

Vorkurs Mathematik

im Wintersemester 2024

Kapitel I – Methode der kleinsten Quadraten

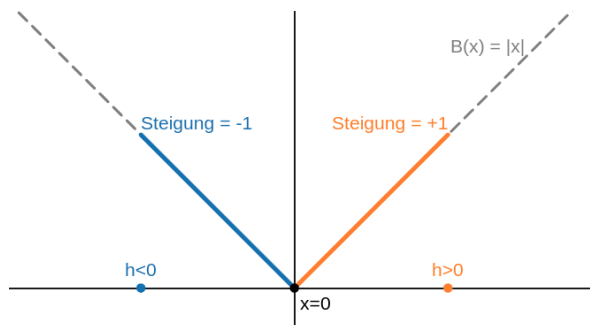
Reflektion

Für eine Funktion $f: (a, b) \rightarrow \mathbb{R}$ haben wir Kandidaten für Minimum bei $x_0 \in (a, b)$ falls gilt

$$\frac{d}{dx} f(x_0) = 0.$$

Darf ich *immer* diese Bedingung benutzen?

Beispiel Für $B(x) = |x|$ liegt das Minimum offensichtlich bei $x = 0$.



Definition der Ableitung

$$\frac{d}{dx} f(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h}$$

($h \neq 0$ ist implizit in der Schreibweise $h \rightarrow 0$) Hintergrund zu Limiten (Grenzwerte) und ihre Rolle in der Differenziation findet man in

E. Cramer and J. Nešlehová. *Vorkurs Mathematik*. 2018, Abschnitte 10.1 – 10.3.

Für unsere B haben wir

$$\frac{d}{dx} B(x) = \lim_{h \rightarrow 0} \frac{B(x+h) - B(x)}{h} = \lim_{h \rightarrow 0} \frac{|x+h| - |x|}{h}$$

Berechne den Bruch für $x = x_0 = 0$, $h \neq 0$

$$\frac{B(x_0+h) - B(x_0)}{h} = \frac{|h| - 0}{h} = \frac{|h|}{h} = \begin{cases} 1 & \text{falls } h > 0 \\ -1 & \text{falls } h < 0 \end{cases}$$

- $h < 0 \Rightarrow$ es ex. ein $\varepsilon > 0$ mit $h = -\varepsilon < 0$ also

$$\frac{f(x_0+h) - f(x_0)}{h} = \frac{f(x_0 - \varepsilon) - f(x_0)}{-\varepsilon} = -\frac{f(x_0) - f(x_0 - \varepsilon)}{\varepsilon} = \frac{f(x_0) - f(x_0 - \varepsilon)}{\varepsilon}$$

- Dagegen ergibt sich bei der *symmetrischen Variante* des Differenzenquotienten

$$\frac{f(x_0 + h) - f(x_0 - h)}{2h} \stackrel{f=B}{=} \frac{|x_0 + h| - |x_0 - h|}{2h} \stackrel{x_0=0, h>0}{=} \frac{h - h}{2h} = 0$$

- In $x = 0$ existiert keine Ableitung von $B(x)$
- Aber die Funktion $B(x) = |x|$ ist konvex (Übung)
 - Das bietet einen alternativen Weg, das Problem anzugreifen!
 - Dies ist ein Thema der Vorlesungen *konvexe Analysis* und *Optimierung*.

Zürück zu Aufgabe 5 - Quadratischer Fehler mit 2 Variablen

Die Fehlerfunktion war gegeben durch

$$Q(a, c) = \sum_{j=1}^N (at_j + c - r_j)^2$$

- Wir berechnen die Ableitung von Q nach der Variable a und der Variable c und suchen einen Punkt (a, c) (oder mehrere solche!) an denen sowohl

$$\frac{d}{dc} Q(a, c) = 0 \tag{5}$$

als auch

$$\frac{d}{da} Q(a, c) = 0 \tag{6}$$

simultan gelten.

- Zunächst zur Ableitung in (5):

$$\begin{aligned} \frac{d}{dc} Q(a, c) &= \frac{d}{dc} \sum_{j=1}^N (at_j + c - r_j)^2 \\ &= \sum_{j=1}^N \frac{d}{dc} (at_j + c - r_j)^2 \\ &= \sum_{j=1}^N 2(at_j + c - r_j) \frac{d}{dc} c \\ &= \sum_{j=1}^N 2(at_j + c - r_j) \cdot 1 \\ &= \sum_{j=1}^N 2(at_j + c - r_j) \end{aligned}$$

- Die Bedingung (5) ist äquivalent zu

$$\begin{aligned}
0 &= \sum_{j=1}^N (at_j + c - r_j) \\
&\Updownarrow \\
0 &= \sum_{j=1}^N at_j + \sum_{j=1}^N c - \sum_{j=1}^N r_j \\
&\Updownarrow \\
0 &= \frac{1}{N} \sum_{j=1}^N at_j + \frac{1}{N} \sum_{j=1}^N c - \frac{1}{N} \sum_{j=1}^N r_j \\
&\Updownarrow \\
0 &= a\bar{t} + c - \bar{r}
\end{aligned}$$

wobei $\bar{t} = \frac{1}{N} \sum_{j=1}^N t_j$, $\bar{r} = \frac{1}{N} \sum_{j=1}^N r_j$.

- Und zur Ableitung (6) nach a :

$$\begin{aligned}
\frac{d}{da} Q(a, c) &= \frac{d}{da} \sum_{j=1}^N (at_j + c - r_j)^2 \\
&= \sum_{j=1}^N \frac{d}{da} (at_j + c - r_j)^2 \\
&= \sum_{j=1}^N 2(at_j + c - r_j) \frac{d}{da} (at_j + c - r_j) \\
&= 2 \sum_{j=1}^N (at_j + c - r_j) t_j \\
&= 2 \sum_{j=1}^N (at_j^2 + ct_j - r_j t_j) \\
&= 2 \left(a \sum_{j=1}^N t_j^2 + c \sum_{j=1}^N t_j - \sum_{j=1}^N r_j t_j \right) \\
&= 2(a\bar{k} + c\bar{t} - \bar{m})
\end{aligned}$$

wobei $m_j = r_j t_j$, $\bar{m} = \frac{1}{N} \sum_{j=1}^N m_j$, $k_j = t_j^2$, $\bar{k} = \frac{1}{N} \sum_{j=1}^N k_j$

- Also ist Bedingung (6) äquivalent zu

$$0 = a\bar{k} + c\bar{t} - \bar{m}$$

- Zusammengefasst: in einem Punkt (a, c) sind beide Ableitungen genau dann gleich Null, wenn (a, c) das Gleichungssystem

$$\begin{aligned}
0 &= a\bar{t} + c - \bar{r} \\
0 &= a\bar{k} + c\bar{t} - \bar{m}
\end{aligned} \tag{7}$$

löst.

Reflektion

Die Werte $\bar{t}$, $\bar{r}$, $\bar{k}$, $\bar{m}$ sind alle aus den uns zu Verfügung stehenden Messdaten berechenbar. Sie hängen nicht von den Parametern a und c ab. Daher ist (7) ein Lineargleichungssystem in 2 Variablen: a und c .

- Multipliziere die erste Gleichung mit $\bar{t}$ (möglich, da $\bar{t} \neq 0$, bei sinnvoller Wahl der Messenzeiten) und bekomme

$$0 = (\bar{t}^2 - \bar{k})a + \bar{m} - \bar{r}\bar{t}$$

und falls $\bar{t}^2 - \bar{k} \neq 0$ dürfen wir die Gleichung umstellen:

$$a = \frac{\bar{m} - \bar{r}\bar{t}}{\bar{k} - \bar{t}^2} \quad (\mathbf{a_0})$$

Was dann ergibt, nach einsetzen von diesem Wert von a in der ersten Gleichung in (7)

$$c = -\frac{\bar{m} - \bar{r}\bar{t}}{\bar{k} - \bar{t}^2}\bar{t} + \bar{r} \quad (\mathbf{c_0})$$

Déjà vu

$$\text{Var}(t) = \bar{k} - \bar{t}^2 = \frac{1}{N} \sum_{j=1}^N t_j^2 - \left(\frac{1}{N} \sum_{j=1}^N t_j \right)^2$$

ist wieder die Varianz vom Datensatz $(t_1, \dots, t_N)$

- Es stellt sich heraus, dass gilt

$$\text{Var}(t) \geq 0 \text{ für sämtliche Datenvektoren}$$

und

$$\text{Var}(t) = 0 \text{ genau dann, wenn sämtliche Einträge gleich sind.}$$

Das letzte bedeutet $t_1 = t_2 = \dots = t_N$, also sämtliche Messungen fanden zum selben Zeitpunkt statt, was in der Realität nicht der Fall ist.

- Die behaupteten Ungleichungen sieht man folgendermassen ein: Offensichtlich gilt

$$\begin{aligned} \text{Var}(t) &= \bar{k} - (\bar{t})^2 = \bar{k} - 2\bar{t} \cdot \bar{t} + (\bar{t})^2 \\ &= \frac{1}{N} \sum_{j=1}^N t_j^2 - 2\frac{1}{N} \sum_{j=1}^N t_j \cdot \bar{t} + \frac{1}{N} \sum_{j=1}^N (\bar{t})^2 \\ &= \frac{1}{N} \sum_{j=1}^N [t_j^2 - 2t_j \cdot \bar{t} + (\bar{t})^2] \\ &= \frac{1}{N} \sum_{j=1}^N [t_j - \bar{t}]^2 \begin{cases} \geq 0 \text{ als Summe von Quadraten} \\ > 0 \text{ sobald für ein } j \text{ gilt } t_j - \bar{t} \neq 0 \end{cases} \end{aligned}$$