Ausgleichsgeraden – Lineare Regression

Durch zwei (Daten-) Punkte kann man eindeutig eine Gerade legen. Kommt ein dritter Punkt dazu, so kann man oft keine eindeutige Gerade zeichnen, da z. B. der dritte Punkt nicht auf der Geraden durch die beiden ersten Punkte liegt. Hat man noch mehr Punkte, so wird es noch schwieriger, durch die zugehörige Punktwolke eine „gut ausgleichende" Gerade zu legen. Oft will man aber eine derartige Ausgleichgerade bestimmen, da man weiß, dass der Zusammenhang zwischen den beiden Koordinaten linear ist. Dies tritt z. B. oft in der Physik (HOOKsches Gesetz, ...) auf
Der Rechner bietet dafür ein Menü, das man zur Zeichnung optimaler Geraden benutzen kann. Diese Geraden nennt man Regressionsgeraden, das Verfahren heißt lineare Regression.

Problemfelder:

1) Wie kann man durch die Punkte A(0|0), B(1|1), C(2|3) eine Ausgleichsgerade zeichnen? Zeichne händisch im Koordinatensystem und lasse auf dem GTR zeichnen.
2) Wie kann man die Gerade optimal einzeichnen, wenn man verlangt, dass der Abstand von Gerade zu Punkt möglichst gering ist? Wie verhält sich die Summe dieser Abstände?
3) Um kleinere Abstände geringer zu gewichten als größere, wählt man häufig das Quadrat der Abstände als „gewichtetes" Maß. Ist der Abstand kleiner als 1, wird er durch das Quadrieren kleiner, sonst größer. Darüber hinaus unterdrückt man das Vorzeichen. Wie wirkt sich das auf die Berechnung aus?
4) Die Ausgleichsgeraden haben zwei Parameter, die man in der Tabellenkalkulation variieren kann. So kann man sich durch gezieltes Experimentieren einer Kombination dieser Parameterwerte nähern, die eine möglichst kleine „Fehlerquadratsumme" erzeugen.
5) In der Sek II kann man die Regressionsgerade für 3 Punkte noch einfach bestimmen. Vergleiche die Werte mit der Regressionsgerade, die der Rechner bestimmt.

Experimentiere mit Hilfe der Tabellenkalkulation. Man kann die Fehlerquadrate auch mittels einer Dynamischer-Geometrie-Software veranschaulichen. Dann sieht man besonders gut, wie sich die Flächen bei Veränderung der Geraden verhalten.

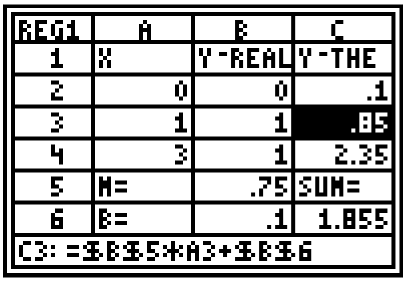
Analyse:

Man kann die drei Punkte in der Tabellenkalkulation eingeben und sich die Punkte plotten lassen. Dann legt man mittels des y=-Menüs entsprechende Geraden durch diese Punktwolke. Noch sinnvoller ist es, sich diese Punkte in ein Koordinatensystem zu zeichnen und eine Ausgleichsgerade händisch nach Augenmaß zu zeichnen. Hierbei kann sich bereits eine Diskussion über die Güte der Ausgleichsgerade ergeben.

Nach Einführung von Abstandsmaßen sind die Untersuchungen mittels CellSheet™ besonders einfach zu realisieren. Hier kann mit (einfachen) Differenzen, den Absolutbeträgen oder den quadratischen Abständen experimentiert werden.

Will man die Ausgleichsgerade analytisch bestimmen, muss man eine Funktion von zwei Variablen definieren, die die Abstandssumme von Punkt und Gerade beschreibt. Von dieser Funktion ist das Minimum gesucht, das sich mittels partieller Ableitungen schnell finden lässt. Die ermittelten Werte werden mit denen des Regressionstools für lineare Regression übereinstimmen. In der Sek I wird man sich nur diesen Werten experimentell nähern können, wobei aber das Grundprinzip trotzdem deutlich wird. So kann man die Black-Box zur Regression etwas erhellen.

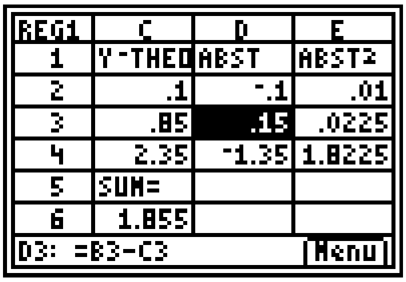
Rechenblatt in CellSheet™ (TI-83 Plus):



REG1	A	B	C
1	X	Y-REAL	Y-THE
2	0	0	.1
3	1	1	.85
4	3	1	2.35
5	M=	.75	SUM=
6	B=	.1	1.855

C3: =$B$5*A3+$B$6

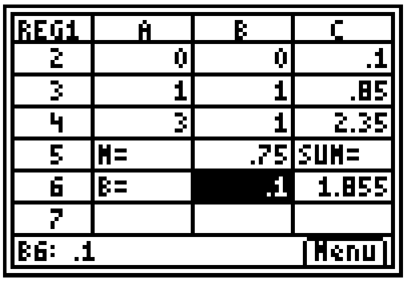
Bild1



REG1	C	D	E
1	Y-THEO	ABST	ABST²
2	.1	⁻.1	.01
3	.85	.15	.0225
4	2.35	⁻1.35	1.8225
5	SUM=		
6	1.855		

D3: =B3-C3 Menu

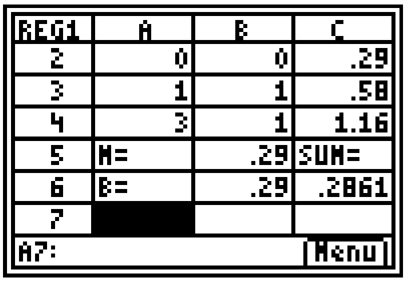
Bild 2



REG1	A	B	C
2	0	0	.1
3	1	1	.85
4	3	1	2.35
5	M=	.75	SUM=
6	B=	.1	1.855
7			

B6: .1 Menu

Bild 3



REG1	A	B	C
2	0	0	.29
3	1	1	.58
4	3	1	1.16
5	M=	.29	SUM=
6	B=	.29	.2861
7			

A7: Menu

Bild 4

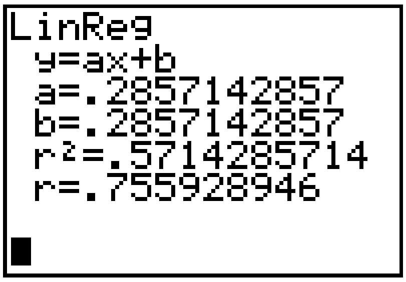


Bild 5

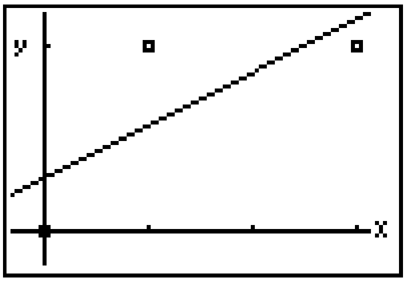


Bild 6

Hinweise

Man kann die Phänomene der Regression bereits an kleinen Datenmengen untersuchen, wird aber auch größere Datenmengen z. B. aus der Physik untersuchen.

Das Problem lässt sich gut in Partnerarbeit bearbeiten. Die Parametervariation ist dabei der zentrale Aspekt.

Die Darstellung der Ausgleichsgeraden auf dem DGS-Modul des TI-83 Plus (Zusatz) ist wenig sinnvoll, auf dem TI-Voyage 200 ist es gerade noch visuell akzeptabel.

Anschlussprobleme sind in den Naturwissenschaften reichlich vorhanden.

Die Realisation auf dem TI-Voyage 200 sieht analog aus, mit der DGS ergibt sich:

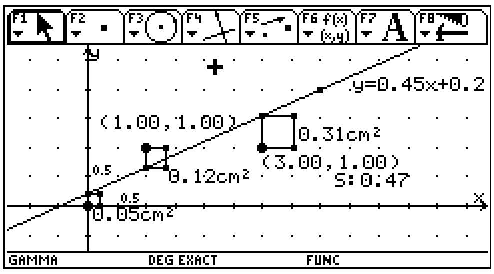


Bild 7