Mathematik für Naturwissenschaftler I
WS 2009/2010

Lektion 21

26. Januar 2010

Kapitel 5. Funktionen mehrerer Veränderlicher, Stetigkeit und partielle Ableitungen

§5.9 Lineare Ausgleichrechnung

Ein fundamentales Problem aus der Praxis des Naturwissenschaftlers ist der Ausgleich von Messfehlern.

Problem:

Zu den (paarweise) verschiedenen Werten $x_1, \dots, x_n$ werden die y-Werte $y_1, \dots, y_n$ gemessen. Also liegen n Messepunkte

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n) \in \mathbb{R}^2$$

vor.

Zur Lösung dieses Problems machen wir den

Ansatz:

Es ist sinnvoll, y als lineare Funktion von x zu suchen.

Frage:

Wie bestimmt man $a, b \in \mathbb{R}$ in

$$y = ax + b,$$

so dass diese Gerade sich den Messpunkten am besten anpasst?

Bemerkung.

Beachte, dass für $n \geqslant 3$ i.a. keine Gerade durch alle Messpunkte gibt!

Methode der kleinsten Quadrate:

Die Abweichung bei x_i ist offenbar

$$\varphi_i := ax_i + b - y_i.$$

Bestimme nun $a, b \in \mathbb{R}$ derart, dass

$$\sum_{i=1}^{n} \varphi_i^2 = \varphi_1^2 + \varphi_2^2 + \cdots + \varphi_n^2$$

den kleinmöglichen Wert annimmt.

Gesucht ist also das absolute Minimum der Funktion

$$\mathbb{R}^2 \ni (a, b) \mapsto f(a, b) := \sum_{i=1}^{n} (ax_i + b - y_i)^2 \,.$$

1. Berechnung der partiellen Ableitungen:

$$\begin{cases} \dfrac{\partial f}{\partial a}(a,b) = 2\sum_{i=1}^{n}(ax_i + b - y_i)\,x_i = 0 \\ \dfrac{\partial f}{\partial b}(a,b) = 2\sum_{i=1}^{n}(ax_i + b - y_i) = 0 \end{cases}$$

$$\Leftrightarrow$$

$$\begin{cases} a\left(\sum_{i=1}^{n} x_i^2\right) + b\left(\sum_{i=1}^{n} x_i\right) = \sum_{i=1}^{n} x_i y_i \\ a\left(\sum_{i=1}^{n} x_i\right) + nb = \sum_{i=1}^{n} y_i \end{cases}$$

$$x_i \neq x_j \qquad \Rightarrow \qquad \sum_{i=1}^{n} x_i^2 > 0.$$

$$\boxed{\mathbf{a} = \frac{\sum\limits_{i=1}^{n} x_i y_i - \mathbf{b}\left(\sum\limits_{i=1}^{n} x_i\right)}{\sum\limits_{i=1}^{n} x_i^2}}$$

$$\frac{\sum\limits_{i=1}^{n} x_i y_i - b\left(\sum\limits_{i=1}^{n} x_i\right)}{\sum\limits_{i=1}^{n} x_i^2} \cdot \left(\sum_{i=1}^{n} x_i\right) + nb = \sum_{i=1}^{n} y_i$$

$$b \cdot \left\{ n\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2 \right\} = \left(\sum_{i=1}^{n} x_i^2\right) \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i y_i$$

$$n\left(\sum_{i=1}^{n} x_i^2\right) - \left(\sum_{i=1}^{n} x_i\right)^2 = \frac{1}{2}\sum_{i,j=1}^{n}(x_i - x_j)^2 = \sum_{1\leqslant i<j\leqslant n}(x_i - x_j)^2 > 0,$$

$$\text{weil} \quad n \geqslant 2 \quad \text{und} \quad x_i \neq x_j$$

$$\Rightarrow$$

$$\boxed{\mathbf{b} = \frac{\left(\sum\limits_{i=1}^{n} x_i^2\right)\sum\limits_{i=1}^{n} y_i - \sum\limits_{i=1}^{n} x_i y_i}{\sum\limits_{1\leqslant i<j\leqslant n}(x_i - x_j)^2}}$$

Es gibt **genau ein** Lösung **a** und **b**.

Um festzustellen, welche Art Extremum vorliegt, berechnen wir

$$\frac{\partial^2 f}{\partial a^2}(a,b) = 2\sum_{i=1}^{n} x_i^2 > 0, \quad \text{weil} \quad n \geqslant 2$$

$$\frac{\partial^2 f}{\partial a \partial b}(a,b) = 2\sum_{i=1}^{n} x_i$$

$$\frac{\partial^2 f}{\partial b^2}(a,b) = 2\sum_{i=1}^{n} = 2n$$

$$\frac{\partial^2 f}{\partial a^2} \cdot \frac{\partial^2 f}{\partial b^2} - \left(\frac{\partial^2 f}{\partial a \partial b}\right)^2 = 4n \sum_{i=1}^{n} x_i^2 - 4 \left(\sum_{i=1}^{n} x_i\right)^2 = 4d > 0$$

Also nimmt f in (a, b) ein lokales Minimum an. Da f keine weiteren Extrema besitzt, ist $f(a, b)$ sogar ein globales Minimum.

D.h. die Gerade

$$y = ax + b$$

ist die gesuchte Lösung. Die durch $y = ax + b$ bestimmte Gerade heißt Regressionsgerade (oder Ausgleichgerade) von y bzgl. x.

Frage:

Wann ist die Regressionsgerade eine gute Näherung für den Zusammenhng zwischen x und y?

Setze für die Mittelwerte

$$\bar{x} := \frac{1}{n}\sum_{i=1}^{n} x_i, \qquad \bar{y} := \frac{1}{n}\sum_{i=1}^{n} y_i;$$

dann nennt man (O.E. nicht alle y_i gleich)

$$r = \frac{\sum\limits_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum\limits_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum\limits_{i=1}^{n}(y_i - \bar{y})^2}}$$

den Korellationskoeffizienten von $x_1, \ldots, x_n$ und $y_1, \ldots, y_n$.

Für den Korellationskoeffizienten r gilt

$r \in [-1, 1]$ und $|r| = 1 \quad \Leftrightarrow \quad$ alle Messpunkte liegen auf eine (nicht horizontalen) Geraden.

Bemerkung.

Die Wahrscheinlichkeit dafür, dass y linear von x abhängt, ist umso größer, desto näher r bei 1 oder -1 liegt.

Um eine feste Wahrscheinlichkeit zu erhalten, kann bei wachsendem n der Abstand zu ± 1 wachsen!