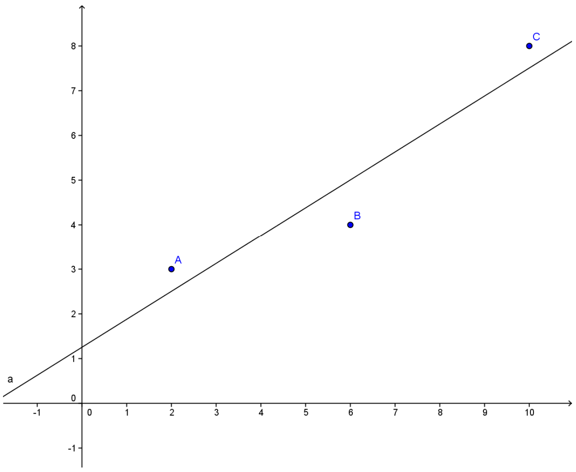
5. Ein Beispiel zur Linearen Regression

Das Problem zu gegebenen Punkten eine optimale Ausgleichsgerade zu finden, kann als Extremalproblem gelöst werden (→ Analysis → Polynomfunktionen → 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.



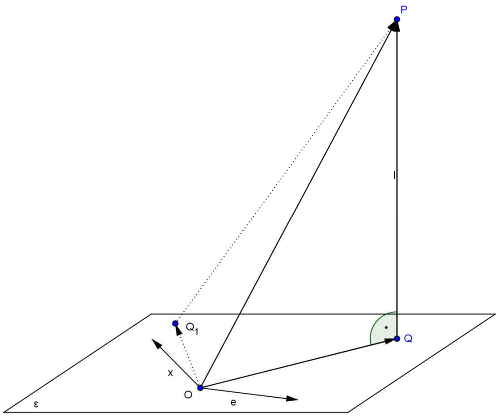
Wir führen den Vektor der x-Koordinaten bzw. der y-Koordinaten ein: $\vec{x} = \begin{pmatrix} 2 \\ 6 \\ 10 \end{pmatrix}$ $\vec{y} = \begin{pmatrix} 3 \\ 4 \\ 8 \end{pmatrix}$

und wählen für die Gleichung der Ausgleichs-geraden den Ansatz: y = mx + q

Führen wir zusätzlich den Vektor $\vec{e} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ ein,

so lässt sich die Summe der Abweichungs-quadrate als Quadrat des absoluten Betrags eines geeigneten Vektors interpretieren:

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



Der Nullpunkt O und die Vektoren $\vec{e}$ und $\vec{x}$ spannen eine Ebene ε auf. Trägt man den Vektor $m\vec{x} + q\vec{e}$ im Nullpunkt ab, so liege der Endpunkt in Q. Für die Komponenten des Verbindungsvektors von Q mit P gilt dann

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

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

Das Problem ist damit darauf zurückgeführt, den Punkt Q 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. D.h. Q muss gleich dem Fusspunkt des von P gefällten Lots sein. Dies ist genau dann der Fall, wenn der Verbindungsvektor von Q und P sowohl auf $\vec{e}$ als auch auf $\vec{x}$ senkrecht steht.

Durchführung für das betrachtete Beispiel:

$\overrightarrow{QP} \perp \vec{e}$ und $\overrightarrow{QP} \perp \vec{x}$

$$\begin{pmatrix} 8-10m-q \\ 4-6m-q \\ 3-2m-q \end{pmatrix}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = 0 \quad \text{und} \quad \begin{pmatrix} 8-10m-q \\ 4-6m-q \\ 3-2m-q \end{pmatrix}\begin{pmatrix} 10 \\ 6 \\ 2 \end{pmatrix} = 0$$

15 - 18m - 3q = 0 | ·(-6) und zu 2. Gleichung addieren
110 -140m - 18q = 0
20 - 32m = 0 und damit m = $^{5}/_{8}$ bzw. durch Einsetzen in die 1. Gleichung q = $^{5}/_{4}$.

Gleichung der Ausgleichsgeraden: $y = \frac{5}{8}x + \frac{5}{4}$

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