

□ Lineare Regression

Dokumentnummer: D1192
 Fachgebiet: Regression, Trendgerade,
 Rechnen mit Summen, Gleichungen,
 lineare Funktion
 Einsatz: 2HAK (erstes Lernjahr)

Figure 1:

PROGRAMMQUELLEN UND PROGRAMMBEILAGEN	
http://maxima.sourceforge.net	http://www.geogebra.org
	

□ 1 Erklärungen zum Verständnis

□ 1.1 Das Summenzeichen

Figure 2: Was bedeutet das Summenzeichen (ein großes griechisches Sigma)? Es ist übrigens auch aus vielen Tabellenkalkulationsprogrammen bekannt.

$$\sum_{i=1}^3 i$$

```
(%i1) sum(i,i,1,3)
      /* wir addieren die Werte
         i=1 + i=2 + i=3 */;
(%o1) 6
```

```
(%i2) ev(i,i=1)+ev(i,i=2)+ev(i,i=3);
(%o2) 6
```

□ 1.2 Ein Beispiel

Figure 3: Ein weiteres Beispiel dazu:

$$\sum_{k=2}^7 k^2$$

□ 1.3 Eine Übung

```
(%i3) sum(k**2,k,2,7)
      /* wir addieren die Werte von k2 für
      k=2 -> 22 = 4,
      k=3 -> 32 = 9,
      k=4 -> 42 = 16,
      k=5 -> 52 = 25,
      k=6 -> 62 = 36 und
      k=7 -> 72 = 49 */;
```

(%o3) 139

Figure 4: Noch drei Übungsmöglichkeiten:

$$\sum_{n=4}^{12} n + 1, \sum_{i=1}^5 i, \sum_{k=2}^4 \frac{1}{k}$$

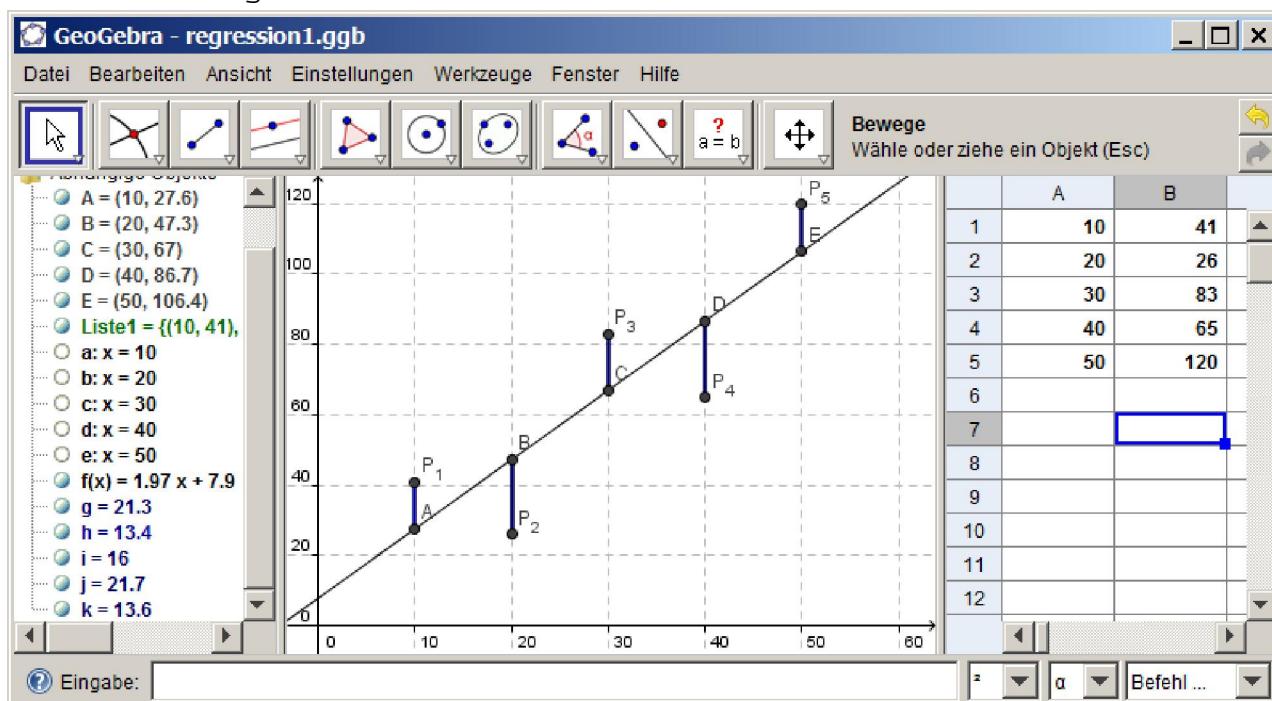
2 Problembeschreibung

2.1 Aufgabe

Gesucht ist ein Maxima-Programm zur Berechnung der linearen Regression. Wir geben eine Erklärung dazu:

2.2 Erklärung

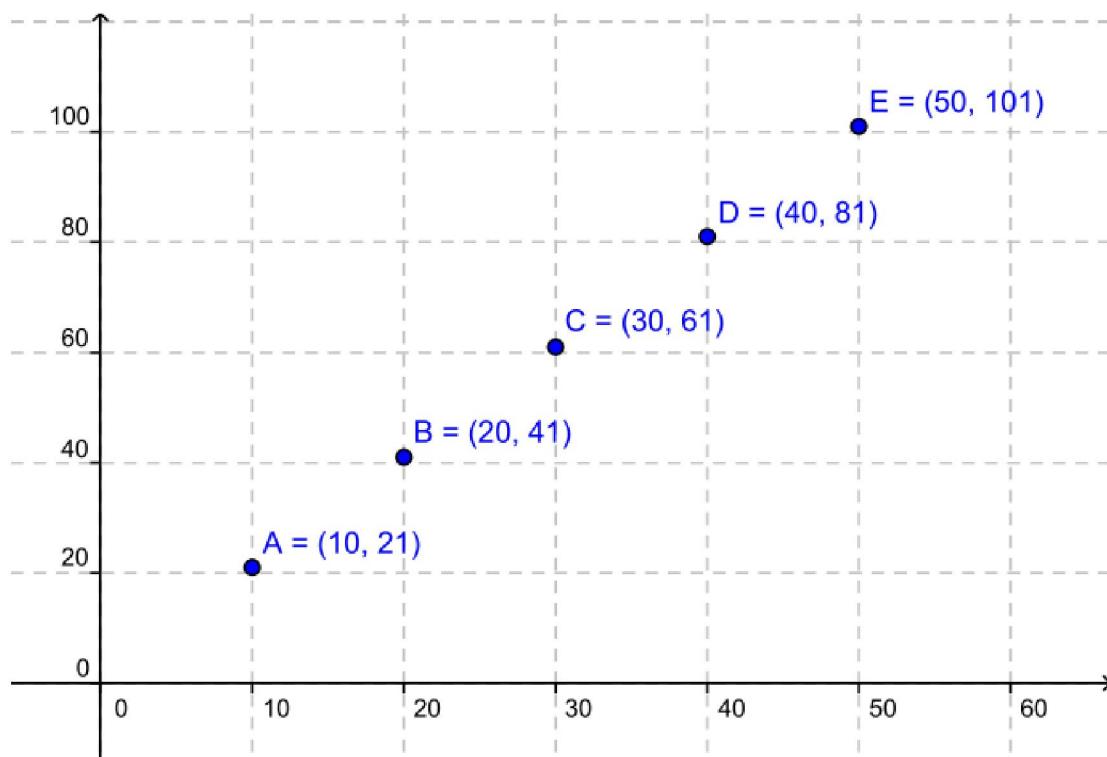
Figure 5: Die Regressionsgerade für die Punkte P₁, P₂, P₃, P₄ und P₅ ist jene Gerade, für die die Summe der Quadrate der Abstände P₁A, P₂B, P₃C, P₄D und P₅E ein Minimum wird. Das lässt sich mit Hilfe der Differentialrechnung verarbeiten. Die Lösung der Extremwertaufgabe wird durch die Regressionsgleichungen ausgedrückt.



3 Problemlösung

□ 3.1 Eingabe

Figure 6: Daten für das Musterbeispiel



```
(%i4) x:[10,20,30,40,50]
      /* Liste mit den x-Werten */;
(%o4) [ 10 , 20 , 30 , 40 , 50 ]
```

```
(%i5) y:[21,41,61,81,101]
      /* Liste mit den y-Werten */;
(%o5) [ 21 , 41 , 61 , 81 , 101 ]
```

□ 3.2 Verarbeitung

```
(%i6) n:length(x)
      /* die Anzahl der Messwerte bestimmen */;
(%o6) 5
```

Figure 7: Die Gleichungen für die lineare Regression,
wir benötigen 4 Summen, die wir als
sx2, sx, sxy und sy bezeichnen wollen:

$$a \cdot \sum_{i=1}^n x_i^2 + b \cdot \sum_{i=1}^n x_i = \sum_{i=1}^n x_i \cdot y_i$$

$$a \cdot \sum_{i=1}^n x_i + b \cdot n = \sum_{i=1}^n y_i$$

Figure 8: Berechnung von sx2

$$\sum_{i=1}^n x_i^2$$

```
(%i7) sx2:=sum(x[i]**2,i,1,n);
(%o7) 5500
```

Figure 9: Berechnung von sx

$$\sum_{i=1}^n x_i$$

```
(%i8) sx:=sum(x[i],i,1,n);
(%o8) 150
```

Figure 10: Berechnung von sxy

$$\sum_{i=1}^n x_i \cdot y_i$$

```
(%i9) sxy:=sum(x[i]*y[i],i,1,n);
(%o9) 11150
```

Figure 11: Berechnung von sy

$$\sum_{i=1}^n y_i$$

```
(%i10) sy:=sum(y[i],i,1,n);
(%o10) 305
```

Figure 12: Die Summen kann man auch mit Tabellenkalkulation berechnen

	A	B	C	D
1	x	y	x2	xy
2		10	21	100
3		20	41	400
4		30	61	900
5		40	81	1600
6		50	101	2500
7		150	305	5500
8	sx	sx	sx2	sxy

Nun können wir die Gleichungen zusammenstellen

```
(%i11) g1:=a*sx2+b*sx=sxy;
(%o11) 150 b + 5500 a = 11150
```

```
(%i12) g2:=a*sx+b*n=sy;
(%o12) 5 b + 150 a = 305
```

Die Lösung des Gleichungssystems erfolgt mit solve() oder algsys()

```

(%i13) l:solve([g1,g2],[a,b]);
(%o13) [[a=2 , b=1 ]]

(%i14) l:algsys([g1,g2],[a,b]);
(%o14) [[a=2 , b=1 ]]

(%i15) A:a,l[1][1];
(%o15) 2

(%i16) B:b,l[1][2];
(%o16) 1

(%i17) kill(x,y)
        /* die gespeicherten Eingabedaten löschen
           und so die Wiederverwendung von x und y
           ermöglichen */;
(%o17) done

(%i18) gerade:y=A*x+B;
(%o18) y=2 x+1

```

3.3 Ausgabe

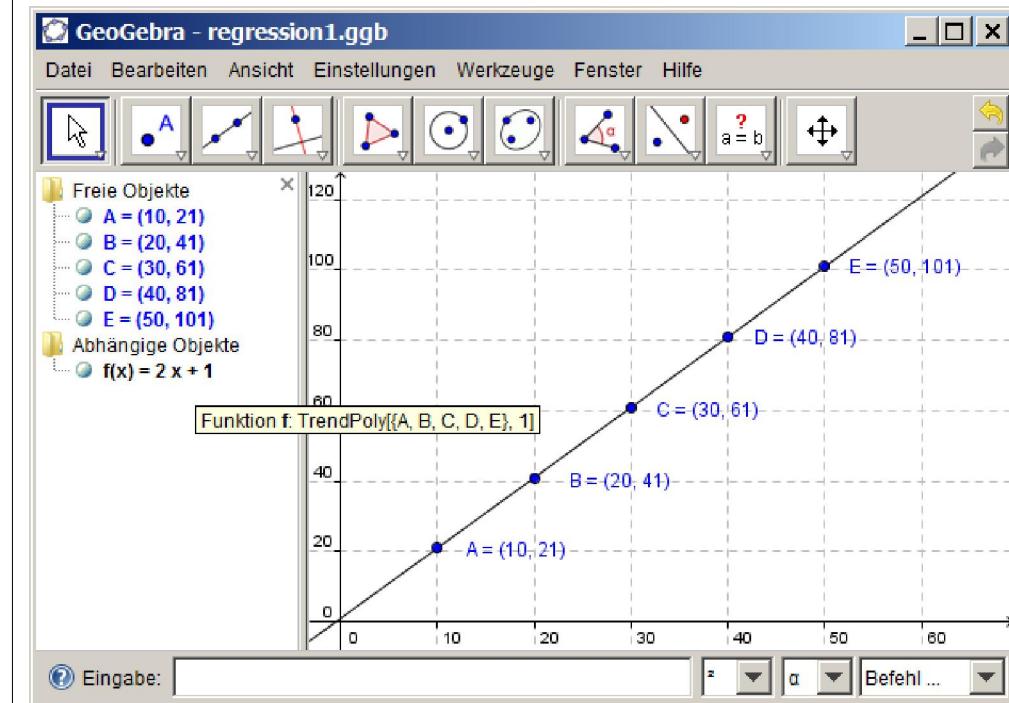
```

(%i19) print("")$ 
        print("Die Regressionsgerade: " ,gerade)$

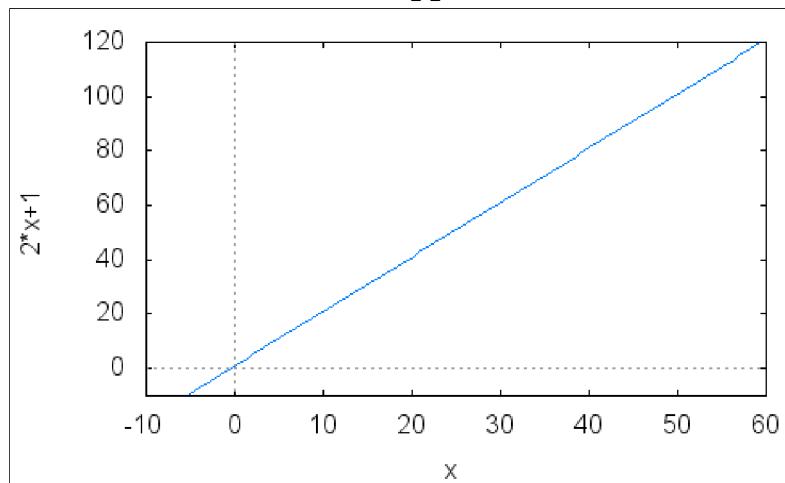
Die Regressionsgerade: y=2 x+1

```

Figure 13: Die Lösung der Musteraufgabe mit Geogebra



```
(%i21) wxplot2d([rhs(gerade)], [x,-10,60], [y,-10,120])$  
plot2d: some values were clipped.
```



□ 4 Übungsaufgabe

Figure 14: Man kontrolliere die Lösung mit wxMaxima

