

Das Gebiet der Optimierung (Extremwertprobleme) beschäftigt sich damit, optimale Parameter für eine Zielfunktion zu finden, wobei diese entweder ein Minimum oder ein Maximum erreichen soll.

Optimierungsprobleme stellen sich generell in allen wissenschaftlichen Disziplinen, in denen mit unbekannten Parametern gearbeitet wird, wie beispielsweise in den Naturwissenschaften oder den Wirtschaftswissenschaften. In vielen Fällen ist eine Lösung von Optimierungsproblemen mit Mitteln der Differenzialrechnung möglich.

Mit Optimierungsproblemen haben sich im Laufe der Jahrhunderte viele bedeutende Mathematiker wie zB Gottfried Wilhelm Leibniz (1646 – 1716), Pierre de Fermat (1601 – 1665), Joseph-Louis de Lagrange (1736 – 1813) und Carl Friedrich Gauß (1777 – 1855) beschäftigt. Bekannt geworden ist zB der Fermat'sche Satz des kürzesten Wegs von Licht.



Pierre de Fermat (Frankreich)

2.1 Optimierungsaufgaben

Viele Problemstellungen haben das Maximieren oder Minimieren von Größen zum Ziel. Dabei kann es sich zum Beispiel um das Aufsuchen eines maximalen Gewinns oder eines möglichst geringen Materialverbrauchs handeln. Die mathematische Formulierung eines solchen Problems nennt man **Extremwertaufgabe**. Ziel dabei ist das Ermitteln, welchen Wert die unabhängige Variable annehmen muss, damit die davon abhängige Größe maximal oder minimal wird.

2.1.1 Extremwertaufgaben ohne Nebenbedingung

ABD

2.1 Ein Ball wird mit einer Anfangsgeschwindigkeit $v_0 = 20 \text{ m/s}$ zum Zeitpunkt $t = 0 \text{ s}$ senkrecht nach oben geworfen. Die Höhe h (in Meter) zum Zeitpunkt t (in Sekunden) ist gegeben durch:

$$h(t) = v_0 \cdot t - \frac{g \cdot t^2}{2} \quad \text{mit } g \approx 10 \text{ m/s}^2$$

Stelle die Funktion grafisch dar.

Lies jenen Zeitpunkt ab, zu dem der Ball die maximale Höhe erreicht.

Ermittle die Steigung der Tangente zu diesem Zeitpunkt.

Beschreibe das Monotonieverhalten der Funktion vor und nach dieser Stelle.

Du kannst aus der Grafik ablesen, dass die maximal erreichbare Höhe 20 m beträgt und zum Zeitpunkt 2 Sekunden nach dem Abwurf erreicht wird.

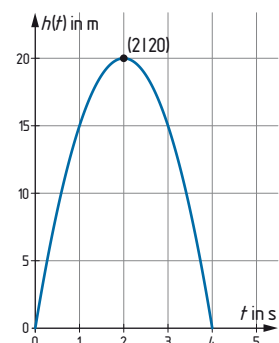
Die Steigung der Tangente $h'(t) = 20 - 10t \Rightarrow k_t(2) = 0$

An dieser Stelle hat die Kurve eine waagrechte Tangente. Das ist die **notwendige Bedingung** für ein **lokales Maximum**.

Die 2. Ableitung an der Stelle 2 ergibt eine negative Zahl (-10). Dies ist die **hinreichende Bedingung** für ein lokales Maximum.

$f'(t) > 0$ im Bereich $[0; 2[$... streng monoton steigend;

$f'(t) < 0$ im Bereich $]2; 4]$... streng monoton fallend



Auf Optimierungsprobleme können die Kriterien, die für die Kurvenuntersuchung gelten, angewendet werden (siehe Abschnitt „1 Differenzieren“).

Lokale Extremwerte findest du über die Bedingung, dass die erste Ableitung an dieser Stelle gleich null ist.

Für das **lokale Maximum** muss zusätzlich die 2. Ableitung an dieser Stelle kleiner als null sein, für das **lokale Minimum** ist die 2. Ableitung an dieser Stelle größer null.

Ist der Definitionsbereich ein abgeschlossenes Intervall $[a; b]$, so kann ein **globaler Extrempunkt** auch am Rand des Intervalls liegen.

Man spricht in diesem Fall von einem **Randmaximum** bzw. bei einem kleinsten Wert von einem **Randminimum**. An dieser Stelle liegt im Allgemeinen keine waagrechte Tangente vor.

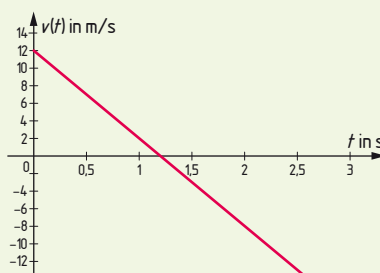
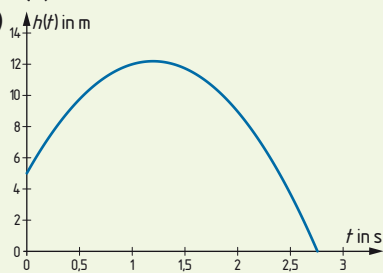
2.2 Für die momentane Höhe h eines senkrecht nach oben geworfenen Balls gilt:
 $h(t) = 5 + 12t - 5t^2$; t ... Zeit in s nach dem Abwurf; h ... Höhe in m nach t s

- Erstelle die Gleichung für die Funktion v .
- Stelle h und v grafisch dar.
- Beschreibe den Zusammenhang zwischen den beiden Graphen.
- Ermittle, nach wie viel Sekunden der Ball die maximale Höhe erreicht hat.

Lösung:

a) $v(t) = 12 - 10t$

b)



- Die Geschwindigkeit wird während der Aufwärtsbewegung des Balls immer kleiner bis sie schließlich null ist. Die Höhenzunahme wird immer kleiner bis der Ball seine größte Höhe erreicht hat. Die Stelle mit der maximalen Höhe entspricht der Nullstelle der Geschwindigkeit. Danach kehrt sich der Vorgang um. Die Geschwindigkeit wird laufend größer und hat wegen der Abwärtsbewegung des Balls ein negatives Vorzeichen. Die Höhe nimmt zunehmend stärker ab, bis der Ball auf dem Boden auftrifft.
- Nach 1,2 s erreicht der Ball die maximale Höhe.

2.3 Die Geschwindigkeit eines Heißluftballons während einer 90-minütigen Fahrt wird durch die Funktion v beschrieben:

$$v(t) = -\frac{7t^3}{5400} + \frac{7t^2}{60}$$

t ... Zeit in Minuten;

$v(t)$... Geschwindigkeit in m/min nach t min

- Stelle v in Abhängigkeit von der Zeit t grafisch dar.
- Ermittle, zu welchem Zeitpunkt der Ballon die maximale Geschwindigkeit erreicht.
- Berechne, wie hoch die Beschleunigung zu dem Zeitpunkt ist, zu dem der Ballon die maximale Geschwindigkeit erreicht hat.



ABC



AB



Optimierung und Regression

BCD



2.4

Tony möchte von einem gleichschenkligen Dreieck mit dem Umfang von 25 cm wissen, welche Seite es haben müsste, damit sein Flächeninhalt möglichst groß ist.

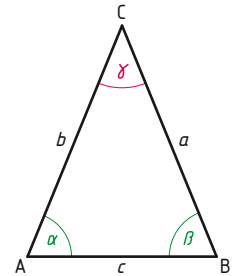
Er hat nach einigen geometrischen Umformungen herausgefunden, dass der Inhalt A der Fläche sich in diesem Fall durch die Länge der Grundlinie c ausdrücken lässt:

$$A(c) = 0,25c \cdot \sqrt{625 - 50c}$$

c ... Grundlinie des Dreiecks in cm;

$A(c)$... Flächeninhalt des Dreiecks mit Grundlinie c in cm^2

- Zeichne die Funktion A in Abhängigkeit von c .
- Berechne, bei welchem Wert von c das Maximum von A vorliegt.
- Beschreibe, wie man die weiteren Seiten des Dreiecks erhält.
- Zeige, wie Tony auf die obige Formel für die Fläche gekommen ist.



BCD



2.5

In einem Schulprojekt sollen Volksschulkinder das Rechteck mit dem größten Umfang herausfinden, das in einem Kreis mit dem Radius 10 cm Platz hat, so dass die Ecken des Rechtecks auf der Kreislinie anstoßen.

Die Kinder schneiden passende Rechtecke aus und legen sie auf den vorgegebenen Kreis. Mit einem Messband messen sie den Umfang.

Es gelingt keinem der Kinder, das wirklich umfanggrößte Rechteck zu finden.

Du kennst aber eine Formel, die den Zusammenhang des Umfangs mit einer der beiden Seiten des Rechtecks darstellt:

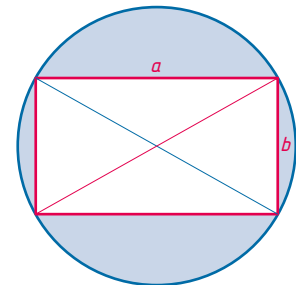
$$u(b) = 2 \cdot \left(b + \sqrt{400 - b^2} \right)$$

b ... Breite des Rechtecks in cm; $u(b)$... Umfang in cm bei einer Breite von b cm

- Berechne die Seite b des umfanggrößten Rechtecks.
- Interpretiere das Ergebnis hinsichtlich der Seite a .

Teile der Kindergruppe mit, welches Rechteck sie ausschneiden müssen, damit die Aufgabe gelöst ist.

- Erkläre, wie die angegebene Formel mit dem Radius r des Kreises zusammenhängt.



BCD



2.6

Ein zylinderförmiger Blumentopf soll 120 Liter fassen. Die Abmessungen sollen so gewählt werden, dass der Materialverbrauch so gering wie möglich ist.

Der Materialverbrauch kann mit der folgenden Formel dargestellt werden.

$$O(r) = r^2 \cdot \pi + \frac{240}{r}$$

r ... Radius in cm;

O ... Oberfläche des oben offenen Zylinders in cm^2

- Ermittle die Abmessungen des Blumentopfs mit minimalem Materialverbrauch.
- Überprüfe stichprobenartig, ob der Materialverbrauch größer wird, wenn die Abmessungen verändert werden.
- Zeige, wie man auf die angegebene Formel kommt.

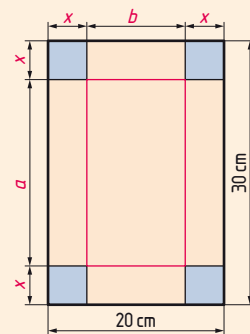


2.1.2 Optimierung mit Nebenbedingungen

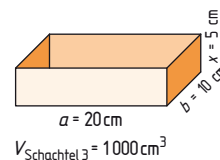
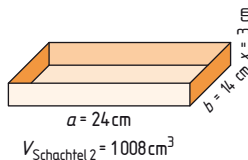
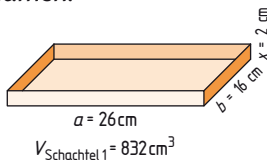
Bisher wurden Extremwerte von Funktionen in **einer Variablen** berechnet. Bei den meisten Extremwertaufgaben handelt es sich jedoch um Aufgaben, in denen eine Größe, die von mehreren Variablen abhängt, wie zum Beispiel ein Volumen oder eine Oberfläche, maximal oder minimal werden soll. Um die gesuchte Größe als Funktion in einer Variablen zu erhalten, müssen zusätzliche Bedingungen angegeben werden, die **Nebenbedingungen** genannt werden.

2.7 Aus einem rechteckigen Stück Karton mit der Größe 20 cm mal 30 cm soll eine oben offene Schachtel hergestellt werden. Dazu schneidet man an den Ecken Quadrate aus und faltet die Seitenwände nach oben. Die Größe des Volumens hängt von der Seitenlänge x der herausgeschnittenen Quadrate ab.

Es soll jener Wert für x ermittelt werden, mit dem die Schachtel das größtmögliche Volumen erreicht.



Wählt man zum Beispiel für die Seitenlängen der weggeschnittenen Quadrate $x = 2$ cm, $x = 3$ cm und $x = 5$ cm, so erhält man drei unterschiedliche Schachteln mit unterschiedlichen Volumen:



$$V_{\text{Quader}} = \text{Länge} \cdot \text{Breite} \cdot \text{Höhe} = a \cdot b \cdot x$$

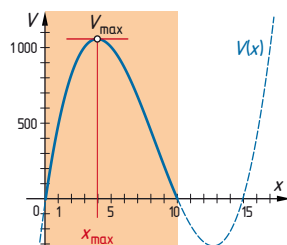
$$\text{Länge: } a = 30 - 2x$$

$$\text{Breite: } b = 20 - 2x$$

$$\text{Höhe: } x \text{ mit } 0 < x < 10$$

$$V(x) = (30 - 2x) \cdot (20 - 2x) \cdot x$$

$$V(x) = 4x^3 - 100x^2 + 600x$$



Aus dem Funktionsverlauf erkennt man: Die Funktion hat im Definitionsbereich $]0; 10[$ an der Stelle $x_{\max}$ ein Maximum. Das bedeutet, dass das Volumen der Schachtel am größten ist, wenn Quadrate der Seitenlänge $x_{\max}$ ausgeschnitten werden. Um das Maximum rechnerisch zu ermitteln, wird die erste Ableitung gebildet und null gesetzt.

$$\frac{dV}{dx} = 12x^2 - 200x + 600$$

$$12x^2 - 200x + 600 = 0 \Rightarrow x_1 \approx 3,9 \text{ cm } (x_2 \approx 12,7 \text{ cm} > 10 \text{ cm})$$

Man erhält das maximale Volumen $V \approx 1\,056,31 \text{ cm}^3$, wenn die ausgeschnittenen Quadrate eine Seitenlänge von rund 3,9 cm haben.

- Volumenformel für den Quader
- Die Zusammenhänge (Nebenbedingungen) werden aus der Skizze abgelesen.
- Da der Karton 20 cm breit ist, muss die Länge x kleiner als 10 cm sein.
- Nun kann das Volumen als Funktion von x angegeben werden.
- Vereinfachen der Funktion

Extremwertaufgaben mit Nebenbedingungen sind Optimierungsaufgaben, bei denen der größtmögliche bzw. kleinstmögliche Wert einer Größe ermittelt wird, die von 2 oder mehreren Variablen abhängt.

Die Funktion, die ein Maximum oder Minimum annehmen soll, wird **Hauptbedingung (= Zielfunktion)** genannt. Zusammenhänge, die das Aufstellen der Zielfunktion in einer Variablen ermöglichen, werden **Nebenbedingungen** genannt.

AB



2.8 Die Zahl 100 soll so in zwei positive Summanden zerlegt werden, dass deren Produkt möglichst groß wird.

Lösung:

x, y ... Summanden

Zielfunktion (ZF):

$$P(x, y) = x \cdot y \rightarrow \text{Maximum}$$

Nebenbedingung (NB):

$$x + y = 100$$

$$\Rightarrow y = 100 - x$$

$$P(x) = x \cdot (100 - x)$$

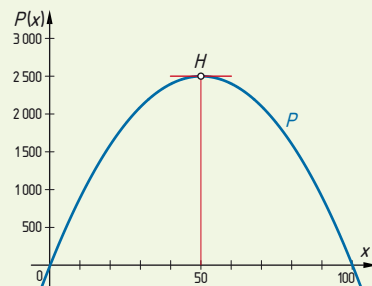
Definitionsbereich: $0 \leq x \leq 100$

$$P(x) = 100x - x^2$$

$$\frac{dP}{dx} = 100 - 2x$$

$$100 - 2x = 0$$

$$\Rightarrow x = 50; y = 100 - x = 50$$



Das Produkt wird am größten, wenn beide Zahlen 50 sind.

- Durch Angeben einer Formel für die Größe, die maximal oder minimal werden soll, erhält man die **Zielfunktion** (ZF), auch **Hauptbedingung** (HB) genannt.
- Die Zielfunktion ist hier eine Funktion in zwei Variablen.
Nun muss mithilfe sachlicher bzw. mathematischer Zusammenhänge eine Gleichung formuliert werden, die es ermöglicht, die **Zielfunktion als Funktion in nur einer Variablen** anzugeben. Diese Gleichung nennt man **Nebenbedingung** (NB).
- Durch Einsetzen der Nebenbedingung in die Hauptbedingung erhält man die Zielfunktion in Abhängigkeit von einer Variablen.
- Der Definitionsbereich muss festgelegt werden.
- Die Zielfunktion wird in möglichst einfacher Form angegeben.
- Im Maximum der Funktion hat deren erste Ableitung den Wert null. Man ermittelt also jene Werte der Funktion P , für die gilt: $\frac{dP}{dx} = 0$
- Stellt man die Zielfunktion grafisch dar, so sieht man, dass das Maximum der Funktion P im angegebenen Definitionsbereich bei 50 liegt.

AB

2.9 Die Zahl 40 soll in zwei natürliche Summanden aufgeteilt werden, sodass die Summe ihrer Quadrate minimal wird.

Erstelle Haupt- und Nebenbedingung. Berechne die Summanden.

2.10 Die reelle Zahl z soll in zwei Faktoren zerlegt werden, sodass deren Summe ein Minimum wird. Erstelle Haupt- und Nebenbedingung. Berechne die Faktoren.

AB

2.11 0,75 Liter eines Getränks sollen in zylinderförmigen geschlossenen Dosen angeboten werden. Über der Flüssigkeit ist ein leeres Volumen von 50 cm^3 .
Ermittle, bei welchen Abmessungen (Radius und Höhe) der Materialverbrauch für eine Dose am geringsten ist.
Erkläre, welcher besondere Zylinder den geringsten Materialverbrauch hat.

ABD



2.12 Aus einem 40 cm langen Drahtstück soll ein Rechteck gebogen werden.
a) Bestimme jeweils die Länge der Seite a und den Flächeninhalt A der entstehenden Rechtecke, wenn für die Seitenlängen $b_1 = 2 \text{ cm}$, $b_2 = 4 \text{ cm}$ bzw. $b_3 = 6 \text{ cm}$ gewählt werden.
b) Erstelle eine Formel, die den Flächeninhalt A als Funktion der Seitenlänge b angibt. Wähle einen sinnvollen Definitionsbereich für die Variable b und stelle die Funktion A in Abhängigkeit von b grafisch dar.
Ermittle, für welchen Wert von b die Funktion A maximal ist.

AB



2.13 Thomas möchte mit einem 6 m langen Weidenzaun ein rechteckiges Blumenbeet mit einem möglichst großen Flächeninhalt einfassen.
Argumentiere, welcher der folgenden Ansätze diesen Sachverhalt beschreibt, und begründe, warum die anderen nicht zutreffen. Berechne x und y .

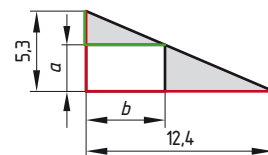
CD



A ☐	ZF: $u(x, y) = 2x + 2y$ soll ein Maximum werden; NB: $6 = 2x + 2y$
B ☐	ZF: $A(x, y) = x \cdot y$ soll ein Maximum werden; NB: $6 = 2x + 2y$
C ☐	ZF: $A(x, y) = x \cdot y$ soll ein Maximum werden; NB: $u(x, y) = 2x + 2y$
D ☐	ZF: $u(x, y) = 2x + 2y$ soll ein Maximum werden; NB: $u = 6$



2.14 Aus der in der Abbildung dargestellten abgeschrägten Hausmauer soll ein rechteckiges Fenster mit den Maßen a und b (in Meter) ausgebrochen werden.
Berechne welche maximale Fensterfläche erzielt werden kann.



AB



2.15 Ein Betrieb verpackt sein Produkt in quaderförmigen Karton-Paketen mit den Seiten a , b und c . Die Volumen betragen jeweils 112 dm^3 . Die Pakete müssen doppelt so lang wie hoch sein.
Der Kartonbedarf zur Herstellung dieser Pakete soll möglichst gering sein.
Erstelle die Haupt- und die Nebenbedingung für diese Pakete.

AB



Ermittle die Abmessungen eines Pakets bei minimalem Kartonbedarf.
(Es werden keine produktionsbedingte Überlappungen berechnet.)

2.2 Regression

Bei Verwendung von Formeln aus der Geometrie oder der Natur- und Wirtschaftswissenschaft hast du es häufig mit dem Zusammenhang zwischen zwei Größen an demselben Untersuchungsobjekt zu tun gehabt. Beispielsweise gibt die Formel $A = r^2 \cdot \pi$ einen exakten Zusammenhang zwischen dem Radius und der Fläche eines Kreises wieder.

Wenn du aber zB eine Versuchsreihe über den von dir als möglich angesehenen Zusammenhang zwischen den Messgrößen Körpermasse und Körpergröße der Kolleginnen und Kollegen in deiner Klasse machst, so sind diese statistisch verteilt. Der Zusammenhang kann aber meist in Form einer Gleichung bzw. einer Funktion näherungsweise beschrieben werden. Dadurch können auch Prognosen über zukünftige bzw. noch nicht gemessene Entwicklungen gemacht werden.

2.2.1 Der lineare Zusammenhang zweier Größen

ABD

2.16 Die Messung der Körpergröße in Zentimeter (cm) und der Körpermasse in Kilogramm (kg) in einer ausgewählten Stichprobe aus deiner Klasse ergab folgende Wertetabelle:

Größe (cm)	167	168	169	170	175	176	177	180	181	182
Masse (kg)	68	70	69	71	78	75	80	78	80	79

Beurteile mit Hilfe einer Grafik, ob ein Zusammenhang zwischen den Messgrößen besteht. Drücke diesen Zusammenhang näherungsweise mithilfe einer linearen Gleichung aus. Erkläre, wie man zu dieser Gleichung mithilfe der Differenzialrechnung gelangt.

Zeichne zunächst die Messpunkte in ein Koordinatensystem ein. Du erhältst ein so genanntes **Streudiagramm** (eine **Punktewolke**). Das Diagramm zeigt den Zusammenhang in losen Punkten an und bestätigt die Tendenz, dass mit einer größeren Körpergröße auch eine größere Körpermasse verbunden ist.

Die Punkte liegen aber nicht auf einem erkennbaren Funktionsgraphen, weder auf einer Geraden noch auf einer Parabel noch auf einer exponentiellen Wachstumskurve.

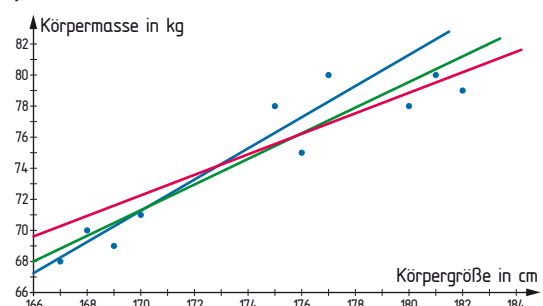
Mit Augenmaß könntest du eine Gerade in die Punktewolke hineinlegen, die den „Trend“ des Zusammenhangs grob beschreibt.

Das Problem: Alle 3 Geraden in der Grafik würden in Frage kommen. Wir brauchen jene Modellgerade, von der die Punkte so wenig wie möglich abweichen sollen.

Es stellt sich zunächst die Frage, was unter „**Abweichung**“ zu verstehen ist. Da du nicht weißt, wo genau die beste Gerade liegt, können ihre Funktionswerte nur allgemein mit $y = k \cdot x + d$ bezeichnet werden.

Als Abweichung der einzelnen Punkte $(x_i | y_i)$ von der von uns gesuchten Modellgeraden definieren wir den Abstand, den der einzelne y_i -Wert eines Punktes vom y -Wert der Modellgeraden hat, also $(y_i - y)$. Diesen vertikalen Abstand nennt man **Residuum**.

Die Residuen der einzelnen Punkte können positiv oder negativ sein, je nachdem, ob der einzelne Punkt oberhalb oder unterhalb der Geraden liegt. Beim Aufsummieren der Abweichungen heben sie sich gegenseitig auf. Durch Quadrieren kann diese Vorzeichen-Problematik beseitigt werden. Außerdem werden größere Abweichungen stärker als kleinere berücksichtigt. Also quadrieren wir die einzelnen Residuen und bilden ihre Summe s_n .



Auf der Bedingung, dass die Summe der quadratischen Abweichungen ein Minimum ist, beruht die **Gauß'sche „Methode der kleinsten Fehlerquadrate“**.

Erinnere dich, dass ein Funktionsgraph nur dann ein lokales Minimum aufweist, wenn die 1. Ableitung null ist. Das ist eine notwendige Bedingung für ein Minimum. Daher muss die Summe der quadratischen Abweichungen differenziert und null gesetzt werden.

Wir suchen das **k** und das **d** der Modellgeraden. Daher musst du die Fehlerquadratsumme einmal nach der Variablen **k** und einmal nach der Variablen **d** differenzieren. Beide Ableitungen werden null gesetzt. Das Gleichungssystem kann anschließend gelöst werden.

Es wird der Rechengang nun im Einzelnen vorgeführt. Dabei verwenden wir die Kurzform des

Summenzeichens ohne Indizes: $\sum x_i$ statt $\sum_{i=1}^n x_i$, **n** ist die Anzahl der Datenpunkte.

$$s_n = \sum (y_i - (k \cdot x_i + d))^2 \dots \text{Fehlerquadratsumme} \rightarrow \text{Minimum}$$

Ableitung von s_n nach **k**:

$$\begin{aligned} 2 \cdot \sum (y_i - (k \cdot x_i + d)) \cdot (-x_i) &= 2 \cdot \sum (-x_i \cdot y_i + k \cdot x_i^2 + d \cdot x_i) = \\ &= 2 \left(-\sum x_i \cdot y_i + k \cdot \sum x_i^2 + d \cdot \sum x_i \right) \end{aligned}$$

Ableitung von s_n nach **d**:

$$2 \cdot \sum (y_i - (k \cdot x_i + d)) \cdot (-1) = 2 \cdot \sum (-y_i + k \cdot x_i + d) = 2 \left(-\sum y_i + k \cdot \sum x_i + d \cdot n \right)$$

Setze nun beide Ableitungen = 0, dividiere durch 2 und ordne nach **k** und nach **d**:

$$\text{I: } k \cdot \sum x_i^2 + d \cdot \sum x_i = \sum x_i \cdot y_i$$

$$\text{II: } k \cdot \sum x_i + d \cdot n = \sum y_i$$

Mithilfe dieses Gleichungssystems kannst du **k** und **d** bestimmen.

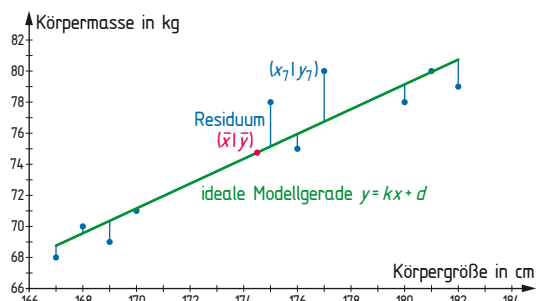
Wir erhalten für die angegebene Tabelle die folgenden Summen:

$$\sum x_i = 1\,745; \quad \sum x_i^2 = 304\,789; \quad \sum y_i = 748; \quad \sum x_i \cdot y_i = 130\,755$$

Eingesetzt in das Gleichungssystem ergibt dies:

$$\begin{cases} \text{I: } 304\,789k + 1\,745d = 130\,755 \\ \text{II: } 1\,745k + 10d = 748 \end{cases} \Rightarrow k = 0,799 \text{ und } d = -64,678$$

Die optimale Gerade – **die Regressionsgerade (= Trendlinie)** – hat daher die Gleichung $y = 0,799x - 64,678$.



In der Grafik ist die Regressionsgerade (= Trendlinie) grün eingezeichnet. Diese Linie ist die ideale Modellgerade, sie ist mithilfe der Gauß'schen Methode den Punkten optimal angenähert. Die Residuen, das sind die Abstände $(y_i - \bar{y})$ der Punkte von der idealen Geraden, sind eingezeichnet.

Die Trendlinie hat noch eine zusätzliche Besonderheit:

Der Punkt $(\bar{x} | \bar{y})$ liegt auf der Regressionsgeraden. Dabei ist $\bar{x}$ der arithmetische Mittelwert aller x_i -Werte und $\bar{y}$ der arithmetische Mittelwert aller y_i -Werte.



Rechengерäte bzw. Computer haben diesen Rechengvorgang im Statistik-Anwendungsprogramm gespeichert.

Wir setzen daher für die Berechnung der Regressionslinie ab nun stets Technologie ein.

TI-Nspire:

Wir geben die Werte beider Größen in zwei Listen l1 und l2 einer Tabellenkalkulation ein und verwenden **Menu/4: Statistik/1: Statistische Berechnungen/3: Lineare Regression (mx+b)**



1: Aktionen
2: Einfügen
3: Daten
4: Statistik
5: Wertetabelle
6: Tipps

Nicht gespeicherte

1: Statistische Berechnungen...
2: Statistische Verteilungen...
3: Konfidenzintervalle...
4: Statistische Tests...

4	170	71	b	-64.6782
5	175	78	r^2	0.873283
6	176	75	r	0.934496
7	177	80	Resid	(-0.80523...

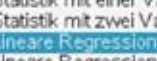


Figure 1: Screenshot of the SPSS 'Analyze' menu. The 'Regression' option is selected, opening a submenu. In this submenu, the 'Linear Regression (mx+b)...' option is highlighted with a mouse cursor. Other visible options in the submenu include 'Linear Regression (a+bx)...', 'Median-Median-Linie...', 'Quadratische Regression...', 'Kubische Regression...', 'Regression vierter Ordnung...', 'Potenzregression...', 'Exponentielle Regression...', 'Logarithmische Regression...', 'Sinusförmige Regression...', and 'Logistische Regression (d=0)...'. The background shows a portion of a data table with columns 'x' and 'y'.

Es öffnet sich ein Fenster, in das die Listen eingegeben werden können, in denen die x- bzw. die y-Werte der Punkte zu finden sind. In unserem Falle geben wir l1 und l2 ein, die als Variable zu verstehen sind (Variablenverweis!). In f1 wird die berechnete **Regressionslinie (= Trendlinie)** als Grafik gespeichert.

A	B	C	D	E
1	167	68	Titel	Lineare R.
2	168	70	RegEqn	$m \cdot x + b$
3	169	69	m	0.799302
4	170	71	b	-64.6782
5	175	78	r^2	0.873283

m steht hier für den Anstieg k und b steht für den Ordinatenabschnitt d . Wir erhalten die „beste Gerade“ durch die Punktwolke als **Regressionslinie** (= **Trendlinie**) mit $y = 0,799x - 64,678$.

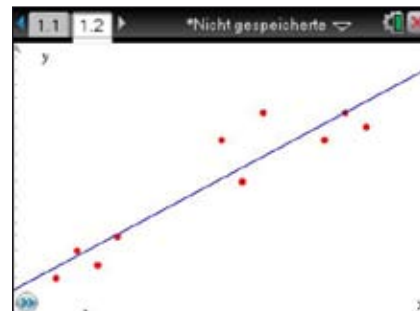
Streudiagramm mit TI-Nspire:

2: Graphs/Menu/3: Grafiktyp/

4: Streudiagramm/x...l1;y...l2

Zeichnen der Regressionslinie:

3: Grafiktyp/1: Funktion/ $f_2(x)=f_1(x)$



Technologieeinsatz zu diesem Thema für TI 82-84, EXCEL und Geogebra-CAS siehe www.hpt.at (Schulbuch Plus für Schüler/innen)

Mithilfe der Trendlinie können verschiedenste **Prognosen** erstellt werden:

Bezogen auf das Eingangsbeispiel zB: Wie viel Kilogramm wiegt ein Jugendlicher in deiner Klasse mit einer Körpergröße von 172 cm? Setze $x = 172$ in die Gleichung der Trendlinie ein.

Man erwartet nach dem berechneten Modell eine Körpermasse von $72,75 \text{ kg} \approx 73 \text{ kg}$.

Wie groß wäre ein Jugendlicher in deiner Klasse, wenn sein Gewicht 79 kg beträgt?

Setze 79 für y ein und berechne $x = 179,82 \text{ cm} \approx 180 \text{ cm}$.

Selbstverständlich handelt es sich hier nur um **Schätzwerte**, weil die tatsächlichen Messwerte im Allgemeinen nie exakt im Trend liegen.

- 2.17** Der Zusammenhang von 2 Merkmalen A und B wurde mit der folgenden Tabelle erfasst:

A	2	3	4	6	7	8	10	12
B	4	3	5	6	8	7	9	13

Berechne die Parameter der Regressionsgeraden.

Berechne den fehlenden Wert für $A = 9$ und für $B = 14$.

- 2.18** Der Zusammenhang von 2 Merkmalen C und D wurde mit der folgenden Tabelle erfasst:

C	1	3	5	6	9	11	14	15
D	10	12	15	14	18	23	27	30

Berechne die Parameter der Regressionsgeraden.

Berechne den fehlenden Wert für $C = 8$ und für $D = 32$.

- 2.19** Aus Angeboten verschiedener Hersteller wurden zufällig 10 Autos ausgewählt und ihre Leistung in Kilowatt (kW) mit der Spitzengeschwindigkeit in Kilometer pro Stunde (km/h) verglichen.

Leistung	44	74	55	136	128	74	59	110	85	115
Spitzen- geschwindigkeit	157	188	160	240	215	195	163	225	200	200

Erstelle den Zusammenhang der beiden Größen mithilfe einer linearen Regression.

Ermittle eine Prognose, welche Spitzengeschwindigkeit nach diesem Modell bei einer Leistung von 139 kW zu erwarten ist.

Schätze mithilfe des Modells ab, welche Leistung bei einer Spitzengeschwindigkeit von 220 km/h benötigt wird.

- 2.20** Für die Schwefeldioxid-Emission (SO_2) in Österreich liegen für eine Periode von 8 Jahren folgende Angaben (in 1 000 t) vor:

Jahr	1	2	3	4	5	6	7	8
SO_2	60,4	56,3	53,8	52,8	50,7	45,8	41,4	40,8

a) Zeichne die Punktwolke und ein Liniendiagramm.

b) Ermittle die Trendlinie.

Berechne die Vorhersagewerte für das 10. Jahr.

Erkläre, wie man aus der Trendlinie ermitteln kann, in welchem Jahr ein Wert von 35 000 t unterschritten wird.



B



B



AB



ABD



ABD



2.21 Das Bruttoinlandsprodukt (kurz BIP) wird als Indikator für die wirtschaftliche Leistung einer Volkswirtschaft verwendet. Es setzt sich aus dem ökonomischen Wert aller im Inland innerhalb eines bestimmten Zeitraums produzierten Güter und Dienstleistungen zusammen.

Für eine bestimmte bereits vergangene Periode lagen in Österreich für 6 Jahre die folgenden Daten für das BIP in Milliarden Euro vor:

Jahr	1	2	3	4	5	6
BIP	123,48	133,60	143,23	151,83	156,94	165,41

- Zeichne ein Liniendiagramm.
- Bestimme die Trendlinie.
Berechne die Vorhersagewerte für das Jahr 10.
- Suche auf der Website der STATISTIK AUSTRIA die aktuellen Werte der letzten 6 Jahre und berechne die Vorhersage für das 10. Jahr neu.

AB



2.22 Ein Lichtstrahl wird in eine Serie von verdünnten Lösungen geschickt. Man misst den Schwächungsgrad der Lichtintensität anhand des (dimensionslosen) Extinktionskoeffizienten E . Die Konzentration c der Lösungen ist in Mol pro Liter (mol/ℓ) angegeben.

Die Messung ergab die folgenden Werte:

Konzentration	0,01	0,05	0,1	0,5	1,0	1,5	2,0
Extinktion	0,003	0,016	0,03	0,15	0,3	0,45	0,6



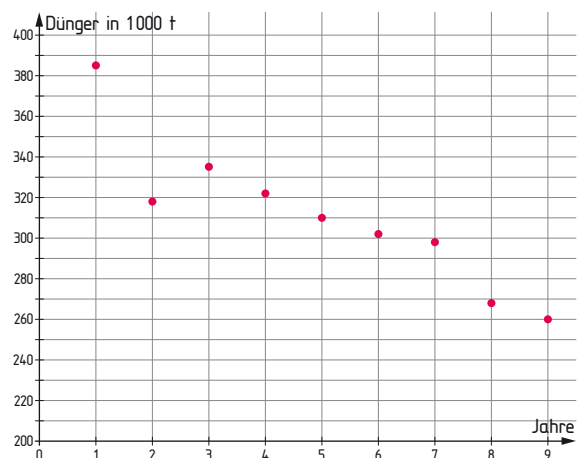
- Erstelle die Gleichung für den linearen Zusammenhang der beiden Größen.
- Schätze, welche Konzentration für eine Lösung mit der Extinktion $E = 0,22$ zu erwarten ist.
- Schätze, welche Extinktion für eine Lösung mit der Konzentration $c = 0,75 \text{ mol}/\ell$ zu erwarten ist.

ABD



2.23 In einer bestimmten Region liegen für 10 Jahre die Daten zum Verbrauch von Dünger in 1 000 Tonnen (t) in Form eines Streudiagramms vor:

- Erstelle die Gleichung der Trendlinie.
Zeichne sie in das Diagramm.
- Berechne die Vorhersagewerte für das 12. Jahr.
Erkläre, wie man aus der Trendlinie ermitteln kann, in welchem Jahr ein Wert von 250 000 t unterschritten wird.



2.2.2 Nichtlineare Zusammenhänge zweier Größen

Die Programme der Rechner liefern auch Regressionskurven für nichtlineare Zusammenhänge. Sie können daher mithilfe von Technologie im Prinzip genauso ermittelt werden wie lineare Zusammenhänge. Das einzige, was du an Wissen beisteuern musst, damit du die richtige Auswahl treffen kannst, ist die Grundkenntnis über den Verlauf einiger häufig vorkommender Kurven. Einige Beispiele, die du möglicherweise mithilfe deiner verfügbaren Technologie einer Punktwolke zuordnen kannst:

Polynomfunktion 2. Grads: $y = ax^2 + bx + c$... parabelförmig

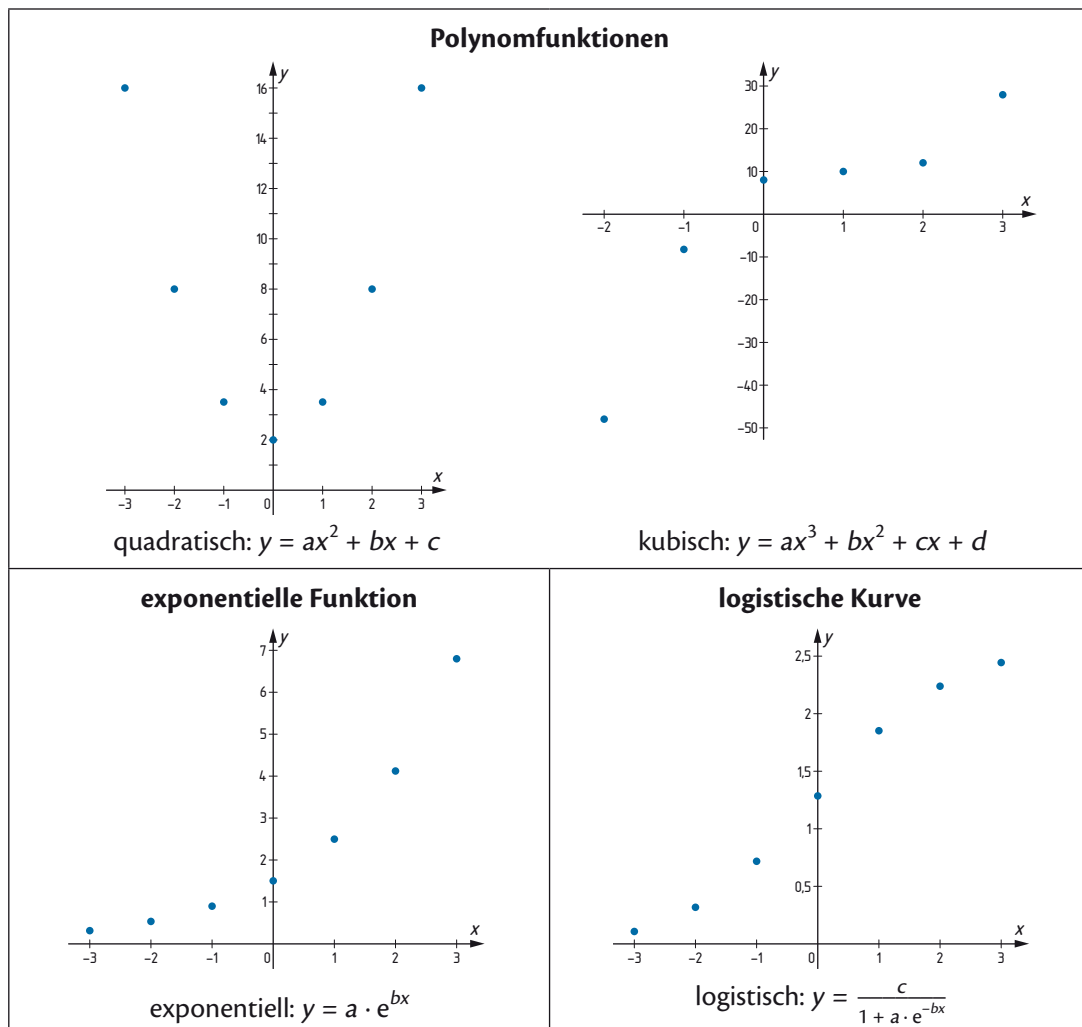
Polynomfunktion 3. Grads: $y = ax^3 + bx^2 + cx + d$... s-förmig

Exponentialfunktion: $y = a \cdot e^{bx}$... monoton steigend oder fallend mit einer waagrechten Asymptote

Logistisches Wachstum: $y = \frac{c}{1 + a \cdot e^{-bx}}$... zusammengesetzt aus einer progressiv steigenden, durch einen Wendepunkt gehenden und schließlich degressiv verlaufenden Kurve mit waagrechten Asymptoten.



Lists/4: Statistik/1: Statistische Berechnungen



Optimierung und Regression

ABD



2.24 Die Auswertung der Untersuchung des Merkmals Y in Abhängigkeit vom Merkmal X ergab die folgenden Messwerte:

X	1	2	3	4	5	6	7
Y	0,38	1,15	2,71	3,92	5,93	8,56	11,24

- Zeichne die Punktwolke.
- Ermittle eine lineare Regressionslinie dieser Punktwolke.
- Ermittle eine exponentielle Regressionslinie.
- Argumentiere, welche der beiden Linien **b)** oder **c)** optisch besser an die Punkte angenähert ist.

BD



2.25 Ein Betrieb hat die folgenden Daten zu den abgesetzten Mengen eines Produkts in Mengeneinheiten (ME) und dem Erlös in Geldeinheiten (GE):

Menge x	0	1	2	3	4	5	6
Erlöse E	0	127	199	227	196	124	10

- Zeichne ein Streudiagramm.
- Ermittle die Gleichung der Erlösfunktion E mithilfe einer quadratischen Regression. Zeichne die Trendlinie in das Streudiagramm. Beurteile anhand der Grafik die Güte der Anpassung.
- Berechne die obere Erlösgrenze.

BD



2.26 Die Gesamtkosten in Geldeinheiten (GE) bei der Erzeugung eines bestimmten Produkts wurden in Abhängigkeit von der Absatzmenge in Mengeneinheiten (ME) untersucht und ergaben die folgenden Daten:

Menge x	0	1	2	3	4	5	6
Kosten K	100	117	134	147	163	200	244

- Zeichne ein Streudiagramm.
- Ermittle die Gleichung der Kostenfunktion K mithilfe einer kubischen Regression. Zeichne die Trendlinie in das Streudiagramm. Beurteile anhand der Grafik die Güte der Anpassung.
- Berechne die Ableitung von K , dh die Grenzkostenfunktion K' .

ABD



2.27 Die Preisgestaltung in Euro pro Mengeneinheit (€/ME) eines Produkts hängt mit der auf dem Markt nachgefragten Menge in Mengeneinheiten (ME) zusammen.

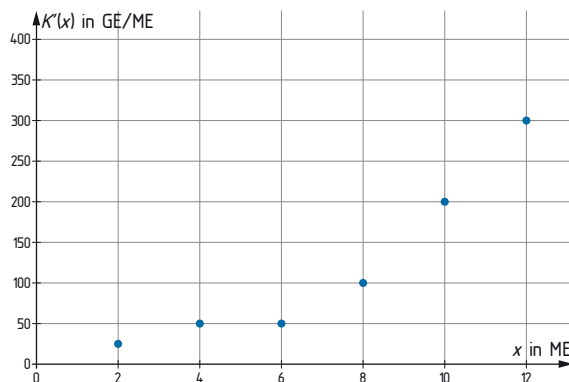
Man stellt für dieses Produkt eine diesbezügliche Untersuchung an und findet die folgenden Werte:

Menge	0	10	20	30	40	50	60	70	80
Preis	330	324	310	288	258	220	174	120	58

- Zeichne die Punktwolke.
- Erstelle die Gleichung der quadratischen Trendlinie. Berechne die Nullstelle dieser Funktion und interpretiere das Ergebnis. Argumentiere, weshalb die von dir berechnete quadratische Regression im Sachzusammenhang angemessen ist.

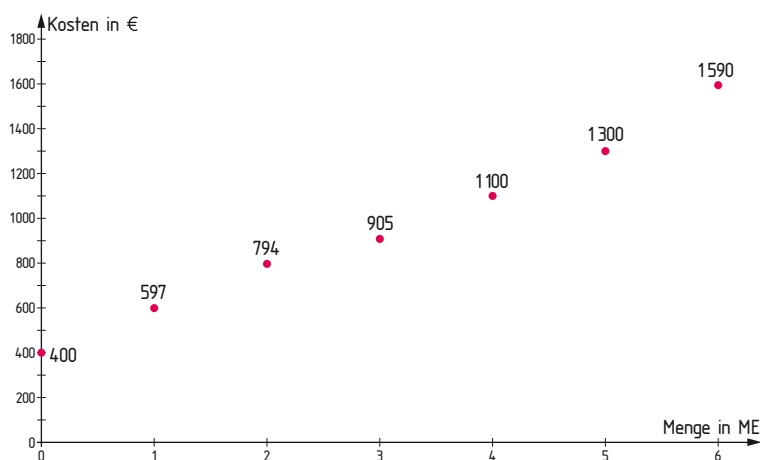


- 2.28** Die Gesamtkostenfunktion K beschreibt alle anfallenden Herstellkosten in Abhängigkeit von der Produktionsmenge x . Unter der Grenzkostenfunktion K' versteht man die Ableitungsfunktion von K . K' gibt näherungsweise Auskunft über die Kostenänderungen, wenn man die Produktionsmengen um eine Einheit steigert oder vermindert. Die Erhebung in einer Firma hat das folgende Streudiagramm für K' ergeben:



Erstelle anhand des Diagramms die quadratische Gleichung der Grenzkostenfunktion. Zeichne die Regressionslinie in das Diagramm. Beurteile anhand der Grafik die Güte der Anpassung an die Datenpunkte. Schätze ab, wo die Regressionslinie ein Minimum hat. Lies die Koordinaten des Minimums der Grenzkostenfunktion ab. Wie kann man diese Stelle interpretieren?

- 2.29** Die Gesamtkosten K , die bei der Erzeugung eines Produkts anfallen, sind in Abhängigkeit von der erzeugten Menge untersucht worden.
 x ... erzeugte Menge in Mengeneinheiten (ME)
 $K(x)$... Gesamtkosten in Euro (€) bei einer Produktion von x ME
 Die Untersuchung ergab das folgende Streudiagramm:



- a) Beschreibe, welche Funktionen in diesem Fall eventuell passend wären. Wähle die beste Kurve. Begründe deine Auswahl.
 b) Erstelle die Gleichung einer Polynomfunktion 3. Grads. Berechne die mittlere Änderungsrate der Funktion im Intervall $[2; 3]$. Was fällt dir bei dieser mittleren Änderungsrate auf?

Zusammenfassung

Optimierungsprobleme

Auf Optimierungsprobleme können die Kriterien, die für die Kurvenuntersuchung gelten, angewendet werden (siehe Abschnitt „1 Differenzieren“).

Lokale Extremwerte erfüllen die notwendige Bedingung, dass die erste Ableitung an dieser Stelle gleich null ist. Hinreichende Bedingungen sind:

Für das **lokale Maximum** muss zusätzlich die 2. Ableitung an dieser Stelle kleiner als null sein, für das **lokale Minimum** ist die 2. Ableitung an dieser Stelle größer null.

Ist der Definitionsbereich ein abgeschlossenes Intervall $[a; b]$, so kann ein **globaler Extrempunkt** auch am Rand des Intervalls liegen.

Man spricht in diesem Fall von einem **Randmaximum** bzw. bei einem kleinsten Wert von einem **Randminimum**. An dieser Stelle liegt im Allgemeinen keine waagrechte Tangente vor.

Extremwertaufgaben mit Nebenbedingungen sind Optimierungsaufgaben, bei denen der größtmögliche bzw. kleinstmögliche Wert einer Größe ermittelt wird, die von 2 oder mehreren Variablen abhängt.

Die Funktion, die ein Maximum oder Minimum annehmen soll, wird **Hauptbedingung (= Zielfunktion)** genannt. Zusammenhänge, die das Aufstellen der Zielfunktion in **einer Variablen** ermöglichen, werden **Nebenbedingungen** genannt.

Die **Regressionsrechnung** erfasst mathematisch den Zusammenhang zwischen zwei messbaren Merkmalen X und Y.

Hat man die Daten in einer **Punktwolke** oder in einem **Streudiagramm** dargestellt, erkennt man in vielen Fällen bereits die Funktion, die den Zusammenhang zwischen den Variablen am besten wiedergibt.

Falls sich die Daten gut linear annähern lassen, so nennen wir die Gerade **Regressionsgerade** oder **Trendlinie** von Y bezüglich X. Die Abweichung $(y_i - y)$ des Funktionswerts y der Trendlinie von einem einzelnen y_i -Wertes eines Datenpunktes nennt man **Residuum**.

Mathematisch gesehen muss man sich für die Berechnung der **Parameter k** und **d** der Geraden eines objektiven Verfahrens bedienen, zB des **Gauß'schen Prinzips der kleinsten Fehlerquadrate**:

Es ist die Gerade so zu legen, dass die Summe der Quadrate der vertikalen Abstände aller Punkte von der Geraden möglichst klein ist.

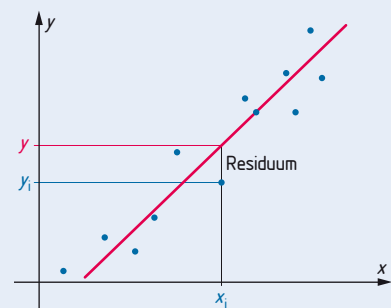
Mit Methoden der Differenzialrechnung wird ein Gleichungssystem gefunden, mit dessen Hilfe man den Anstieg k und den Ordinaten-Abschnitt d der Geraden berechnen kann.

Die Parameter der idealen Modellgeraden, der **Regressionsgeraden (= Trendlinie)**, erhält man aus:

$$\text{I: } k \cdot \sum x_i^2 + d \cdot \sum x_i = \sum x_i \cdot y_i$$

$$\text{II: } k \cdot \sum x_i + d \cdot n = \sum y_i$$

Mithilfe von Technologieeinsatz können lineare und nichtlineare Zusammenhänge zwischen 2 Größen durch die Regressionsrechnung ermittelt werden. Allerdings darf nicht übersehen werden, dass die Trendlinie nur ein angenähertes Modell zur Beschreibung der Daten darstellt.



Vermischte Aufgaben zur Vertiefung

2.30 Skizziere jeweils die Funktionen f mit $f(x) = \frac{x^2}{4}$ und g in ein gemeinsames Koordinatensystem. Vergleiche die Lage der Extremstellen. Erkläre die Veränderung.

a) $g(x) = 2 \cdot f(x)$ b) $g(x) = f(x) - 2$ c) $g(x) = [f(x)]^2$ d) $g(x) = \frac{1}{f(x)}$

2.31 Berechne das Seitenverhältnis eines Rechtecks mit dem Umfang $u = 18$ cm, wenn die Länge der Diagonale möglichst kurz sein soll.

2.32 Aus 82 gleich langen Zündhölzern wird ein Rechteck gelegt. Ermittle, wie viele Zündhölzer für die Länge bzw. Breite des Rechtecks benötigt werden, damit der Flächeninhalt maximal wird.

2.33 Für ein bestimmtes Gut wurden die anfallenden Gesamtkosten K in Euro (€) in Abhängigkeit von der erzeugten Menge x in Mengeneinheiten (ME) festgestellt:

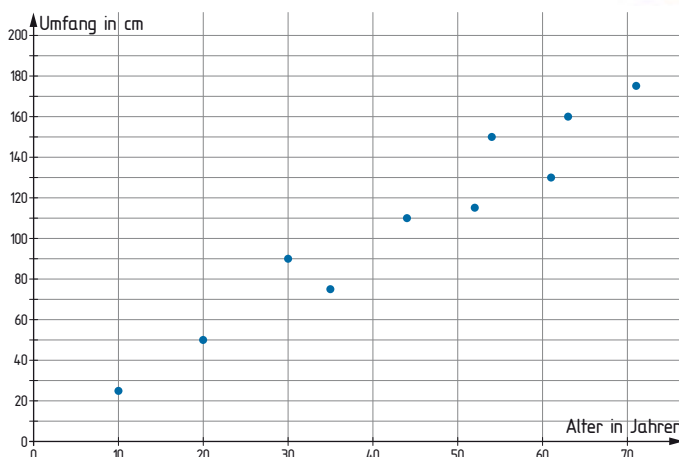
Menge x	100	200	300	400	500
Kosten K	3.450	5.210	7.400	9.180	10.940

a) Zeichne die Punktwolke.

Ermittle eine lineare Trendlinie und zeichne sie in die Punktwolke ein.

b) Schätze mithilfe der Trendlinie, wie hoch die Gesamtkosten bei einer Produktion von 600 ME des Guts wären.

2.34 Es wird der Umfang von verschiedenen alten Bäumen der gleichen Art gemessen. Die Messpunkte sind in das Streudiagramm eingezeichnet. Lies die Punkte aus dem Diagramm ab und fertige für die weiteren Aufgaben eine Tabelle an.



a) Skizziere die Regressionsgerade mit Augenmaß in das Streudiagramm.

Lies die Gleichung der Regressionsgeraden ab.

b) Berechne die Trendlinie mithilfe der abgelesenen Werte.

c) Beurteile, ob eine nichtlineare Kurve den Trend ebenfalls beschreibt.

Gib die deiner Meinung nach „beste“ Näherungskurve an.

BCD

TE

AB

TE

AB

TE

AB

TE

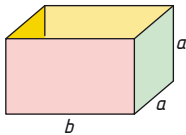
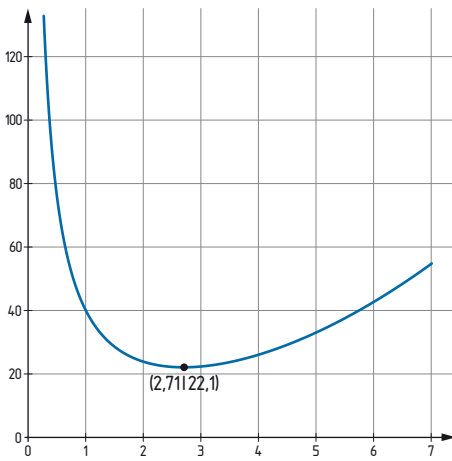
ABD

TE

Optimierung und Regression

Wissens-Check

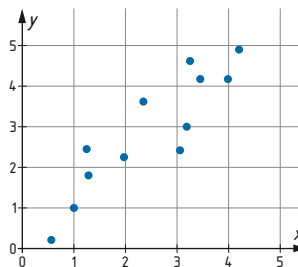
Bearbeite die Aufgaben. **Begründe** jeweils deine Auswahl.

		gelöst
1	Eine oben offene Kartonschachtel hat 2 quadratische Seitenflächen mit der Seitenlänge a sowie die Längsseite b . Ordne den gegebenen Extremwertaufgaben 1 und 2 jeweils die richtige Nebenbedingung aus A bis D zu.	
		
	1	Die Schachtel mit $V = 8 \text{ dm}^3$ soll einen minimalen Kartonverbrauch haben.
	2	Die Schachtel mit $O = 8 \text{ dm}^2$ soll ein minimales Fassungsvermögen haben.
	A	$a^2 = \frac{8}{b}$
2	C	$8a^2 = b$
	B	$2a^2 + 4ab = 8$
	D	$2a + 3b = \frac{8}{a}$
	Für eine Extremwertaufgabe wurden die folgenden Bedingungen ermittelt: $s^2 \cdot h = 10$ und $A = s^2 + 4sh \Rightarrow \text{Minimum (h in cm)}$ Um die Aufgabe zu lösen wurde die folgende Grafik erzeugt:	
		
Um welchen geometrischen Körper handelt es sich?		
Beschrifte die Achsen dieser Grafik korrekt.		
Kreuze die richtige Aussage dazu an. Ergänze die fehlenden Einheiten.		
A	☐	Minimum: $s = 2,71$ _____ ; $h = 22,1$ _____
B	☐	Minimum: $s = 2,71$ _____ ; $A = 22,1$ _____
C	☐	Minimum: $h = 2,71$ _____ ; $s = 22,1$ _____
D	☐	Minimum: $A = 2,71$ _____ ; $h = 22,1$ _____
E	☐	Minimum: $A = 2,71$ _____ ; $s = 22,1$ _____

gelöst

Ergänze die Textlücken durch Ankreuzen der jeweils richtigen Satzteile so, dass eine korrekte Aussage entsteht.

Das dargestellte Streudiagramm lässt sich durch die Regressionsgerade f mit $f(x) = k \cdot x + d$ annähernd beschreiben. Für diese Gerade gilt ①, weil dann ② ist.



3

①		②	
$k = -1$ und $d = 5$	A ☐	die Summe der quadratischen Abweichungen der Punkte von der Geraden f minimal	A ☐
$k = 1$ und $d = 0,27$	B ☐	die Summe der Abweichungen der Punkte von der Geraden f minimal	B ☐
$k = 2$ und $d = -0,5$	C ☐	die Summe der Abweichungen der Punkte von der Geraden f null	C ☐

4

Eine Regressionsgerade hat die Gleichung $y = 0,7x + 20$. Kreuze den richtigen Wert des Residuums für den Messpunkt (90|80) an.

A ☐	B ☐	C ☐	D ☐	E ☐
3	-3	-2	2	-1

5

Eine Datenmenge aus den Merkmalen X und Y lässt sich mit einer Regressionsgeraden darstellen, deren Steigung 1,5 beträgt. Der Punkt (10|9) liegt auf der Geraden. Kreuze an, wie groß der Achsenabschnitt d der Regressionsgeraden ist.

A ☐	$d = -6$	D ☐	$d = 8$
B ☐	$d = -8$	E ☐	$d = 6$
C ☐	$d = 0$		

6

Die Gleichung der quadratischen Regressionslinie lautet $y = 1,5x^2 + 3$. Kreuze den richtigen Wert des Residuums für den Messpunkt (4|26) an.

A ☐	B ☐	C ☐	D ☐	E ☐
-3	-2	-1	0	1

CLIL-Review: Optimization and regression

Get in pairs and work on the following tasks (2.E1, 2.E2). You may use your formula collection and an online dictionary.

Please find important vocabulary and other supporting material to these tasks as well as solutions and example answers in the solutions book. For further information have a look at page 2 of this book.



BCD

2.E1 The following task is a maximum/minimum optimization problem. It illustrates one of the most important applications of the first derivative. Perhaps you will find this task difficult because there are so many “words” to transfer into mathematical language. It might be easier for you to follow these guidelines¹⁾:



1. **Read the problem slowly and carefully.** It is important to know exactly what the problem is asking. If you misread the problem or hurry through it, you have NO chance of solving it correctly.



2. **If possible draw a sketch of the problem** to be solved. Pictures are a great help in sorting out your thoughts.



3. **Define variables** to be used and carefully **label your diagram** with these variables. This step leads to the creation of mathematical equations.



4. Write down all equations which are related to your problem or diagram. Clearly **denote that equation which you are asked to maximize or minimize.**



Experience will show you that MOST optimization problems will begin with two equations. One equation is a “**constraint**” equation and the other is the “**optimization**” equation. The “constraint” equation is used to solve for one of the variables. This is then substituted into the “optimization” equation before differentiation occurs. Some problems may have NO constraint equation. Some problems may have two or more constraint equations.



5. **Before differentiating, make sure that the optimization equation is a function of only one variable.** Then differentiate using the rules of differentiation.

6. **Verify that your result is a maximum or minimum value** using the second derivative test for extremes.

Solve the following task on your own by following the guidelines above. Then compare your solution with your partner. Prepare this topic for a presentation in class.

A closed rectangular box is made of very thin sheet metal. Its volume is 416 cm^3 . Its length is 4 times as long as its width (w). Obtain an expression for its height.

Show that the total surface area A is equal to $8w^2 + \frac{1040}{w} \text{ cm}^2$.

Calculate the dimensions of the box for which the surface area is a minimum.

Explain why your answer gives a minimum rather than a maximum.

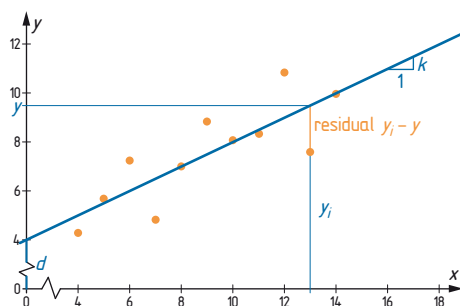
¹⁾ created by Duane Kouba,

<https://www.math.ucdavis.edu/~kouba/CalcOneDIRECTORY/maxmindirectory/MaxMin.htm>

2.E2 First read the text about linear regression. Look up all the mathematical terms that are new to you. Get familiar with the content.

If the n pairs (x_i, y_i) for an experiment are plotted on a graph, then the points are scattered about a straight line. The method of least squares fits the best line to the points by making the sum of the squared deviations $(y_i - y)$ of each point from this line a minimum. This line is called the **line of regression** and is used for predicting y -values by given x -values.

The equation of the line of regression is $y = k \cdot x + d$. It can be calculated by the statistic programme of your calculator or computer.



The differences $(y_i - y)$ between the observed values and the fitted values of the regression line are called **residuals**.

Plotting the residuals to the regression line is a check for the adequacy of the model used. If the plot of residuals shows a small random scatter, this is a good indication that the regression line is an adequate model.

Solve the following task on your own.
Then compare your solution with your partner.
Prepare this topic for a presentation in class.

Carbon dioxide is an important greenhouse gas. The burning of carbon-based fuels has rapidly increased the concentration of CO_2 in the atmosphere.

The maximum value of atmospheric carbon dioxide in parts per million (ppm) for each given year is shown in the table below.



Year x	1995	1998	2001	2004	2007	2010	2013
Carbon dioxide y	318	320	322	325	329	333	335

Plot these points on a scatter diagram.

Calculate the equation of the line of regression with the statistic programme of your calculator.

Draw the line of regression into the scatter diagram.

Estimate the amount of carbon dioxide, when $x = 2020$.