Die Frage lautete, wo denn die Mittelwerte, die in der Lösung zum ersten Mal auftauchen, herkommen. Um diese Frage zu beantworten, muss man die Summen umstellen und die Gleichungen ein wenig umformen. Im Lösungsweg sieht dies wie folgt aus:

Gesucht ist das Minimum der Funktion: $f(b,m) = \sum_{i=1}^{n} (mx_i + b - y_i)^2$

Dank der verwendeten Fehlerfunktion $\left(\sum_{i=1}^{n} r_i^2\right)$ ist diese Funktion konvex mit einem einzigen Extremum - dem Minimum das wir suchen.

Für unsere Rechnung bedeutet dies, dass es genügt, die Nullstellen der ersten Ableitungen zu bestimmen.

Die Partiellen Ableitungen unserer Zielfunktion nach m und b lauten wie folgt:

$$f_b'(b,m) = \sum_{i=1}^{n} 2\,(b + mx_i - y_i) \quad = \sum_{i=1}^{n} (2b + 2mx_i - 2y_i)$$

$$f_m'(b,m) = \sum_{i=0}^{n} 2\,(b + mx_i - y_i)\,x_i \quad = \sum_{i=1}^{n} (2bx_i + mx_i^2 - x_i y_i)$$

Wenn wir die Summen umschreiben erhalten wir:

$$f_b'(b,m) = 2bn + 2m\sum_{i=1}^{n} x_i - 2\sum_{i=1}^{n} y_i = 0$$

$$f_m'(b,m) = 2b\sum_{i=1}^{n} x_i + 2m\sum_{i=1}^{n} x_i^2 - 2\sum_{i=1}^{n} x_i y_i = 0$$

Um die Nullstellen zu bestimmen, setzen wir die Gleichungen gleich null.

Dieser Schritt entspricht einer Umsortierung der Summen. Die dabei entstehenden Teilsummen, wie bsw. $\sum_{i=1}^{n} 2b$ lassen sich dabei sofort auflösen $\sum_{i=1}^{n} 2b = 2b + 2b + \ldots + 2b = n \cdot 2b$. Diese Terme sind nicht von der Laufvariable i abhängig und daher in Bezug auf die Summe konstant.

Diese Gleichunge sollen jetzt in ein Gleichungssystem überführt werden. Dazu stellen wir die Gleichungen zunächst wie folgt um:

$$2nb + 2\sum_{i=1}^{n}(x_i)\, m = 2\sum_{i=1}^{n} y_i$$

$$2\sum_{i=1}^{n}(x_i)\, b + 2\sum_{i=1}^{n}(x_i^2)\, m = 2\sum_{i=1}^{n} x_i y_i$$

Dies können wir jetzt als lineares Gleichungssystem mit m und b als Unbekannte schreiben:

$$\begin{pmatrix} 2n & 2\sum_{i=1}^{n} x_i \\ 2\sum_{i=1}^{n} x_i & 2\sum_{i=1}^{n} x_i^2 \end{pmatrix} \begin{pmatrix} b \\ m \end{pmatrix} = \begin{pmatrix} 2\sum_{i=1}^{n} y_i \\ 2\sum_{i=1}^{n} x_i y_i \end{pmatrix} \begin{array}{l} \mid :2n \\ \mid :2 \end{array}$$

Die erste Zeile teilen wir durch 2n, die zweite Zeile durch 2 und erhalten so:

$$\begin{pmatrix} 1 & \frac{1}{n}\sum_{i=1}^{n} x_i \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{pmatrix} \begin{pmatrix} b \\ m \end{pmatrix} = \begin{pmatrix} \frac{1}{n}\sum_{i=1}^{n} y_i \\ \sum_{i=1}^{n} x_i y_i \end{pmatrix}$$

Die Terme, die wir hier erhalten haben: $\frac{1}{n}\sum_{i=1}^{n} x_i$ und $\frac{1}{n}\sum_{i=1}^{n} y_i$ beschreiben die Mittelwerte der x_i und y_i. Die Notation kann wie folgt vereinfacht werden: $\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$ und $\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i$.

$$\Leftrightarrow \begin{pmatrix} 1 & \bar{x} \\ \sum_{i=1}^{n} x_i & \sum_{i=1}^{n} x_i^2 \end{pmatrix} \begin{pmatrix} b \\ m \end{pmatrix} = \begin{pmatrix} \bar{y} \\ \sum_{i=1}^{n} x_i y_i \end{pmatrix}$$

An dieser Stelle lässt sich die Lösung für b auch bereits ablesen: $b = \bar{y} - m\bar{x}$.

Um m zu bestimmen Bedarf es noch einiger Umformungen.