Interpolation, lineare Gleichungen (mit und ohne Lösungen) und lineare Regression

Franz Pauer
Institut für Mathematik, Universität Innsbruck
Technikerstr. 13/7, A-6020 Innsbruck, Österreich
franz.pauer@uibk.ac.at

1 Einleitung

„Lineare Regression“ ist ein Thema des Lehrplans von berufsbildenden höheren Schulen (4. und 5. Jahrgang der HTL, 2. und 5. Jahrgang der HAK). Es könnte auch in der AHS als interessante Anwendung von Kenntnissen über lineare Gleichungen und über das Skalarprodukt besprochen werden.

Leider wird in den mir bekannten Schulbüchern (zum Beispiel [1] oder [2]) darauf verzichtet zu erklären, wie die Gleichung der Regressionsgeraden berechnet werden kann und warum man dazu gerade die Summe der „Fehlerquadrate“ minimiert. Auch in [3] wird zur zweiten Frage nur auf rechentechnische bzw. historische Gründe verwiesen. Dort wird aber erläutert, wie man die Gleichung der Regressionsgeraden mit Hilfe des Skalarprodukts leicht herleiten kann.

In diesem Beitrag betrachten wir zuerst Interpolationsaufgaben, die durch Systeme linearer Gleichungen $Ax = b$ beschrieben werden können (Abschnitt 2). Wenn diese Gleichungssysteme keine Lösung haben (Abschnitt 3), ersetzen wir die Spalte b durch die „am nächsten liegende“ Spalte, mit der das Gleichungssystem eine Lösung hat. Dieser Vorgang entspricht der linearen Regression. Um zu präzisieren, was „am nächsten liegend“ bedeutet, müssen wir festlegen, was wir unter dem Abstand zweier Spalten verstehen. Verwenden wir dazu den euklidischen Abstand, landen wir bei der „Methode der kleinsten Quadrate“.

Im Abschnitt 4 wird dargestellt, wie man auf einfache Weise mit Hilfe des Skalarprodukts die Gleichung der Regressionsgeraden erhält. Da in der Handelsakademie das Skalarprodukt nicht Pflichtstoff ist, wird im Abschnitt 5 die Gleichung der Regressionsgeraden ohne Verwendung des Skalarprodukts hergeleitet. Abschnitt 6 geht auf die Herleitung dieser Gleichung mit Hilfe der Differenzialrechnung ein (und weist darauf hin, dass man diese dazu eigentlich nicht braucht).

2 Interpolationsaufgaben

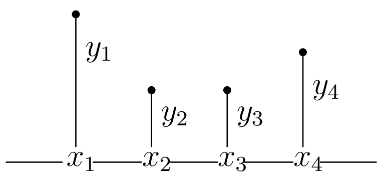
Wir betrachten die folgenden *Interpolationsaufgaben*:
Gegeben sind

- Funktionen $f_1, \ldots, f_k$ von $\mathbb{R}$ nach $\mathbb{R}$,
- paarweise verschiedene reelle Zahlen $x_1, \ldots, x_n \in \mathbb{R}$ und
- reelle Zahlen $y_1, \ldots, y_n \in \mathbb{R}$.

Gesucht sind reelle Zahlen $c_1, \ldots, c_k$ so, dass die Funktion $f := \sum_{i=1}^{k} c_i f_i$ die Bedingungen

$$f(x_1) = y_1,\ f(x_2) = y_2,\ \ldots,\ f(x_n) = y_n$$

erfüllt.



Durch die Funktionen $f_1, \ldots, f_k$ wird der „Typ“ der Interpolationsaufgabe vorgegeben. Die reellen Zahlen $x_1, \ldots, x_n$ heißen *Stützstellen*, die reellen Zahlen $y_1, \ldots, y_n$ *Werte* der Interpolationsaufgabe. Die gesuchte Funktion f heißt *interpolierende Funktion*.

Wir suchen also eine Funktion f des vorgegebenen Typs so, dass die Funktionswerte von f in den Stützstellen die vorgegebenen Werte der Interpolationsaufgabe sind.

Anders formuliert: Wir suchen Zahlen $c_1, \ldots, c_k$ so, dass

$$\begin{array}{rcl}
f_1(x_1)c_1 + f_2(x_1)c_2 + \ldots + f_k(x_1)c_k & = & y_1 \\
f_1(x_2)c_1 + f_2(x_2)c_2 + \ldots + f_k(x_2)c_k & = & y_2 \\
\ldots & = & \ldots \\
f_1(x_k)c_1 + f_2(x_k)c_2 + \ldots + f_k(x_n)c_k & = & y_n
\end{array}$$

ist. Das ist ein System von k linearen Gleichungen mit n Unbekannten $c_1, \ldots, c_n$.

Beispiel 1. Interpolation durch lineare Funktionen.
Wenn f_1 die Identität (also die Funktion, die jeder Zahl sich selbst zuordnet) und f_2 die konstante Funktion 1 (also die Funktion, die jeder Zahl die Zahl 1 zuordnet) ist, dann suchen wir eine Funktion $f := c_1 f_1 + c_2 f_2$ mit

$$(f(x_i) =)\ c_1 x_i + c_2 = y_i\,,\ 1 \leq i \leq n\,.$$

In diesem Fall schreiben wir wie üblich k statt c_1 und d statt c_0. Dann ist f die Funktion, die jeder reellen Zahl z die Zahl $f(z) := k.z + d$ zuordnet.

Um die Zahlen k und d mit den Eigenschaften

$$\begin{array}{rcl} k.x_1 + d.1 & = & y_1 \\ \vdots & \vdots & \vdots \\ k.x_n + d.1 & = & y_n \end{array}$$

zu finden, müssen wir ein System von n linearen Gleichungen mit zwei Unbekannten lösen.

Beispiel 2. Interpolation durch Polynomfunktionen.
Wenn die gegebenen Funktionen $f_1, ..., f_k$ die Potenzfunktionen vom Grad 0 bis $k-1$ sind, dann ist die gesuchte Funktion f eine Polynomfunktion vom Grad $k-1$, also $f : \mathbb{R} \longrightarrow \mathbb{R}$, $z \longmapsto c_1 + c_2.z + \ldots + c_k.z^{k-1}$.

Wir suchen reelle Zahlen $c_1, c_2, \ldots, c_k$ mit der Eigenschaft, dass

$$\begin{array}{rcl} c_1 + x_1c_2 + \ldots + x_1^{k-1}c_k & = & y_1 \\ \vdots & \vdots & \vdots \\ c_1 + x_nc_2 + \ldots + x_n^{k-1}c_k & = & y_n \end{array}$$

ist, müssen also ein System von n Gleichungen mit k Unbekannten lösen.

Beispiel 3. Interpolation durch Linearkombinationen von Sinus und Cosinus.
Ist f_1 die Sinusfunktion und f_2 die Cosinusfunktion, dann suchen wir eine (2π-periodische) Funktion $f := c_1 \sin + c_2 \cos$ mit

$$(f(x_i) =)\ c_1 \sin(x_i) + c_2 \cos(x_i) = y_i\,,\ 1 \leq i \leq n\,.$$

Wir suchen daher zwei reelle Zahlen c_1 und c_2 so, dass

$$\begin{array}{rcl} c_1 \sin(x_1) + c_2 \cos(x_1) & = & y_1 \\ \vdots & \vdots & \vdots \\ c_1 \sin(x_n) + c_2 \cos(x_n) & = & y_n \end{array}$$

ist. Um sie zu berechnen, müssen wir ein System von n linearen Gleichungen mit zwei Unbekannten lösen.

3 Systeme linearer Gleichungen (mit und ohne Lösung)

Das System

$$\begin{array}{rcl} f_1(x_1)c_1 + f_2(x_1)c_2 + \ldots + f_k(x_1)c_k & = & y_1 \\ f_1(x_2)c_1 + f_2(x_2)c_2 + \ldots + f_k(x_2)c_k & = & y_2 \\ \ldots & = & \ldots \\ f_1(x_k)c_1 + f_2(x_k)c_2 + \ldots + f_k(x_n)c_k & = & y_n \end{array}$$

von k linearen Gleichungen mit den Unbekannten $c_1, \ldots, c_n$ kann auch in der Form

$$c_1 \begin{pmatrix} f_1(x_1) \\ \vdots \\ f_1(x_n) \end{pmatrix} + \ldots + c_k \begin{pmatrix} f_k(x_1) \\ \vdots \\ f_k(x_n) \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix},$$

oder, mit den Abkürzungen

$$\mathbf{y} := \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \quad \text{und} \quad \mathbf{f_i}(\mathbf{x}) := \begin{pmatrix} f_i(x_1) \\ \vdots \\ f_i(x_n) \end{pmatrix},\ 1 \leq i \leq k,$$

kurz als

$$\sum_{i=1}^{k} c_i \mathbf{f_i}(\mathbf{x}) = \mathbf{y}$$

angeschrieben werden.

Diese Schreibweise legt es nahe, das System linearer Gleichungen als die folgende Aufgabe zu interpretieren: Schreibe die Spalte $\mathbf{y}$ als Linearkombination $\sum_{i=1}^{k} c_i \mathbf{f_i}(\mathbf{x})$ der Spalten $\mathbf{f_i}(\mathbf{x})$, $1 \leq i \leq k$. Das ist aber nur dann möglich, wenn $\mathbf{y}$ in dem von den Spalten $\mathbf{f_i}(\mathbf{x})$, $1 \leq i \leq k$, erzeugten Untervektorraum U (des Vektorraums aller Spalten) enthalten ist. Wenn das nicht der Fall ist, ist dieses System linearer Gleichungen nicht lösbar.

Im Fall von Beispiel 1 (Interpolation durch lineare Funktionen) ist U die von

$$\mathbf{1} := \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \quad \text{und} \quad \mathbf{x} := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

erzeugte Ebene.

Im Fall von Beispiel 2 (Interpolation durch Polynomfunktionen vom Grad $\leq k$) ist U der von

$$\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \ldots, \begin{pmatrix} x_1^{k-1} \\ \vdots \\ x_n^{k-1} \end{pmatrix}$$

erzeugte Vektorraum.

Im Fall von Beispiel 3 (Interpolation durch Linearkombinationen von Sinus und Cosinus) ist U die von

$$\begin{pmatrix} \sin(x_1) \\ \vdots \\ \sin(x_n) \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} \cos(x_1) \\ \vdots \\ \cos(x_n) \end{pmatrix}$$

erzeugte Ebene.

Wenn die Interpolationsaufgabe einerseits eine Situation beschreibt, von der man weiß , dass es eine Lösung gibt, andererseits die Aufgabe aber nicht lösbar ist, weil $\mathbf{y}$ mit Mess- oder Rundungsfehlern behaftet ist, liegt es nahe, dass $\mathbf{y}$ „eigentlich“ ein Element von U sein sollte. Wir erzwingen die Lösbarkeit der Aufgabe, indem wir $\mathbf{y}$ durch eine Spalte $\mathbf{y}'$ in U ersetzen!

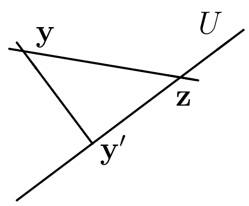
Wie sollen wir diese Spalte $\mathbf{y}'$ aber wählen? Am einfachsten ist es, $\mathbf{y}'$ in U so zu wählen, dass der Abstand zwischen $\mathbf{y}'$ und $\mathbf{y}$ möglichst klein ist. Wir suchen also ein Element des Vektorraums (im Fall von Beispiel 1 der Ebene) U so, dass der Abstand $\|\mathbf{y}' - \mathbf{y}\|$ zwischen $\mathbf{y}$ und $\mathbf{y}'$ so klein wie möglich ist. Wir müssen uns nun aber entscheiden, welchen Abstand wir meinen: wir wählen den (durch das Standardskalarprodukt auf dem $\mathbb{R}^n$ induzierten) euklidischen Abstand im $\mathbb{R}^n$, also

$$\|\mathbf{y}' - \mathbf{y}\| := \sqrt{\sum_{i=1}^{n} (y_i' - y_i)^2}\,.$$

Für positive reelle Zahlen a und b ist $a \leq b$ genau dann, wenn $a^2 \leq b^2$ ist. Daher ist der $\|\mathbf{y}' - \mathbf{y}\|$ genau dann minimal, wenn die Summe $\sum_{i=1}^{n}(y_i' - y_i)^2$ der „Fehlerquadrate“ minimal ist.

4 Lineare Regression (wenn das Skalarprodukt bekannt ist)

Genau dann ist der Abstand zwischen $\mathbf{y}'$ und $\mathbf{y}$ kleiner oder gleich dem Abstand zwischen $\mathbf{y}$ und jedem anderen Element $\mathbf{z}$ von U, wenn die Gerade durch $\mathbf{y}'$ und $\mathbf{y}$ normal zum Untervektorraum U steht. Das folgt leicht aus dem Satz von Pythagoras:



Das Dreieck mit den Eckpunkten $\mathbf{y}'$, $\mathbf{y}$ und $\mathbf{z}$ hat bei $\mathbf{y}'$ einen rechten Winkel. Der Abstand zwischen $\mathbf{y}$ und $\mathbf{z}$ ist die Länge der Hypotenuse, also größer oder gleich der Länge einer Kathete, also dem Abstand zwischen $\mathbf{y}$ und $\mathbf{y}'$.

Das Skalarprodukt zweier Spalten

$$\mathbf{a} := \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \quad \text{und} \quad \mathbf{b} := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

ist die Zahl

$$\mathbf{a} \cdot \mathbf{b} := \sum_{i=1}^{n} a_i b_i \,.$$

Die Gerade durch $\mathbf{y}'$ und $\mathbf{y}$ steht genau dann normal zu U, wenn alle Skalarprodukte von $\mathbf{y} - \mathbf{y}'$ mit den erzeugenden Spalten von U gleich 0 sind. Für $\mathbf{y}' = \sum_{i=1}^{k} c_i \mathbf{f_i}(\mathbf{x}) \in U$ muss also gelten:

$$(\mathbf{y} - \mathbf{y}') \cdot \mathbf{f_j}(\mathbf{x}) = 0,\ 1 \leq j \leq k\,.$$

Anders geschrieben:

$$\sum_{i=1}^{k} c_i(\mathbf{f_i}(\mathbf{x}) \cdot \mathbf{f_j}(\mathbf{x})) = \mathbf{y} \cdot \mathbf{f_j}(\mathbf{x}),\ 1 \leq j \leq k\,.$$

Wenn wir dieses System von k linearen Gleichungen mit k Unbekannten $c_1, \ldots, c_k$ lösen, dann erhalten wir die „annähernd" interpolierende Funktion $f = \sum_{i=1}^{k} c_i f_i$.

Im Fall von Beispiel 1 (lineare Interpolation) ist

- $\mathbf{y}' = k.\mathbf{x} + d.\mathbf{1} \in U$ und
- die Gerade durch $\mathbf{y}$ und $\mathbf{y}'$ steht normal auf der von $\mathbf{x}$ und $\mathbf{1}$ erzeugten Ebene U.

Also ist

- $(\mathbf{y} - (k.\mathbf{x} + d.\mathbf{1})) \cdot \mathbf{x} = 0$ und
- $(\mathbf{y} - (k.\mathbf{x} + d.\mathbf{1})) \cdot \mathbf{1} = 0$.

Daraus erhalten wir das folgende System von zwei linearen Gleichungen mit zwei Unbekannten k und d:

- $k(\mathbf{x} \cdot \mathbf{x}) + d(\mathbf{1} \cdot \mathbf{x}) = \mathbf{x} \cdot \mathbf{y}$
- $k(\mathbf{x} \cdot \mathbf{1}) + d(\mathbf{1} \cdot \mathbf{1}) = \mathbf{1} \cdot \mathbf{y}$

Als Lösung erhalten wir

$$k = \frac{n(\mathbf{x}\cdot\mathbf{y}) - (\mathbf{1}\cdot\mathbf{x}).(\mathbf{1}\cdot\mathbf{y})}{n(\mathbf{x}\cdot\mathbf{x}) - (\mathbf{1}\cdot\mathbf{x})^2} \quad \text{und} \quad d = \frac{(\mathbf{x}\cdot\mathbf{x}).(\mathbf{1}\cdot\mathbf{y}) - (\mathbf{1}\cdot\mathbf{x}).(\mathbf{x}\cdot\mathbf{y})}{n.(\mathbf{x}\cdot\mathbf{x}) - (\mathbf{1}\cdot\mathbf{x})^2} \,.$$

Wegen $\mathbf{x}\cdot\mathbf{y} = \sum_{i=1}^n x_i y_i$, $\mathbf{1}\cdot\mathbf{x} = \sum_{i=1}^n x_i$, $\mathbf{1}\cdot\mathbf{y} = \sum_{i=1}^n y_i$, $\mathbf{1}\cdot\mathbf{1} = n$, $\mathbf{x}\cdot\mathbf{x} = \sum_{i=1}^n x_i^2$ und $\mathbf{y}\cdot\mathbf{y} = \sum_{i=1}^n y_i^2$ ist

$$k = \frac{n\sum_{i=1}^n x_i y_i - (\sum_{i=1}^n x_i)(\sum_{i=1}^n y_i)}{n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

und

$$d = \frac{(\sum_{i=1}^n x_i^2)(\sum_{i=1}^n y_i) - (\sum_{i=1}^n x_i)(\sum_{i=1}^n x_i y_i)}{n\sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2} \,.$$

Wir haben damit die Funktion $f : \mathbb{R} \longrightarrow \mathbb{R},\ z \longmapsto k.z + d$, so bestimmt, dass der (euklidische) Abstand vom n-Tupel der gegebenen (gemessenen oder gerundeten) ungenauen Funktionswerte $(y_1, \ldots, y_n)$ zum n-Tupel der berechneten Funktionswerte $(f(x_1), \ldots, f(x_n))$ möglichst klein ist, also
$\sum_{i=1}^n (y_i - (kx_i + d))^2$ möglichst klein ist. Der Graph dieser Funktion heißt *Regressionsgerade* oder *Trendlinie*.

Dazu mussten wir nur das Skalarprodukt und das Lösen von Systemen von zwei linearen Gleichungen mit zwei Unbekannten als bekannt voraussetzen.

5 Lineare Regression, wenn das Skalarprodukt noch nicht bekannt ist

Die Aufgabe
„Gegeben sind paarweise verschiedene reelle Zahlen $x_1, \ldots, x_n \in \mathbb{R}$ und reelle Zahlen $y_1, \ldots, y_n \in \mathbb{R}$. Berechne eine Funktion $f : \mathbb{R} \longrightarrow \mathbb{R},\ z \longmapsto k.z + d$, so , dass der (euklidische) Abstand zwischen $(y_1, \ldots, y_n)$ und $(f(x_1), \ldots, f(x_n))$ möglichst klein ist.“
kann auch ohne Verwendung des Skalarproduktes gelöst werden. Man muss dazu - etwas mühsam, aber ohne wesentliche Schwierigkeiten - einige Rechnungen ausführen.

Wir bestimmen Zahlen k und d so, dass $\sum_{i=1}^n (y_i - (kx_i + d))^2$, das Quadrat des Abstandes von $(y_1, \ldots, y_n)$ zu $(kx_1 + d, \ldots, kx_n + d)$, möglichst klein ist.

Das heißt: k und d werden so gewählt, dass für alle reellen Zahlen r und s

$$\sum_{i=1}^{n}(y_i - ((k+r)x_i + (d+s)))^2 \geq \sum_{i=1}^{n}(y_i - (kx_i + d))^2$$

ist, also die Differenz

$$D(r,s) := \sum_{i=1}^{n}(y_i - ((k+r)x_i + (d+s)))^2 - \sum_{i=1}^{n}(y_i - (kx_i + d))^2 \geq 0$$

ist. Ausrechnen ergibt:

$$D(r,s) = \sum_{i=1}^{n}((y_i - kx_i - d) - (rx_i + s))^2 - \sum_{i=1}^{n}(y_i - kx_i - d))^2 =$$

$$= \sum_{i=1}^{n}(rx_i + s)^2 - 2\sum_{i=1}^{n}(y_i - kx_i - d)(rx_i + s)\,.$$

Als Summe von Quadraten ist die Zahl $\sum_{i=1}^{n}(rx_i + s)^2$ immer ≥ 0 . Wählt man daher k und d so, dass für alle r, s

$$\sum_{i=1}^{n}(rx_i + s)(y_i - kx_i - d) = 0$$

ist, dann ist $D(r,s) \geq 0$ für alle $r, s \in \mathbb{R}$. Wegen

$$\sum_{i=1}^{n}(rx_i + s)(y_i - kx_i - d) = r\sum_{i=1}^{n}(x_i y_i - kx_i^2 - dx_i) + s\sum_{i=1}^{n}(y_i - kx_i - d)$$

ist (k, d) Lösung des Systems

- $\sum_{i=1}^{n}(x_i y_i - kx_i^2 - dx_i) = 0$
- $\sum_{i=1}^{n}(y_i - kx_i - d) = 0$

von zwei linearen Gleichungen in zwei Unbekannten. Dessen Lösung ist dieselbe wie die in Abschnitt 4.

6 Lineare Regression, wenn partielles Differenzieren und seine Anwendung auf Extremwertaufgaben bekannt sind

Gegeben sind paarweise verschiedene reelle Zahlen $x_1, \ldots, x_n \in \mathbb{R}$ und reelle Zahlen $y_1, \ldots, y_n \in \mathbb{R}$. Differenziert man die quadratische Funktion $q := \sum_{i=1}^{n}(y_i - (ux_i + v))^2$ an der Stelle (k, d) partiell nach u und v, erhält man eine Darstellung dieser Funktion in der Form

$$q = q(k,d) + \frac{\partial q}{\partial u}(k,d)(u-k) + \frac{\partial q}{\partial v}(k,d)(v-d) + \frac{\partial^2 q}{\partial u \partial v}(k,d)(u-k)(v-d) +$$

$$+\frac{1}{2}\frac{\partial^2 q}{\partial^2 u}(k,d)(u-k)^2 + \frac{1}{2}\frac{\partial^2 q}{\partial^2 v}(k,d)(v-d)^2 =$$

$$= q(k,d) - 2\sum_{i=1}^{n} x_i(y_i - (kx_i + d))(u-k) - 2\sum_{i=1}^{n}(y_i - (kx_i + d))(u-d) +$$

$$+(2\sum_{i=1}^{n} x_i)(u-k)(v-d) + (\sum_{i=1}^{n} x_i^2)(u-k)^2 + n(v-d)^2 \,.$$

Dieselbe Darstellung der quadratischen Funktion q als Linearkombination von $1, (u-k), (v-d), (u-k)(v-d), (u-k)^2$ und $(v-d)^2$ hätte man auch erhalten, wenn man $\sum_{i=1}^{n}(y_i-(ux_i+v))^2$ als $\sum_{i=1}^{n}(y_i-((k+(u-k))x_i+(d+(v-d))))^2$ geschrieben hätte, dann ausmultipliziert und die Koeffizienten von $1, (u-k)$, $(v-d), (u-k)(v-d), (u-k)^2, (v-d)^2$ zusammengefasst hätte.

Die Funktion q hat in (k, d) ein Minimum, wenn $\frac{\partial q}{\partial u}(k,d) = 0$, $\frac{\partial q}{\partial v}(k,d) = 0$ und die Hesse'sche Matrix

$$\begin{pmatrix} \frac{\partial^2 q}{\partial^2 u}(k,d) & \frac{\partial^2 q}{\partial u \partial v}(k,d) \\ \\ \frac{\partial^2 q}{\partial u \partial v}(k,d) & \frac{\partial^2 q}{\partial^2 v}(k,d) \end{pmatrix} = 2 \cdot \begin{pmatrix} \sum_{i=1}^{n} x_i^2 & \sum_{i=1}^{n} x_i \\ \\ \sum_{i=1}^{n} x_i & n \end{pmatrix}$$

positiv definit ist. Diese Matrix ist bekanntlich genau dann positiv definit, wenn $\sum_{i=1}^{n} x_i^2 > 0$ und $n.\sum_{i=1}^{n} x_i^2 > (\sum_{i=1}^{n} x_i)^2$ ist. Die zweite Ungleichung folgt aus der (quadrierten) Cauchy-Schwarz'schen Ungleichung $\|\mathbf{1}\|^2.\|\mathbf{x}\|^2 > (\mathbf{1} \cdot \mathbf{x})^2$.

Auch dafür braucht man keine Differenzialrechnung: Wenn die Koeffizienten bei $u-k$ und bei $v-d$ gleich 0 sind, ist $q(k,d)$ der kleinste Funktionswert von q genau dann, wenn die (homogene) quadratische Funktion

$$q' := (2\sum_{i=1}^{n} x_i)(u-k)(v-d) + (\sum_{i=1}^{n} x_i^2)(u-k)^2 + n(v-d)^2$$

keine negativen Funktionswerte hat. Wir ergänzen quadratisch und erhalten

$$q' = n.((v-d) + \frac{1}{n}\sum_{i=1}^{n} x_i(u-k))^2 + \frac{1}{n}(n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2)(u-k)^2 .$$

Diese Funktion hat somit genau dann keine negativen Funktionswerte, wenn $n(\sum_{i=1}^{n} x_i^2) - (\sum_{i=1}^{n} x_i)^2 \geq 0$ ist.

Literatur

[1] Tinhof, F. et al.: Mathematik II HAK/LW. 3. Auflage, Trauner Verlag, Linz, 2009.

[2] Timischl, W., Prugger, E.: Mathematik & Wirtschaft 1. 2. Auflage, Dorner Verlag, Wien, 2006.

[3] Götz, S.: Von Pferden, Ziegen und unmöglichen Würfeln. Didaktikhefte 37, 19-46, 2005.