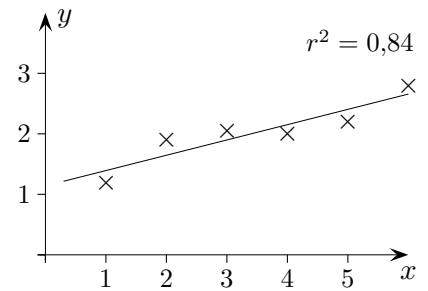


Regressionsgerade

x	x_1	x_2	x_3	\dots	x_n
y	y_1	y_2	y_3	\dots	y_n

Beim Auswerten von Messreihen wird häufig eine durch theoretische Überlegungen nahegelegte lineare Beziehung zwischen den x - und y -Werten gesucht, d. h. eine Gerade $y = mx + b$ (Regressionsgerade), die die Datenpunkte möglichst gut approximiert.



Als Abweichungsmaß kann (nach Gauss) die Summe der Quadrate der Differenzen

$$Q = (mx_1 + b - y_1)^2 + (mx_2 + b - y_2)^2 + \dots + (mx_n + b - y_n)^2$$

genommen werden. Hierbei werden m und b so gewählt, dass Q einen kleinsten Wert annimmt.

Q kann als Funktion der Variablen m und b betrachtet werden.

Im Minimum müssen dann die Ableitungen nach b und m Null ergeben.

$$Q(b) = \sum (mx_i + b - y_i)^2$$

Die Summe erstreckt sich stets von 1 bis n .

$$Q'(b) = \sum 2(mx_i + b - y_i)$$

$$\implies m \sum x_i + nb - \sum y_i = 0$$

$$\implies \bar{y} = m\bar{x} + b$$

Mittelwerte $\bar{x} = \frac{1}{n} \sum x_i$, $\bar{y} = \frac{1}{n} \sum y_i$

$P(\bar{x} | \bar{y})$ liegt auf der Ausgleichsgeraden.

m ist noch zu bestimmen.

Um die Rechnung einfach zu halten, wählen wir den Schwerpunkt als Ursprung:

$$d_i = x_i - \bar{x}$$

$$e_i = y_i - \bar{y}$$

b ist dann Null, die Steigung hat sich nicht verändert.

$$Q(m) = \sum (md_i - e_i)^2$$

$$Q'(m) = \sum 2(md_i - e_i) \cdot d_i$$

$$0 = 2m \sum d_i^2 - 2 \sum e_i d_i$$

$$\implies m_{\min} = \frac{\sum d_i e_i}{\sum d_i^2}$$

$$m = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sum (x_i - \bar{x})^2}$$

Die Gleichung der Regressionsgeraden lautet daher: $y = m(x - \bar{x}) + \bar{y}$

Korrelationskoeffizient

Wir benötigen ein Maß dafür, wie stark die Datenpunkte um die Regressionsgerade streuen. Dazu rechnen wir die quadratische Abweichung aus.

$$\begin{aligned} Q &= \sum (e_i - md_i)^2 & m &= \frac{\sum e_i d_i}{\sum d_i^2} \\ &= \sum (e_i^2 - 2e_i m d_i + (m d_i)^2) \\ &= \sum e_i^2 - 2m \sum d_i e_i + m^2 \sum d_i^2 \\ &= \sum e_i^2 - \frac{2(\sum e_i d_i)^2}{\sum d_i^2} + \frac{(\sum e_i d_i)^2 \sum d_i^2}{(\sum d_i^2)^2} && \text{kürzen durch } \sum d_i^2 \\ &= \sum e_i^2 - \frac{(\sum e_i d_i)^2}{\sum d_i^2} \end{aligned}$$

Je kleiner der Term $\frac{(\sum e_i d_i)^2}{\sum d_i^2}$ ist, desto größer ist die Quadratsumme.

Diese ist Null, falls $\frac{(\sum e_i d_i)^2}{\sum d_i^2} = \sum e_i^2$ ist.

$$\begin{aligned} \text{Da } Q \geq 0 \text{ ist, folgt} \quad & 0 \leq \frac{(\sum e_i d_i)^2}{\sum d_i^2} \leq \sum e_i^2 \\ \implies & 0 \leq \frac{(\sum e_i d_i)^2}{\sum d_i^2 \sum e_i^2} \leq 1 \end{aligned}$$

Der mittlere Term heißt *Bestimmtheitsmaß*.

Gebräuchlicher ist der *Korrelationskoeffizient*

$$r = \frac{\sum e_i d_i}{\sqrt{\sum d_i^2 \sum e_i^2}} = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}}$$

r ist die Wurzel aus dem Bestimmtheitsmaß und hat im Gegensatz zu diesem stets dasselbe Vorzeichen wie die Steigung m (leicht zu sehen). Es ist: $-1 \leq r \leq 1$.

Für $r = 0$ ist die Steigung m auch Null. Es liegt kein linearer Zusammenhang vor, für $r = 1$ und $r = -1$ liegen die Datenpunkte auf der Regressionsgeraden.

Zu beachten ist, dass ein hoher Korrelationskoeffizient nicht eine *kausale* Abhängigkeit bedeuten muss.

In Excel können Ausgleichsgeraden ohne Aufwand ausgegeben werden:

Auf einen Datenpunkt klicken, mit rechter Maustaste Trendlinie hinzufügen, Trendlinie formatieren, Optionen, Gleichung und Bestimmtheitsmaß im Diagramm darstellen.

Regressionsgerade, vektorielle Lösung

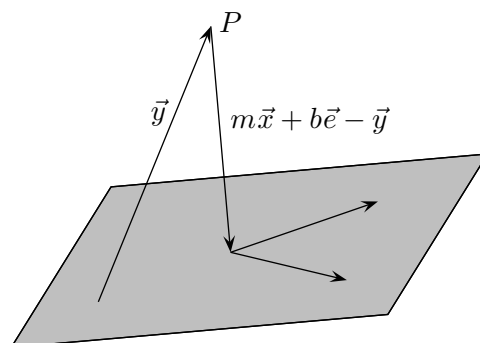
Als mathematisch fruchtbar wird es sich erweisen, das Ausgleichsproblem vektoriell zu lösen. Hierzu betrachten wir die Quadratsumme als Skalarprodukt, dass wir kühn auf n Dimensionen verallgemeinern. Wenn auch die Anschauung aufgegeben werden muss, so bleiben doch die Rechenregeln (Linearität, usw.) erhalten.

$$Q(m, b) = \left(\begin{array}{c} mx_1 + b - y_1 \\ mx_2 + b - y_2 \\ \vdots \\ mx_n + b - y_n \end{array} \right)^2 = \left[m \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + b \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} - \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} \right]^2 = [m\vec{x} + b\vec{e} - \vec{y}]^2$$

$\vec{OX} = m\vec{x} + b\vec{e}$ kann als Ebene durch den Ursprung im n -dimensionalen Raum gedeutet werden und $Q(m, b)$ als Quadrat des Abstandes des zum Vektor \vec{y} gehörenden Punktes P zu einem Punkt der Ebene.

Dieser Abstand wird minimal (zumindest sind wir im Fall $n = 3$ davon überzeugt), falls der Differenzvektor $m\vec{x} + b\vec{e} - \vec{y}$ orthogonal (d.h. das Skalarprodukt ist Null) zu den Richtungsvektoren der Ebene ist.

$$\begin{array}{r} (m\vec{x} + b\vec{e} - \vec{y}) \cdot \vec{x} = 0 \\ (m\vec{x} + b\vec{e} - \vec{y}) \cdot \vec{e} = 0 \\ \hline m\vec{x}^2 + b\vec{e}\vec{x} - \vec{y}\vec{x} = 0 \\ m\vec{x}\vec{e} + b\vec{e}^2 - \vec{y}\vec{e} = 0 \\ \hline m \sum x_i^2 + b \sum x_i - \sum x_i y_i = 0 \\ m \sum x_i + bn - \sum y_i = 0 \\ \hline m \sum x_i^2 + nb\bar{x} - \sum x_i y_i = 0 \\ \hline m\bar{x} + b - \bar{y} = 0 \end{array}$$



Diese letzten beiden Gleichungen ergaben sich auch schon bei der Lösung mit der Differentialrechnung, so dass das weitere Vorgehen identisch ist.

Wir wollen noch der Frage nachgehen, wie innerhalb der Vektorrechnung gezeigt werden kann, dass diese Orthogonalitäts-Schlussweise für allgemeines n zum Minimum der Quadratsumme führt. Die Tragweite dieser Überlegungen wird auf der nächsten Seite sichtbar.

Der Nachweis gelingt, wenn wir uns die betrachtete Ebene von zwei zueinander orthogonalen Einheitsvektoren aufgespannt denken:

$$\vec{OX} = \lambda \vec{a}^\circ + \mu \vec{b}^\circ, \quad \vec{a}^\circ \perp \vec{b}^\circ, \quad |\vec{a}^\circ| = |\vec{b}^\circ| = 1$$

Das noch Fehlende soll im Hinblick auf Späteres etwas allgemeiner formuliert werden.

Weg in die Abstraktion

Sei \vec{a}_i° , $i = 1 \dots m$, ein Orthonormalsystem, d.h. die Vektoren haben die Länge 1 und sind paarweise orthogonal. Wir rechnen das Quadrat des Abstands eines Punkts P zu einer Linearkombination der \vec{a}_i° aus.

$$\begin{aligned} \left(\vec{OP} - \sum_{i=1}^m \lambda_i \vec{a}_i^\circ \right)^2 &= \vec{OP}^2 - 2 \vec{OP} \cdot \sum_{i=1}^m \lambda_i \vec{a}_i^\circ + \left(\sum_{i=1}^m \lambda_i \vec{a}_i^\circ \right)^2 \\ &= \vec{OP}^2 - 2 \sum_{i=1}^m \lambda_i \vec{OP} \cdot \vec{a}_i^\circ + \sum_{i=1}^m \lambda_i^2 \\ &= \vec{OP}^2 - \sum_{i=1}^m \left(\vec{OP} \cdot \vec{a}_i^\circ \right)^2 + \sum_{i=1}^m \left(\lambda_i - \vec{OP} \cdot \vec{a}_i^\circ \right)^2 \end{aligned}$$

Offensichtlich wird das Minimum für die Wahl von $\lambda_i = \vec{OP} \cdot \vec{a}_i^\circ$ angenommen.

Nun kann abschließend noch der Nachweis der Orthogonalität erbracht werden.

Der Vektor, der den Fußpunkt mit dem Punkt P verbindet,

$$\vec{OP} - \sum_{i=1}^m \left(\vec{OP} \cdot \vec{a}_i^\circ \right) \cdot \vec{a}_i^\circ$$

ist orthogonal zu jedem \vec{a}_k° und damit auch zu jeder Linearkombination der \vec{a}_k° , im Fall $n = 3$ also zu allen Vektoren der Ebene:

$$\left(\vec{OP} - \sum_{i=1}^m \left(\vec{OP} \cdot \vec{a}_i^\circ \right) \cdot \vec{a}_i^\circ \right) \cdot \vec{a}_k^\circ = \vec{OP} \cdot \vec{a}_k^\circ - \vec{OP} \cdot \vec{a}_k^\circ = 0$$

Damit wurde unser Vorgehen nachträglich gerechtfertigt.

Bei der vektoriellen Bearbeitung des Ausgleichsproblems haben wir uns letztendlich von der Anschauung getrennt und uns nur noch auf die rechnerischen Umformungen verlassen.

Zum Verständnis weiterer Methoden zur Lösung von Approximationsproblemen, z.B. der Approximation einer Funktion durch eine Linearkombination trigonometrischer Funktionen (Fourieranalyse) ist ein zusätzlicher Abstraktionsschritt erforderlich.

Blicken wir noch einmal zurück und versuchen, das Wesentliche herauszustellen. Wir bewegten uns in einer Menge von Elementen (Vektoren), auf denen Rechenoperationen und ein Skalarprodukt, d.h. eine Zuordnung mit gewissen Eigenschaften, definiert sind. Das Skalarprodukt legt für die Elemente des Vektorraums eine Länge (Norm), einen Abstand und eine Orthogonalitätsbeziehung fest. In diesem mathematischen Kontext kann zu einem Element die ihm nächstliegende Linearkombination eines Orthonormalsystems gefunden werden.

Statt der Vektoren betrachten wir nun Funktionen auf einem Intervall $[a, b]$. Funktionen können addiert, mit einer Zahl multipliziert und auf ihnen kann ein Skalarprodukt definiert werden, und zwar durch:

$$\langle f | g \rangle = \int_a^b f \cdot g \, dx, \quad \langle f | g \rangle \text{ ist eine von mehreren gebräuchlichen Schreibweisen.}$$

Weg in die Abstraktion

Das Skalarprodukt hat eine Norm $\|f\| = \sqrt{\langle f | g \rangle}$ und eine Abstandsdefinition $d(f, g) = \|f - g\|$ zur Folge.

Übertrage

- a) $\|f\| = 1$
- b) $f \perp g$
- c) $\|f - g\| \leq \frac{1}{100}$

in Aussagen über Flächeninhalte.

Das für Approximationen am häufigsten verwendete Orthonormalsystem besteht aus bestimmten trigonometrischen Funktionen; es sind jedoch auch andere Systeme bestimmter Polynome bekannt.

