Regressionsgerade wenn x und y-Werte fehlerbehaftet sind

Michael Kopp

Ich gehe von einer Menge von N Punkten $P^1, ..., P^N$ aus, durch welche man eine ideale Gerade mit

$$f(x) = m \cdot x \tag{1}$$

legen soll. Dabei soll die Summe der Quadratischen Abweichungen von $f(x)$ und den Punkten insgesamt minimal werden. Dabei ist jedoch zu beachten, dass man nicht nur die Fehler in y-Richtung einbeziehen muss, sondern genauso auch die Fehler in x-Richtung.

Dazu betrachte ich einen Punk $P(x_p, y_p)$. Um den Abstand von $f(x)$ zu bestimmen, kann man die Gleichung der Normalen durch P (die senkrecht zu $f(x)$ verläuft – schließlich will man den Abstand Punkt-Gerade messen und dazu braucht man eine senkrechte Hilfsgerade) angeben via

$$n: \; y = -\frac{1}{m} \cdot (x - x_p) + y_p \; . \tag{2}$$

Der Schnittpunkt der beiden Geraden ist gegeben durch:

$$\tilde{x} = \frac{\frac{x_p}{m} + y_p}{m + \frac{1}{m}} \text{ und } \tilde{y} = m \cdot \tilde{x} \; . \tag{3}$$

Um nun den quadratischen Abstand von P zu $f(x)$ zu bestimmen, verwendet man Pythagoras:

$$d_p^2 = \Delta x^2 + \Delta y^2 = \left(x_p - \frac{\frac{x_p}{m} + y_p}{m + \frac{1}{m}} \right)^2 + \left(y_p - m \cdot \frac{\frac{x_p}{m} + y_p}{m + \frac{1}{m}} \right)^2 \; . \tag{4}$$

Nun kann man für jeden Punkt diesen quadratischen Abstand bestimmen und die Summe davon als zu minimierende gesamte Quadratische Abweichung verwenden:

$$D(P^1, ..., P^N, m) = D(m) = \sum_{p=1}^{N} d_p^2 \; . \tag{5}$$

Jetzt müsste man – zumindest theoretisch – einfach das gesuchte m als m_0 über

$$\frac{\mathrm{d}\,D}{\mathrm{d}\,m}(m_0) = 0 \tag{6}$$

bestimmen können...