call *empirical formulas*.

The most often used method of obtaining empirical formulas is described by Gauss *method of least squares*.

Let us assume that based on measurements for several values of the independent variable x:

$$x_1, x_2, \ldots, x_n$$

there were found values corresponding to the variable y:

$$y_1, y_2, \ldots, y_n.$$

On the plane we determine points

$$P_1(x_1, y_1), P_2(x_2, y_2), \ldots, P_n(x_n, y_n),$$

the orthogonal coordinates of which are corresponding values of x and y. Points $P_1, P_2, \ldots, P_n$ due to experimental errors may be irregularly scattered on the plane. Quite often though the drawing made or the sense of the problem may lead us to suppose that between the quatities x i y there should be some linear relationship

$$y = ax + b, \tag{1}$$

i.e. that the points $P_1, P_2, \ldots, P_n$ should be placed on a straight line having the equation (1). We want to determine this line, and therefore to set the coefficients a i b.

The relationship (1) we can write also as

$$y - ax - b = 0. \tag{2}$$

The points $P_1, P_2, \ldots, P_n$ in general are not situated on this line but in its vicinity, then the coordinates of these points do not have to fulfill the relationship (2).

We receive, however, that

$$y_1 - ax_1 - b = \varepsilon_1, y_2 - ax_2 - b = \varepsilon_2, \ldots, y_n - ax_n - b = \varepsilon_n, \tag{3}$$

where e.g. the number ε_1 denotes the difference of the ordinates of a point P_1 and the point Q_1, which lies on the straight line (1) and it has the same abscissa as point P_1 has. Similar meaning have the numbers $\varepsilon_2, \varepsilon_3, \ldots, \varepsilon_n$. We call these numbers *deviations of ordinates.*

The straight line (1), and then its coeficients a i b, we want to choose in such a way that the sum of squares of deviations

$$\begin{aligned} S = \varepsilon_1^2 + \varepsilon_2^2 + \ldots + \varepsilon_n^2 = \\ = (y_1 - ax_1 - b)^2 + (y_2 - ax_2 - b)^2 + \ldots + (y_n - ax_n - b)^2 \end{aligned} \tag{4}$$

is as small as possible. We look for such values a i b, for which the expression (4) assumes minimum. For this we need to solve the system of equations

$$\begin{aligned} \frac{\partial S}{\partial b} = 2(y_1 - ax_1 - b)(-1) + 2(y_2 - ax_2 - b)(-1) + \ldots + \\ +2(y_n - ax_n - b)(-1) = 0, \\ \frac{\partial S}{\partial a} = 2(y_1 - ax_1 - b)(-x_1) + 2(y_2 - ax_2 - b)(-x_2) + \ldots + \\ +2(y_n - ax_n - b)(-x_n) = 0, \end{aligned} \tag{5}$$

which after rearranging assumes the form

$$\begin{aligned} bn + a(x_1 + x_2 + \ldots + x_n) = y_1 + y_2 + \ldots + y_n, \\ b(x_1 + x_2 + \ldots + x_n) + a(x_1^2 + x_2^2 + \ldots + x_n^2), = \\ = x_1 y_1 + x_2 y_2 + \ldots + x_n y_n. \end{aligned}$$

Introducing the $\sum$ sign for abbreviated sum notation, i.e. denoting $n = \sum_{i=1}^{n} 1$ oraz

$$x_1^2 + x_2^2 + \ldots + x_n^2 = \sum_{i=1}^{n} x_i^2,$$

$$x_1 + x_2 + \ldots + x_n = \sum_{i=1}^{n} x_i,$$

$$y_1 + y_2 + \ldots + y_n = \sum_{i=1}^{n} y_i,$$

$$x_1 y_1 + x_2 y_2 + \ldots + x_n y_n = \sum_{i=1}^{n} x_i y_i,$$

we can write this system of equations as

$$
\begin{aligned}
b\sum_{i=1}^{n}1 + a\sum_{i=1}^{n}x_i &= \sum_{i=1}^{n}y_i, \\
b\sum_{i=1}^{n}x_i + a\sum_{i=1}^{n}x_i^2 &= \sum_{i=1}^{n}x_iy_i.
\end{aligned} \tag{6}
$$

These equations we call *system of normal equations of the method of least squares.* From the geometrical sense of the formula (4) it follows that S assumes minimum. It occurs then for the obtained values of a i b. We check this fact by computing the second partial derivatives and the discriminant. We have

$$
\begin{aligned}
\frac{\partial^2 S}{\partial a^2} &= 2(x_1^2 + x_2^2 + \ldots + x_n^2) > 0 \\
\frac{\partial^2 S}{\partial a \partial b} &= 2(x_1 + x_2 + \ldots + x_n) \\
\frac{\partial^2 S}{\partial b^2} &= 2 + 2 + \ldots + 2 = 2n
\end{aligned} \tag{7}
$$

Then the discriminant $w = S''^2_{a,b} - S''_{a,a}S''_{b,b}$

$$
w = 4[(x_1 + x_2 + \ldots + x_n)^2 - n(x_1^2 + x_2^2 + \ldots + x_n^2)] \tag{8}
$$

We check the discriminant e.g. for $n = 3$ and we obtain

$$
\begin{aligned}
w = 4[(x_1 + x_2 + x_3)^2 - 3(x_1^2 + x_2^2 + x_3^2)] = \\
= 4[(x_1 + x_2 + x_3)(x_1 + x_2 + x_3) - 3(x_1^2 + x_2^2 + x_3^2)] = \\
= 4[x_1^2 + x_1x_2 + x_1x_3 + x_1x_2 + x_2^2 + x_2x_3 + x_1x_3 + x_2x_3 + x_3^2 - \\
-3x_1^2 - 3x_2^2 - 3x_3^2] = \\
= 4[-x_1^2 + 2x_1x_2 - x_2^2 - x_1^2 + 2x_1x_3 - x_3^2 - x_2^2 + 2x_2x_3 - x_3^2] = \\
= 4[-(x_1^2 - 2x_1x_2 + x_2^2) - (x_1^2 - 2x_1x_3 + x_3^2) - (x_2^2 - 2x_2x_3 + x_3^2)] = \\
= -4[(x_1 - x_2)^2 + (x_1 - x_3)^2 + (x_2 - x_3)^2]
\end{aligned} \tag{9}
$$

Because always

$$
w < 0 \ \text{ i } \ \frac{\partial^2 S}{\partial a^2} > 0 \tag{10}
$$

we obtain the minimum.

References

[1] Romanowski, S., Wrona. W (1968) *Matematyka wyzsza dla studiow technicznych, Part III* PWN Warszawa

Pawel Jan Piskorz (paweljs@gmail.com)