4.1.3* Ein *inverses* Maximum-Likelihood-Argument von Gauß

(*Theoria motus corporum coelestium in sectionibus conicis solem ambientium*, 1809, Artikel 177; der Titel bedeutet: „Theorie der Himmelskörper-Bewegungen um die Sonne längs Kegelschnitten“)

Carl Friedrich Gauß untersuchte die mathematischen Gesetzmäßigkeiten von Messfehlern, ist der Begründer der *Fehlerrechnung*. Er erkannte als erster die allgemeine Rolle der sogenannten Normalverteilung, deren Dichte daher auch *Gaußsche Glockenkurve* genannt wird. Hier seine raffiniert einfache und doch bahnbrechende Herleitung:

Durch n unabhängige gleich genaue Messungen zur Bestimmung einer Größe x habe man die Messwerte $x_1, \ldots, x_n$ erhalten.

Ist φ die Wahrscheinlichkeitsdichte der Messfehler, so ist zur Bestimmung des *Maximum-Likelihood-Schätzers* von x die Funktion

$$L(x) = \varphi(x_1 - x)\varphi(x_2 - x) \cdots \varphi(x_n - x)$$

zu maximieren. Mit $\tilde{\varphi} := \varphi'/\varphi$ ergibt sich nach Logarithmieren durch Ableiten nach x die Bestimmungsgleichung für den ML-Schätzer:

$$\tilde{\varphi}(x_1 - x) + \tilde{\varphi}(x_2 - x) + \cdots + \tilde{\varphi}(x_n - x) = 0.$$

Wenn das *arithmetische Mittel* $\bar{x}$ der *beste* Schätzer bei gleich genauen Einzelmessungen ist (wovon, betonte Gauß, man allgemein stets ausgehe), sollte es *mit dem ML-Schätzer übereinstimmen*, also die gerade aufgestellte Gleichung für $x = \bar{x}$ erfüllt sein.

Das arithmetische Mittel $\bar{x}$ ist charakterisiert durch

$$(x_1 - \bar{x}) + (x_2 - \bar{x}) + \cdots + (x_n - \bar{x}) = 0.$$

Gauß bemerkte nun*):

Gilt $\sum_{k=1}^{n} u_k = 0 \;\Rightarrow\; \sum_{k=1}^{n} f(u_k) = 0$ für eine *stetige* Funktion f, ist $\dfrac{f(u)}{u}$ konstant.

Beweis:
Mit untereinander gleichen Werten $u_2 = u_3 = \cdots = u_n = u$, so dass also $u_1 = (1-n)u$, folgt $f((1-n)u) = -(n-1)f(u)$, insbesondere (Fall $n=2$) $f(-u) = -f(u)$ und damit $f(ku) = kf(u)$ $(k \in \mathbb{Z})$, also $f(ru) = rf(u)$ $(r \in \mathbb{Q})$, und $\frac{f(ru)}{ru} = \frac{f(u)}{u}$ ist *konstant*. ∎

Dies aber bedeutet $\tilde{\varphi}(x) = ax$ (a eine Konstante), also $\varphi'(x) = ax\varphi(x)$ und $a < 0$, d.h. $a = -2h^2$, und damit $\varphi(x) = C\mathrm{e}^{-h^2x^2}$. Da das Gesamtintegral über φ den Wert 1 ergeben muss (Fehler-Wahrscheinlichkeitsdichte!), folgt mit dem Laplaceschen Integral $\int_{-\infty}^{\infty} \mathrm{e}^{-x^2}\mathrm{d}x = \sqrt{\pi}$, dass $C = h/\sqrt{\pi}$.

Fazit: Die Fehlerdichte ist $\boldsymbol{\varphi(x) = \frac{h}{\sqrt{\pi}}\, \mathrm{e}^{-h^2x^2}}$, die *Normalverteilung* (mit $\mu = 0$).

Gauß merkte an, h entspreche der *Messgenauigkeit*, und diese Fehlerdichte könne natürlich nur *näherungsweise* richtig sein, da ja *beliebig große* Fehler unmöglich sind.

Mit der Annahme annähernd normalverteilter Messfehler konnte nun Gauß die *Kleinste-Quadrate-Methode* durch ein Maximum-Likelihood-Argument einfach begründen. Mehr dazu im folgenden Abschnitt 4.1.4.

Laplace wandte ein, auch *andere* Fehlerverteilungen seien zu berücksichtigen, und meinte, *durchschnittlich möglichst geringe Fehler* seien wichtiger beim Schätzen als maximale Wahrscheinlichkeit. F.W. Bessel wies empirisch nach, dass astronomische Beobachtungsfehler fast immer in etwa normalverteilt sind, und Gauß zeigte später, dass Kleinste-Quadrate-Schätzer den minimalen durchschnittlichen Fehler aller linearen Schätzer besitzen. Und im 20. Jahrhundert wurde das Maximum-Likelihood-Prinzip durch die Untersuchungen von R.A. Fisher und anderen zu einem der bestbegründeten allgemeinen Schätzverfahren.

*) Dies ist äquivalent mit „Cauchys Funktionalgleichung“; *Analyse algébrique*, Paris 1821, p. 104ff.