

Vorkurs Mathematik

im Wintersemester 2024

Kapitel I – Methode der kleinsten Quadraten

Nachweis des Minimums anhand der quadratischen Ergänzung

Wir zeigen mit einer anderen Methode, dass in dem schon berechneten Punkt (a_0, c_0) tatsächlich ein Minimum vorliegt. Diese Methode kommt vollständig ohne Ableitung aus.

Da wir den Punkt (a_0, c_0) gerade so bestimmt haben, dass dort beide erste Ableitungen verschwinden, gilt

$$D_a := 2 \sum_{j=1}^N (a_0 t_j^2 + c_0 t_j - r_j t_j) = 0$$

$$D_c := 2 \sum_{j=1}^N (a_0 t_j + c_0 - r_j) = 0$$

Dass sich durch Einsetzen der bestimmten Werte (a_0, c_0) in der beiden Gleichungen Null ergibt lässt sich direkt zeigen (Übungsblatt 6, Aufgabe 2).

Wir wollen nun für beliebige $(a, c) \in \mathbb{R}^2$ die Funktionswerte $Q(a, c)$ und $Q(a_0, c_0)$ vergleichen. Die binomische Formel liefert:

$$\begin{aligned} (at_1 + c - r_1)^2 &= [(at_1 + c - a_0 t_1 - c_0) + (a_0 t_1 + c_0 - r_1)]^2 \\ &= (at_1 + c - a_0 t_1 - c_0)^2 + (a_0 t_1 + c_0 - r_1)^2 + 2(at_1 + c - a_0 t_1 - c_0)(a_0 t_1 + c_0 - r_1) \end{aligned}$$

Analog gilt für alle $j = 2, 3, \dots, N$. Wir erhalten:

$$(at_j + c - r_j)^2 = (at_j + c - a_0 t_j - c_0)^2 + (a_0 t_j + c_0 - r_j)^2 + 2(at_j + c - a_0 t_j - c_0)(a_0 t_j + c_0 - r_j)$$

Addiere alle N Gleichungen auf:

$$\begin{aligned} Q(a, c) &= \sum_{j=1}^N (at_j + c - r_j)^2 \\ &= \sum_{j=1}^N [(at_j + c - a_0 t_j - c_0)^2 + (a_0 t_j + c_0 - r_j)^2 + 2(at_j + c - a_0 t_j - c_0)(a_0 t_j + c_0 - r_j)] \\ &= \sum_{j=1}^N (at_j + c - a_0 t_j - c_0)^2 + \underbrace{\sum_{j=1}^N (a_0 t_j + c_0 - r_j)^2}_{Q(a_0, c_0)} + \\ &\quad + 2 \underbrace{\sum_{j=1}^N t_j (a_0 t_j + c_0 - r_j)}_{(a-a_0)D_a} + 2(c-c_0) \underbrace{\sum_{j=1}^N (a_0 t_j + c_0 - r_j)}_{(c-c_0)D_c} \end{aligned}$$

Die letzte beide Terme sind Null, da $D_a = D_c = 0$. Also erhalten wir insgesamt

$$Q(a, c) = Q(a_0, c_0) + \underbrace{\sum_{j=1}^N (at_j + c - a_0t_j - c_0)^2}_{\geq 0 \text{ (Summe von Quadraten)}} \geq Q(a_0, c_0)$$

Reflektion

Wir haben ganz ohne Benutzung von Ableitungen gezeigt, dass in (a_0, c_0) ein Minimum vorliegt. Das lag daran, dass die Funktion $(a, c) \mapsto Q(a, c)$ eine sehr spezielle Struktur hat, insbesondere, dass nur quadratische Terme auftauchen. Diese spezielle Funktion kann man sowohl mit den Methoden aus der "Analysis" (Stichwort Ableitung) als auch mit den aus der "Algebra" (Stichwort Binomische Formel) untersuchen.

Frage: Ist (a_0, c_0) das einzige Minimum von Q ?

Zwei Möglichkeiten dies zu beantworten:

1. Mit Vorwissen aus der Schule und etwas Beweislust kann man nachweisen, dass

$$\begin{cases} \frac{d}{da}Q = 0 \\ \frac{d}{dc}Q = 0 \end{cases}$$

tasächlich notwendige Bedingungen für ein Minimum sind. Diese sind **nur** im Punkt (a_0, c_0) erfüllt. Wenn es überhaupt ein Minimum gibt, dann ist es in (a_0, c_0) .

2. Mit Widerspruchsbeweis: angenommen es existiert ein weiteres Minimum (a_*, c_*) mit $Q(a_0, c_0) = Q(a_*, c_*)$. Dann gilt

$$\sum_{j=1}^N (a_*t_j - a_0t_j + c_* - c_0)^2 = 0$$

Damit eine Summe von Quadraten Null ergibt, müssen alle Summanden einzeln Null sein. Dies bedeutet:

$$\begin{cases} c_* - c_0 = a_0t_1 - a_*t_1 = t_1(a_0 - a_*) \\ c_* - c_0 = a_0t_2 - a_*t_2 = t_2(a_0 - a_*) \\ \vdots \\ c_* - c_0 = a_0t_N - a_*t_N = t_N(a_0 - a_*) \end{cases}$$

- o Falls $a_* \neq a_0$ darf ich durch $a_0 - a_*$ dividieren:

$$\begin{cases} \frac{c_* - c_0}{a_0 - a_*} = t_1 \\ \frac{c_* - c_0}{a_0 - a_*} = t_2 \\ \vdots \\ \frac{c_* - c_0}{a_0 - a_*} = t_N \end{cases}$$

also sind alle Messzeiten gleich, im Widerspruch zu unserer (Modell-)Annahme.

- o Falls $a_0 = a_*$ so folgt

$$0 = a_0t_1 - a_*t_1 = c_* - c_0 \Rightarrow c_0 = c_*$$

also sind (a_*, c_*) und (a_0, c_0) keine unterschiedliche Punkte.

Kapitel II. Der Vektorraum $\mathbb{R}^n$

Schon aus der Schule kennt man die Ebene $\mathbb{R}^2$, also den 2-dimensionalen Raum, und ebenso kennt man $\mathbb{R}^3$, den 3-dim. Raum.

Elemente von $\mathbb{R}^2$ bezeichnet man sowohl mit Zeilenvektoren (x_1, x_2) oder (x, y) oder (a, b) oder (u, v) als auch mit Spaltenvektoren $\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ oder $\begin{pmatrix} x \\ y \end{pmatrix}$ oder $\begin{pmatrix} a \\ b \end{pmatrix}$ oder $\begin{pmatrix} u \\ v \end{pmatrix}$.

Man sagt, dass diese Spaltenvektoren die "Transponierte Vektoren" zu den entsprechenden obigen Zeilenvektoren sind, und umgekehrt. Geschrieben wird

$$(x, y)^T = \begin{pmatrix} x \\ y \end{pmatrix} \text{ sowie } \begin{pmatrix} a \\ b \end{pmatrix}^T = (a, b)$$

Vektoren werden oft (z.B. in der Physik) mit Pfeilen bezeichnet, z.B. $\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ und die Zahlen $x_1, x_2 \in \mathbb{R}$ heißen die erste bzw. zweite Komponente von $\vec{x}$.

In Mathematikvorlesung werden die Pfeile normalerweise weggelassen, und wir schreiben einfach etwa

$$v = (v_1, v_2, v_3) \in \mathbb{R}^3$$

Oft trifft man Spalten (oder Zeilen) mit mehr als 2 oder 3 Komponenten an, z.B. in Excel Tabellen.

$$\begin{array}{ccc} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \\ x_4 & y_4 & z_4 \\ x_5 & y_5 & z_5 \end{array}$$

hier können die Spalten $(x_1, \dots, x_5)$ und $(y_1, \dots, y_5)$ vorgegeben sein, und die Elemente der Spalte $z_1, \dots, z_5$ als Summe definiert sein, d.h. $z_i = x_i + y_i$.

Dann ist der Vektor $z \in \mathbb{R}^5$ die Summe $x + y$ der Vektoren $x, y \in \mathbb{R}^5$:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \\ x_4 + y_4 \\ x_5 + y_5 \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{pmatrix}$$

Definition 1. (Vektoren in $\mathbb{R}^n$)

Ein Vektor mit n reellen Komponenten ist ein n -Tupel reeller Zahlen, darstellbar als

$$\vec{v} = (v_1, \dots, v_n) = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}$$

Für Vektoren in $\mathbb{R}^n$ definieren wir zwei Operationen

Definition 2. (Vektoraddition und Skalarmultiplikation)

Seien

$$\vec{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{R}^n$$

Die *Summe* $\vec{u} + \vec{v}$ ist definiert als

$$\vec{u} + \vec{v} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \vdots \\ u_n + v_n \end{pmatrix} \in \mathbb{R}^n$$

Sei nun $\lambda \in \mathbb{R}$ eine reelle Zahl. Die *Skalarmultiplikation*⁰ ist definiert als

$$\lambda \vec{v} = \lambda \cdot \vec{v} = \lambda \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \begin{pmatrix} \lambda v_1 \\ \lambda v_2 \\ \vdots \\ \lambda v_n \end{pmatrix} \in \mathbb{R}^n.$$

⁰Nicht mit *Skalarprodukt* zu verwechseln!