

Fehler- und Ausgleichsrechnung

Daniel Gerth

Überblick Fehler- und Ausgleichsrechnung

Dieses Kapitel erklärt:

- Wie man Ausgleichsrechnung betreibt
- Wie man Messreihen statistisch auswertet und interpretiert
- Wie man Fehlerfortpflanzung kontrolliert

Regressionsgerade

Wir betrachten die folgende Aufgabe: Von einem gegebenen naturwissenschaftlichen Prozess sei bekannt, dass er durch einen eindimensionalen linearen Zusammenhang

$$y = f(x) = mx + c$$

modelliert werden kann. Unbekannt jedoch sind die Geradenparameter m (Steigung) und c (y -Achsenabschnitt). Ausgehend von den x -Werten $x_1, \dots, x_r$, die allesamt **verschieden** voneinander sein sollen und welche wir als **Messstellen** interpretieren wollen, und den an diesen Stellen vorliegenden y -Werten $y_1, \dots, y_r$, interpretierbar als **Messwerte**, wollen wir auf die unbekannten Parameter m und c schließen.

Wo liegt das Problem? (theoretisch reichen zwei Punktepaae (x_1, y_1) und (x_2, y_2) aus, um die Parameter m und c festzulegen). **Bedingt durch Fehler oder Ungenauigkeiten beim Messvorgang** liegen die Punkte

$$(x_1, y_1), \dots, (x_r, y_r)$$

typischerweise jedoch nicht auf einer Geraden, d.h. $y_i = m \cdot x_i + c + \epsilon_i$ wobei ϵ_i “klein” ist.

Wir suchen nach **Kompromisse**: Zu bestimmen ist dann eine Gerade

$$y = mx + c$$

welche den Punktepaaen zumindest leidlich nahekommt. **Hinweis: Die hier präsentierte Theorie gilt analog für beliebige Polynome!**

Für jedes Paar (x_i, y_i) setzen wir trotz des Fehlers ϵ an:

$y_i = m \cdot x_i + c = \begin{pmatrix} x_i & 1 \end{pmatrix} \begin{pmatrix} m \\ c \end{pmatrix}$ Man kann damit das gesamte Problem schreiben als

$$\vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_r \end{pmatrix} = \begin{pmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_r & 1 \end{pmatrix} \begin{pmatrix} m \\ c \end{pmatrix} = A\vec{v}$$

Es entsteht also eine Matrixgleichung mit $\vec{y} \in \mathbb{R}^r$, $A \in \mathbb{R}^{r \times n}$ und $\vec{v} \in \mathbb{R}^n$ in der typischerweise $r > n$. Bedingt durch die Messfehler hat diese Gleichung keine Lösung. Stattdessen sucht man nach Werten m und c , so dass der resultierende Vektor $\vec{v}^*$ einen kleinen Fehler produziert:

$$\vec{v}^* = \min_{\vec{v} \in \mathbb{R}^n} \|\vec{y} - A\vec{v}\|^2$$

(Man kann dies für alle linearen Systeme $y = Ax$ machen, wir behandeln hier jedoch nur den oben genannten Spezialfall)

Nun ist $||\vec{y} - A\vec{v}||^2$ eine Funktion von $\mathbb{R}^n$ nach $\mathbb{R}$, fällt also in ein bekanntes Schema. Wir benötigen noch folgende hilfreiche Werkzeuge:

Theorem 1

Sei $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$ und $M \in \mathbb{R}^{m \times n}$. Dann gilt für das Skalarprodukt in $\mathbb{R}^m$ bzw. $\mathbb{R}^n$:

$$\langle Mx, y \rangle = \langle x, M^T y \rangle$$

Desweiteren gilt, falls $M = A^T$:

$$\text{grad}_x(\langle Mx, x \rangle) = 2Mx$$

und

$$\text{grad}_x(\langle x, z \rangle) = z$$

Damit gilt:

$$\begin{aligned} ||A\vec{v} - \vec{y}||^2 &= \langle A\vec{v} - \vec{y}, A\vec{v} - \vec{y} \rangle \\ &= \langle A\vec{v}, A\vec{v} \rangle - 2\langle A\vec{v}, \vec{y} \rangle + \langle \vec{y}, \vec{y} \rangle \\ &= \langle A^T A\vec{v}, \vec{v} \rangle - 2\langle A^T \vec{y}, \vec{v} \rangle + ||\vec{y}||^2 \end{aligned}$$

Um $f(\vec{v}) = \|\vec{y} - A\vec{v}\|^2$ zu minimieren, müssen wir den Gradienten gleich Null setzen:

$$\begin{aligned} 0 &= \text{grad}(\|\vec{y} - A\vec{v}\|^2) \\ &= \text{grad}(\langle A^T A\vec{v}, \vec{v} \rangle - 2\langle A^T \vec{y}, \vec{v} \rangle + \|\vec{y}\|^2) \\ &= 2A^T A\vec{v} - 2A^T \vec{y} \end{aligned}$$

Damit ist $\vec{v}^*$ Lösung der so genannten **Normalengleichung**

$$A^T A\vec{v} = A^T \vec{y}$$

Die Lösung ist eindeutig, falls $\text{rang}(A) = n$. Dann ist $A^T A$ symmetrisch und positiv semidefinit, und die Lösung stellt tatsächlich ein Minimum dar.

Gegeben sei die Messreihe (x_i, y_i) :

$$(1, 0.9) \quad (3, 5.5) \quad (5, 9.1)$$

Wir suchen den linearen Zusammenhang $y = mx + c$, also ist

$$\vec{y} = \begin{pmatrix} 0.9 \\ 5.5 \\ 9.1 \end{pmatrix}, \quad \vec{v} = \begin{pmatrix} m \\ c \end{pmatrix}, \quad A = \begin{pmatrix} 1 & 1 \\ 3 & 1 \\ 5 & 1 \end{pmatrix}$$

Die Normalengleichung $A^T A \vec{v} = A^T \vec{y}$ lautet

$$\begin{pmatrix} 1 & 3 & 5 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 3 & 1 \\ 5 & 1 \end{pmatrix} \begin{pmatrix} m \\ c \end{pmatrix} = \begin{pmatrix} 1 & 3 & 5 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0.9 \\ 5.5 \\ 9.1 \end{pmatrix}$$
$$\begin{pmatrix} 35 & 9 \\ 9 & 3 \end{pmatrix} \begin{pmatrix} m \\ c \end{pmatrix} = \begin{pmatrix} 61 \\ 15 \end{pmatrix}$$

Die Lösung lautet $\vec{v}^* = \begin{pmatrix} 2.05 \\ -0.98 \end{pmatrix}$, also gilt die Funktionsgleichung

$y = 2.05x - 0.98$, zum Vergleich: die Gleichung mit fehlerlosen Daten ist $y = 2x - 1$.

Messreihenbewertung

Wir nehmen an, wir messen eine Größe x (z.B. die Dicke eines Blatt Papiers, den PH-Wert einer Lösung, usw). Für eine einzelne Messung haben wir folgende Begriffe:

Definition 2.1

Gegeben sei eine Messgröße x , für welche ein Messwert $\bar{x}$ existiere. Der wirkliche, unbekannte Wert der Messgröße sei ebenfalls mit x bezeichnet. Dann heißt die Differenz

$$\Delta x := x - \bar{x}$$

*der **absolute Fehler** der Größe x . Der Quotient*

$$\frac{\Delta x}{\bar{x}} = \frac{x - \bar{x}}{\bar{x}}$$

*heißt **relativer Fehler** der Größe x .*

Angenommen, wir haben nun n Messwerte $x_1, x_2, \dots, x_n$ für unsere Größe x .

Definition 2.2

Der **Mittelwert** der Messung ist

$$\bar{x} := \frac{1}{n} \sum_{i=1}^n x_i$$

die **Standardabweichung** (*“gemittelter Fehler der Einzelmessungen”*) beträgt

$$\sigma := \sqrt{\frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2}$$

und der mittlere Fehler des Mittelwertes, also die zu erwartende Unsicherheit für die Schätzung des Mittelwertes, ist $\frac{\sigma}{\sqrt{n}}$

Damit gilt

$$\mathbb{P}(\bar{x} - \sigma \leq x_i \leq \bar{x} + \sigma) = 0.682$$

bzw.

$$\mathbb{P}(\bar{x} - 0.677\sigma \leq x_i \leq \bar{x} + 0.677\sigma) = 0.5$$

Fehlerfortpflanzung

Als wichtige Anwendung des totalen Differentials wollen wir die unabhängigen Variablen $x_1, \dots, x_n$ einer Funktion $f(x_1, \dots, x_n)$ als **Messgrößen eines naturwissenschaftlichen Experiments interpretieren**. Abhängig davon soll eine weitere Größe

$$z = f(x_1, \dots, x_n)$$

errechnet werden.

Ideal wäre es, die obigen Messgrößen exakt zu

$$\vec{x}_{\text{Messung}} = (\bar{x}_1, \dots, \bar{x}_n)$$

bestimmen zu können. Unter diesen Umständen resultierte daraus die ebenfalls exakt bestimmbare abhängige Größe

$$z_{\text{Messung}} = f(\bar{x}_1, \dots, \bar{x}_n).$$

Die Wirklichkeit sieht anders aus. Bedingt durch Ungenauigkeiten der Messmethoden o.Ä. sind die Messwerte $\bar{x}_1, \dots, \bar{x}_n$ **fehlerbehaftet**. Die wirklichen, aber leider unbekannten Messgrößen $x_1, \dots, x_n$ **weichen von den gemessenen Werten** $\bar{x}_1, \dots, \bar{x}_n$ geringfügig ab.

Welche Auswirkung haben kleine Abweichungen der realen Größen von den Messwerten auf das zu berechnende Endergebnis z ? Wir wirken sich Fehler in den Messwerten auf die davon abhängige Größe $z = f(x_1, \dots, x_n)$ auf?

Satz 3.1

Die Größe $x_1, \dots, x_n$ seien gemessen zu $\bar{x}_1, \dots, \bar{x}_n$ und fehlerbehaftet mit maximalen Abweichungen

$$\Delta_{\max} x_1, \dots, \Delta_{\max} x_n.$$

Ist f differenzierbar in $(\bar{x}_1, \dots, \bar{x}_n)$, so gilt für den maximalen absoluten Fehler der abhängigen Größe $z = f(x_1, \dots, x_n)$

$$\Delta_{\max} z \approx |f_{x_1}(\bar{x}_1, \dots, \bar{x}_n)| \Delta_{\max} x_1 + \dots + |f_{x_n}(\bar{x}_1, \dots, \bar{x}_n)| \Delta_{\max} x_n. \quad (1)$$

Wir betrachten zwei Größen x und y , deren Werte zu $\bar{x} \neq 0$ und $\bar{y} \neq 0$ ermittelt wurden, aber $\Delta_{\max}x$ bzw. $\Delta_{\max}y$ fehlerbehaftet sind.

- **Fehlerfortpflanzung bei Summation und Differenzbildung**

$$\Delta_{\max}z = \Delta_{\max}x + \Delta_{\max}y,$$

wobei $z = f(x, y) = x \pm y$.

- **Fehlerfortpflanzung bei Produkt- und Quotientenbildung**

$$\Delta_{\max}p = |\bar{y}|\Delta_{\max}x + |\bar{x}|\Delta_{\max}y,$$

wobei $p(x, y) = xy$.

$$\Delta_{\max}q = \frac{1}{|\bar{y}|}\Delta_{\max}x + \frac{|\bar{x}|}{\bar{y}^2}\Delta_{\max}y,$$

wobei $q(x, y) = \frac{x}{y}$.