Пусть опытным путём получены данные:

x	x_1	x_2	…	x_n
y	y_1	y_2	…	y_n

При выравнивании их по параболе получили зависимость $y=ax^2+bx+c$.
Требуется, используя метод наименьших квадратов (далее – МНК), аппроксимировать эти данные линейной зависимостью $y=kx+l$ и выяснить, какая из этих зависимостей лучше, в смысле МНК, выравнивает полученные данные.

Решение.

Согласно МНК, $\sum_{i=1}^{n}\varepsilon_i^2=\sum_{i=1}^{n}(y_i-f(x_i))^2\to min$, где f(x) – аппроксимирующая зависимость. Пусть $f(x)=kx+l$, тогда $\sum_{i=1}^{n}\varepsilon_i^2=\sum_{i=1}^{n}(y_i-kx_i-l)^2$. Рассмотрим функцию

$\omega(k,l)=\sum_{i=1}^{n}(y_i-kx_i-l)^2$: нам необходимо найти те значения аргументов, при которых она принимает наименьшее значение. Для этого приравняем к нулю её частные производные и решим полученную систему двух линейных уравнений с двумя неизвестными. Имеем:

$$\begin{cases}\omega'_k=\sum_{i=1}^{n}2(y_i-kx_i-l)(-x_i)=0,\\ \omega'_l=\sum_{i=1}^{n}2(y_i-kx_i-l)(-1)=0;\end{cases}\Leftrightarrow\begin{cases}\sum_{i=1}^{n}(x_iy_i-kx_i^2-lx_i)=0,\\ \sum_{i=1}^{n}(y_i-kx_i-l)=0;\end{cases}\Leftrightarrow$$

$$\Leftrightarrow\begin{cases}\sum_{i=1}^{n}x_iy_i-k\sum_{i=1}^{n}x_i^2-l\sum_{i=1}^{n}x_i=0,\\ \sum_{i=1}^{n}y_i-k\sum_{i=1}^{n}x_i-nl=0;\end{cases}\Leftrightarrow\begin{cases}k\sum_{i=1}^{n}x_i^2+l\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}x_iy_i,\\ k\sum_{i=1}^{n}x_i+ln=\sum_{i=1}^{n}y_i.\end{cases}\Leftrightarrow$$

$$\Leftrightarrow\begin{cases}k\sum_{i=1}^{n}x_i^2+l\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}x_iy_i,\\ l=\frac{1}{n}\sum_{i=1}^{n}y_i-\frac{k}{n}\sum_{i=1}^{n}x_i;\end{cases}\Leftrightarrow\begin{cases}k\sum_{i=1}^{n}x_i^2+(\overline{y}-k\overline{x})\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}x_iy_i,\\ l=\overline{y}-k\overline{x};\end{cases}\Leftrightarrow$$

$$\Leftrightarrow\begin{cases}k\sum_{i=1}^{n}x_i^2-k\overline{x}n\frac{1}{n}\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}x_iy_i-\overline{y}n\frac{1}{n}\sum_{i=1}^{n}x_i,\\ l=\overline{y}-k\overline{x};\end{cases}\Leftrightarrow$$

$$\Leftrightarrow\begin{cases}k(\sum_{i=1}^{n}x_i^2-\overline{x}n\overline{x})=\sum_{i=1}^{n}x_iy_i-\overline{y}n\overline{x},\\ l=\overline{y}-k\overline{x};\end{cases}\Leftrightarrow\begin{cases}k(\sum_{i=1}^{n}x_i^2-\sum_{i=1}^{n}\overline{x}^2)=\sum_{i=1}^{n}x_iy_i-\sum_{i=1}^{n}\overline{x}\overline{y},\\ l=\overline{y}-k\overline{x};\end{cases}\Leftrightarrow$$

$$\Leftrightarrow\begin{cases}k\sum_{i=1}^{n}(x_i-\overline{x})^2=\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y}),\\ l=\overline{y}-k\overline{x};\end{cases}\Leftrightarrow\begin{cases}k=\dfrac{\sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})}{\sum_{i=1}^{n}(x_i-\overline{x})^2},\\ l=\overline{y}-k\overline{x}.\end{cases}$$

Для того чтобы выяснить, какая зависимость лучше, в смысле МНК, аппроксимирует опытные данные, нужно для каждой из них подсчитать сумму $\sum_{i=1}^{n}\varepsilon_i^2=\sum_{i=1}^{n}(y_i-f(x_i))^2$. Чем меньше значение суммы, тем лучше аппроксимация.

$$\boxed{\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} \overline{xy} = \sum_{i=1}^{n} \left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right)}$$

Доказательство

$$\sum_{i=1}^{n} \left(x_i - \overline{x}\right)\left(y_i - \overline{y}\right) = \sum_{i=1}^{n} \left(x_i y_i - x_i \overline{y} - y_i \overline{x} + \overline{xy}\right) = \sum_{i=1}^{n} x_i y_i - \overline{y} \sum_{i=1}^{n} x_i -$$

$$- \overline{x} \sum_{i=1}^{n} y_i + n\overline{xy} = \sum_{i=1}^{n} x_i y_i - \frac{1}{n} \sum_{i=1}^{n} y_i \sum_{i=1}^{n} x_i - \frac{1}{n} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i + n\overline{xy} =$$

$$= \sum_{i=1}^{n} x_i y_i - \frac{2}{n} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i + n \frac{1}{n} \sum_{i=1}^{n} x_i \frac{1}{n} \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} x_i y_i - \frac{2}{n} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i +$$

$$+ \frac{1}{n} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} x_i y_i - \frac{1}{n} \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i = \sum_{i=1}^{n} x_i y_i - n \frac{1}{n} \sum_{i=1}^{n} x_i \frac{1}{n} \sum_{i=1}^{n} y_i =$$

$$= \sum_{i=1}^{n} x_i y_i - n\overline{xy} = \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} \overline{xy}$$

В частности $\boxed{\sum_{i=1}^{n} x_i^2 - \sum_{i=1}^{n} \overline{x}^2 = \sum_{i=1}^{n} \left(x_i - \overline{x}\right)^2}$