

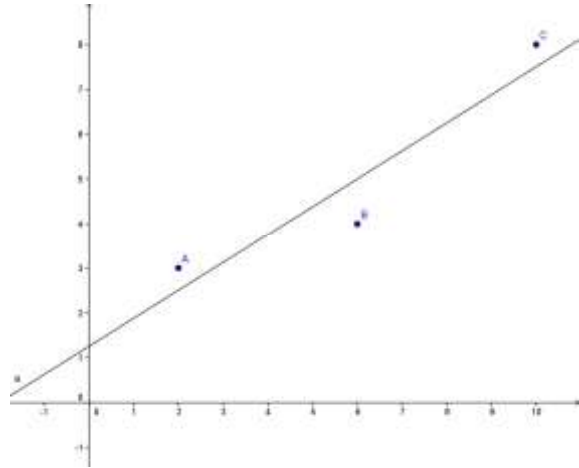
5. Ein Beispiel zur Linearen Regression

Das Problem zu gegebenen Punkten eine optimale Ausgleichsgerade zu finden, kann als Extremalproblem gelöst werden (\rightarrow Analysis \rightarrow Polynomfunktionen \rightarrow Extremalprobleme). Im Folgenden wird an einem Beispiel ein anderer Weg zur Lösung erläutert.

Beispiel:

Gegeben die drei Punkte (2, 3), (6, 4), (10, 8).

In der Abbildung ist die vermutete Ausgleichsgerade dargestellt.



Wir führen den Vektor der x-Koordinaten und den der y-Koordinaten ein:

$$\vec{x} = \begin{pmatrix} 2 \\ 6 \\ 10 \end{pmatrix}, \quad \vec{y} = \begin{pmatrix} 3 \\ 4 \\ 8 \end{pmatrix} \text{ und zusätzlich den Vektor } \vec{e} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Für die Gleichung der Ausgleichsgeraden wählen wir die Form $y = mx + q$

Die Summe der Abweichungsquadrate kann damit als Quadrat des absoluten Betrags eines geeigneten Vektors interpretiert werden:

$$D(m, q) = \sum_{i=1}^3 (y_i - (mx_i + q))^2$$

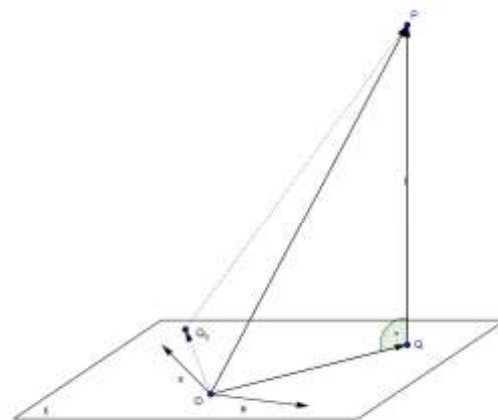
Der Nullpunkt O und die Vektoren \vec{x} und \vec{e} spannen eine Ebene ε auf. Trägt man den Vektor $m\vec{x} + q\vec{e}$ im Nullpunkt ab, so liegt der Endpunkt in einem Ebenenpunkt Q_1

Für den Komponenten des

Verbindungsvektors $\overrightarrow{Q_1P}$ gilt dann

$$\overrightarrow{Q_1P} = \vec{y} - (m\vec{x} + q\vec{e}) \text{ und damit}$$

$$D(m, q) = |\overrightarrow{Q_1P}|^2 = (\vec{y} - (m\vec{x} + q\vec{e}))^2$$



Das Problem ist damit darauf zurückgeführt, den Punkt Q_1 so zu bestimmen, dass $D(m, q)$ minimal wird. Dies ist genau dann der Fall, wenn der in der betrachteten Ebene liegende Punkt Q möglichst nahe bei P liegt. Q muss also gleich dem Fußpunkt des von P gefällten

Lots sein. Dies ist genau dann der Fall, wenn der Verbindungsvektor von Q und P sowohl \vec{e} als auch auf \vec{x} senkrecht steht.

\vec{e} auf als auch auf x senkrecht steht.

$$\begin{aligned} \overrightarrow{QP} \cdot \vec{e} &= 0 & \overrightarrow{QP} \cdot \vec{x} &= 0 \\ \begin{pmatrix} 8 - 10m - q \\ 4 - 6m - q \\ 3 - 2m - q \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} &= 0 & \begin{pmatrix} 8 - 10m - q \\ 4 - 6m - q \\ 3 - 2m - q \end{pmatrix} \cdot \begin{pmatrix} 10 \\ 6 \\ 2 \end{pmatrix} &= 0 \end{aligned}$$

vereinfacht:

$$\begin{aligned} 15 - 10m - 3q &= 0 & \cdot(-6) \text{ und zur zweiten Gleichung addiert} \\ 110 - 140m - 18q &= 0 \end{aligned}$$

$$20 - 32m = 0 \quad \text{und damit} \quad m = \frac{5}{8}.$$

$$\text{Einsetzen in die erste Gleichung ergibt} \quad q = \frac{5}{4}.$$

$$\text{Die Ausgleichsgerade hat also die Gleichung} \quad y = \frac{5}{8}x + \frac{5}{4}$$

Die Lösungsidee kann analog auf 4 und mehr Punkte angewendet werden.