2 Auswertung von Messdaten

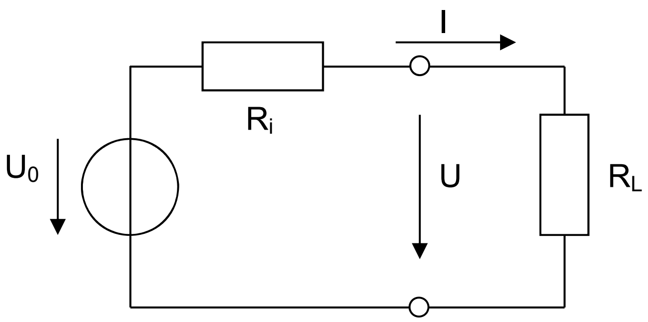
2.1 Messreihendarstellung, Regression und Parameterbestimmung

2.1.1 Beispiel "Belastungskennlinie"

Die Belastungskennlinie ist kennzeichnend für die Qualität einer Spannungsquelle. Aus ihr lassen sich die Leerlaufspannung, der Innenwiderstand und der Kurzschlussstrom entnehmen.

Messtechnisch wird die Belastungskennlinie bestimmt, indem eine variable Last an die Klemmen der Spannungsquelle angeschlossen wird. Durch Verändern der Last werden unterschiedliche Lastströme eingestellt und die jeweils zugehörigen Klemmenspannungen der Quelle gemessen.

Messschaltung und beispielhafte Messreihe:



Laststrom I [A]	Spannung U [V]
0,4	13,9
0,8	13,6
1,2	13,4
2,0	13,3
3,0	12,7
4,5	12,3

Der Schaltung entnimmt man die Gleichung

$U = U_0 - R_i \cdot I$ Bestimmung von R_i aus $(I_1, U_1), (I_2, U_2)$: $U_1 = U_0 - R_i \cdot I_1$

$U_2 = U_0 - R_i \cdot I_2$

Bei der graphischen Darstellung wird man die Messwerte also auf einer Geraden $U_1 - U_2 = -R_i (I_1 - I_2)$

$U(I) = m \cdot I + b$

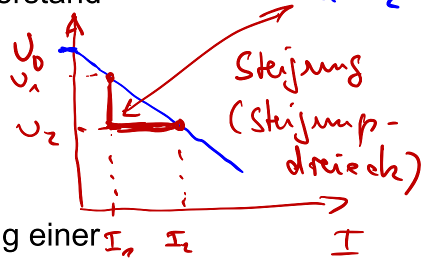
$R_i = -\frac{U_1 - U_2}{I_1 - I_2}$

erwarten. Dabei ist die Steigung gleich dem negativen Innenwiderstand

$m = -R_i$ bzw. $R_i = -m$

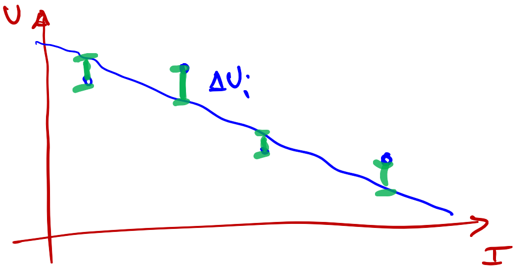
und der y-Achsabschnitt gleich der Leerlaufspannung

$U_0 = b$.



Theoretisch genügen schon zwei Messungen für die Bestimmung einer Geradengleichung. Aber bei jeder praktischen Messung weichen die Messwerte vom idealen Verlauf ab. Die Aufgabe ist dann, eine Gerade zu suchen, die "irgendwie bestmöglich" zu den aufgenommenen Messwerten passt.

Dazu bildet man die Differenzen zwischen den gemessenen Spannungen U_i und den Spannungswerten, die die Geradengleichung bei dem entsprechenden Strom I_i liefert, also



$$\Delta U_i = U_i - (m \cdot I_i + b)$$

m und b werden dann so bestimmt, dass

$$\sum_i (\Delta U_i)^2$$

Minimierungsproblem für eine Funktion von zwei Variablen

also die Summe der Quadrate dieser Differenzen minimal wird. Dieses Verfahren heißt lineare Regression und die so bestimmte Gerade heißt **Ausgleichsgerade**.

Die Suche nach Funktionen, die Datenreihen bestmöglich approximieren – engl.: "best fit" – ist nicht auf Geraden beschränkt. Mit der MATLAB-Function **`polyfit`** lassen sich Datenreihen beispielsweise durch **beliebige Polynome** annähern.

Zur obigen Messaufgabe lässt sich die Ausgleichsgerade mit folgendem MATLAB-Code berechnen und darstellen.

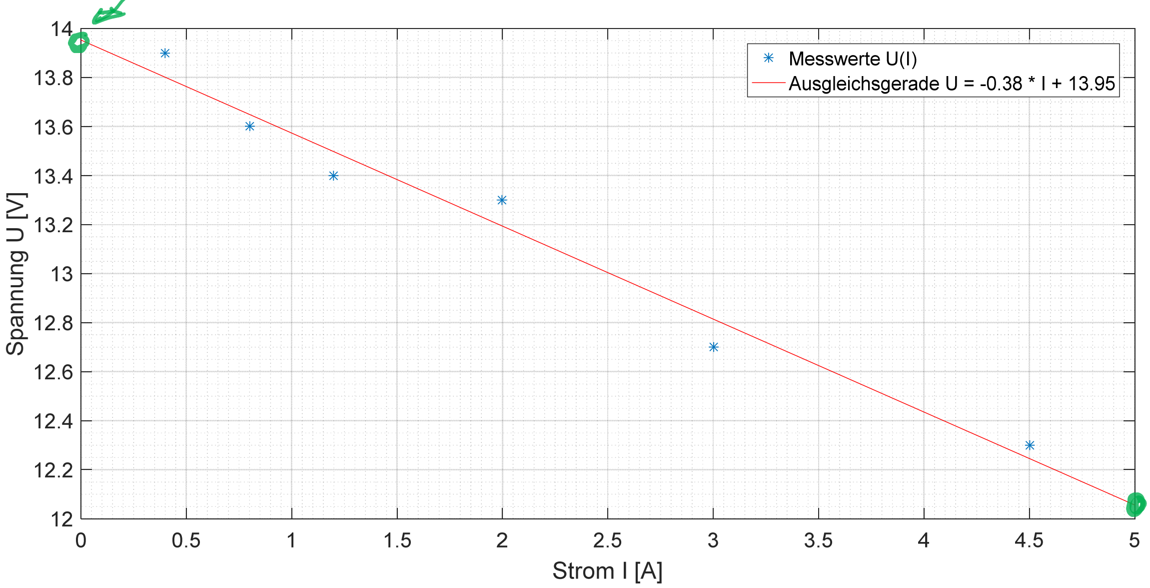
```
%% Messwertpaare (Strom, Spannung) als Spaltenvektoren
I = [0.4; 0.8; 1.2; 2.0; 3.0; 4.5];
U = [13.9; 13.6; 13.4; 13.3; 12.7; 12.3];

%% Ausgleichsgerade mittels polyfit
p = polyfit(I,U,1)
m = p(1), b = p(2)

%% Grafik
plot(I,U,'*', [0;5],polyval(p,[0;5]),'r')
grid minor
set(gca,'xminorgrid','on','yminorgrid','on')
xlabel('Strom I [A]'), ylabel('Spannung U [V]')
legend('Messwerte U(I)',sprintf('Ausgleichsgerade U = %0.2f * I + %0.2f',m,b))
```

$\begin{bmatrix} 0{,}4 \\ 0{,}8 \\ \vdots \\ 4.5 \end{bmatrix}$

Polynom 1. Grades, also Geradenglg

$U(I) = p(1) \cdot I + p(2)$

$p = \begin{bmatrix} p(1) \\ p(2) \end{bmatrix}$



Aus der Messung ermittelt man auf diese Weise also die Leerlaufspannung $U_0 \approx 13{,}95\,V$ und den Innenwiderstand $R_i \approx 0{,}38\,\Omega$.

Die Streuung der Messwerte um die gesuchte Gerade bedeutet, dass diese Werte – wie alle Messwerte – eine Unsicherheit aufweisen.

Hier: $\Delta U_0 \approx 0{,}08\,V$ und $\Delta R_i \approx 0{,}032\,\Omega$

(Berechnung siehe nächste Seite)

2.1.2 "Best Fit", Lineare Regression, Ausgleichsgerade

Aufgabenstellung:

Sei eine Menge von N Wertepaaren (x_i, y_i) gegeben, wobei weder die x_i noch die y_i jeweils alle untereinander gleich sind.

Bestimme die Koeffizienten des Polynoms

$$y = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

so, dass der Graph des Polynoms möglichst gut zu den Wertepaaren "passt".

Das geschieht durch Minimieren des Gütemaßes

$$P = \sum_{i=1}^{N} (y(x_i) - y_i)^2$$

"Methode der kleinsten Quadrate" / "Least Squares" (C. F. Gauss, 1777 – 1855).

1.) Ausgleichsgerade

Für eine Ausgleichsgerade

$$y = mx + b$$

ist dies eine Optimierungsaufgabe mit 2 Variablen mit den Lösungen

$$\boxed{\begin{gathered} m = \frac{s_{xy}}{s_x^2} = \frac{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})}{\underline{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2}} \\ b = \bar{y} - m\bar{x} \end{gathered}}$$

eventuell Bekannt

$$S = \sqrt{\frac{1}{N-1} \sum_i (x_i - \bar{x})^2}$$

Standardabweichung

Diese Berechnung verwendet die **Mittelwerte**

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} y_i \ , \quad \bar{y} = \frac{1}{N}\sum_{i=1}^{N} x_i \ ,$$

die **Kovarianz**

$$s_{xy} = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})(y_i - \bar{y})$$

cov(x,y) liefert $\begin{bmatrix} s_x^2 & s_{xy} \\ s_{xy} & s_y^2 \end{bmatrix}$

und die **Varianz** der x-Werte

$$s_x^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2 .$$

Die Varianz ist das Quadrat der Standardabweichung

$$s_x = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - \bar{x})^2} .$$

Die Ausgleichsgerade hat damit einen starken Bezug zu klassischen Kenngrößen der Statistik.

Unsicherheiten

Da die Wertepaare um die Ausgleichsgerade streuen, werden auch die Steigung und der Achsabschnitt als Zufallsgrößen mit einer Standardabweichung aufgefasst.

Es gilt

$$s_m = \frac{s_y}{s_x}\sqrt{\frac{1-B_{xy}}{N-2}}$$

$$s_b = s_m\sqrt{\frac{1}{N}\sum_{i=1}^{N}x_i^2} .$$

Dabei werden zusätzlich die Standardabweichung der y-Werte

$$s_y = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(y_i - \bar{y})^2}$$

der **Korrelationskoeffizient**

$$r_{xy} = \frac{s_{xy}}{s_x \cdot s_y}$$

$-1 \leq r_{xy} \leq 1$

und das **Bestimmtheitsmaß**

$$B_{xy} = r_{xy}^2$$

$0 \leq B \leq 1$, $B \lessapprox 1$ bedeutet gute Annäherung der Wertepaare durch eine Gerade.

benötigt.

2.) Kurvenapproximation durch beliebige Polynome

$P(x) = a_n x^n + a_{n-1} x^{n-1} + \ldots + a_1 x + a_0$

Die Forderung, dass die Messpunkte (x_i, y_i) alle die Polynomgleichung erfüllen sollen, liefert ein – in der Regel überbestimmtes – Gleichungssystem.

$$\begin{bmatrix} x_1^n & x_1^{n-1} & \cdots & x_1 & 1 \\ x_2^n & x_2^{n-1} & \cdots & x_2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_N^n & x_N^{n-1} & \cdots & x_N & 1 \end{bmatrix} \begin{pmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_1 \\ a_0 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}$$

$x_1^n \cdot a_n + x_1^{n-1} \cdot a_{n-1} + \ldots + x_1 \cdot a_1 + 1 \cdot a_0 = y_1$

$P(x_1) = y_1$

Das lässt sich kurz in der Form

$$X \cdot \vec{a} = \vec{y}$$

schreiben und das Gütemaß lautet dann

$$\boxed{P = \sum_{i=1}^{N} (y(x_i) - y_i)^2} = (X \cdot \vec{a} - \vec{y})^T (X \cdot \vec{a} - \vec{y})$$

$\Rightarrow \sum_{i=1}^{N} (P(x_i) - y_i)^2$

vergleiche $u^T u = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \cdot \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = u_1^2 + u_2^2 + \ldots u_n^2$

$$= \vec{a}^T \cdot X^T X \cdot \vec{a} - 2\vec{a}^T \cdot X \cdot \vec{y} + \vec{y}^T \vec{y}$$

Durch mehrdimensionales (vektorielles) Ableiten nach $\vec{a}$ und Nullsetzen erhält man für den optimalen Koeffizientenvektor das (eindeutig lösbare) Gleichungssystem

$$X^T X \cdot \vec{a} = X^T \vec{y}\,.$$

$X^T X$ ist eine quadratische Matrix

Diese Methode benutzt zum Beispiel die MATLAB-Funktion `polyfit`.

Für das Polynom 1. Ordnung liefert diese Rechnung die Parameter der Ausgleichsgerade wie oben.

$$\begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_N & 1 \end{bmatrix} \begin{pmatrix} a_1 \\ a_0 \end{pmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$X^T X = \begin{bmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & N \end{bmatrix}$$

$$X^T Y = \sum x_i \cdot y_i$$

etc.

2.1.3 Approximation von Messreihen durch allgemeine Funktionen

Viele Zusammenhänge lassen sich nicht sinnvoll durch Polynome annähern. Man kann dann auch andere Funktionsansätze wählen und die Parameter durch numerische Optimierung so bestimmen, dass die Funktion die Menge der Wertepaare im Sinne eines "best fit" möglichst gut annähert.

Beispiel:

Laut Theorie folgt die Temperaturabhängigkeit eines NTC (Widerstand mit negativem Temperaturkoeffizienten) der Funktion

von Widerständen wird häufig als lineare Näherung angegeben

$R(\vartheta) = R(20°C) \cdot (1 + \alpha(\vartheta - 20°C))$

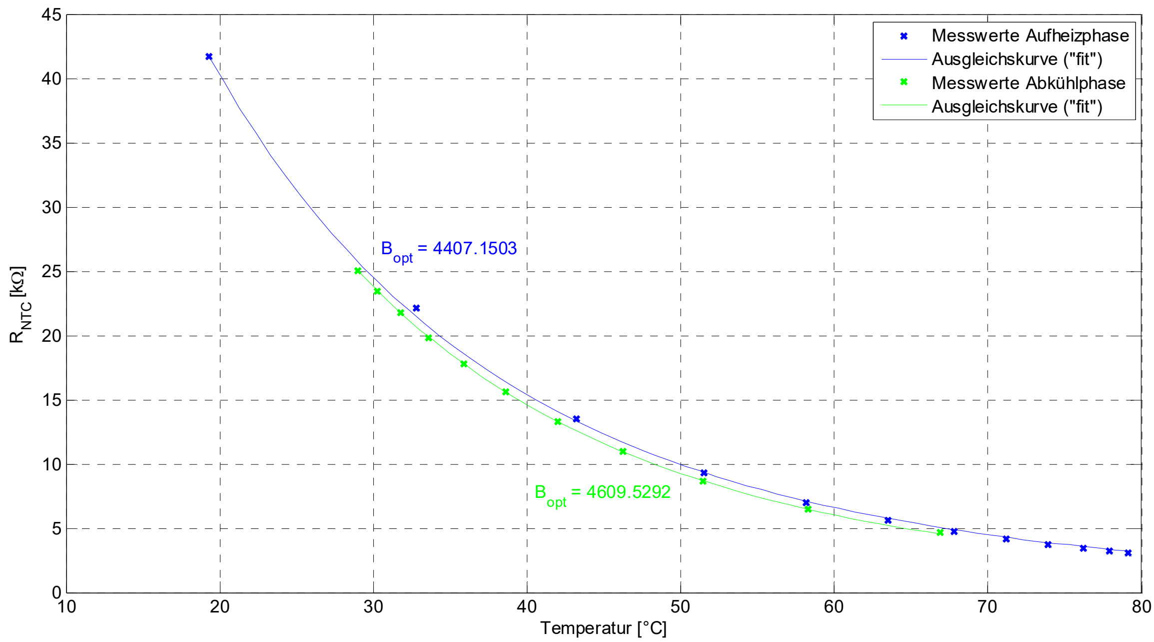
Temperaturkoeffizient

$$R(T) = R(T_0) \cdot e^{B \cdot \left(\frac{1}{T} - \frac{1}{T_0}\right)}$$

mit Angabe der Temperaturen jeweils in Kelvin.

Der Widerstand $R(T_0)$ bei der Temperatur T_0 und eine Reihe weiterer Wertepaare (T_i, R_i) sei gegeben. Dann kann der Parameter B durch Optimierung so bestimmt werden, dass die Funktion $R(T)$ die Wertepaare bestmöglich wiedergibt.

In MATLAB ist dies z. B. mit der Funktion `lsqcurvefit` möglich. Für die obige Aufgabe erhält man bei einer typischen Messreihe die unten stehende Lösung.



Dieses Beispiel (`rntc_lsqfit.m`) und diverse weitere sind in AULIS bereitgestellt.

`lsqcurvefit` bestimmt durch einen Optimierungsalgorithmus die freien Parameter in dem gewählten Funktionsansatz so, dass der Graph der Funktion sich bestmöglich (im Sinne kleinster Abweichungsquadrate) an die Messpunkte anpasst.

Die Optimierung führt umso sicherer zum Erfolg je bessere, vorab geschätzte Startwerte für die gesuchten Parameter angegeben werden.

2.2 Übungen "Ausgleichsgerade, Kurvenapproximation (Fit)"

2.2.1 Erstellen von Diagrammen

Die Datei "Temperaturmessung-10-2008.dat" enthält (fiktive) Messdaten eines Temperaturmessversuchs..

```
# Temperaturmessung 10/2008
# Zeit    T_Diode         R_NTC          T_U
# (min)     (°C)         (OHM)           (C)
   0       19.30       41735.0       21.09
   5       32.79       22142.0       22.63
  10       43.23       13503.0       23.85
```

etc.

Das Format ist für die Speicherung von Messreihen typisch. Jede Zeile besteht aus dem Zeitpunkt und den zu diesem Zeitpunkt erfassten Werten verschiedener Messgrößen. Alle Einträge sind durch Tabulator, Komma, Semikolon oder einfach durch Leerzeichen getrennt. Für jede Messgröße finden sich die aufgenommenen Werte also in einer Spalte der Datei wieder. Am Dateianfang kann es beschreibende Textzeilen geben, die mit einem Kommentarzeichen (//, #, %, …) beginnen.

MATLAB bietet für das Einlesen der Daten die Funktion `textscan.` (In einigen bereitgestellten Beispielen wurde allerdings das ältere `textread` verwendet.)

Übung:

Erzeugen Sie zur oben genannten Datendatei ein Diagramm, in dem die zweite und die vierte Spalte über der Zeit dargestellt werden.

Das Diagramm soll üblichen messtechnischen Anforderungen genügen, also mit **Raster, Achsbeschriftungen und aussagekräftige Legende** (in MATLAB: `xlabel, ylabel, legend` und `grid` oder auch `grid minor`.

2.2.2 "Fit", Ausgleichsfunktion zu einer Messreihe

Aufgabe: Es ist ein Diagramm zu erstellen, in dem der ansteigende Teil des Temperaturverlaufs "T_Diode" dargestellt und durch den Graphen einer Funktion angenähert wird, die dem theoretisch zu erwartenden Verlauf

$$\vartheta(t) = \vartheta_0 + (\vartheta_{end} - \vartheta_0)\left(1 - e^{-\frac{t}{\tau}}\right)$$

entspricht.

Anmerkungen:

- Dies gehört auch zu den Auswertungen im Labor.
- Eine Lösung dieser Aufgabe in MATLAB finden Sie auf AULIS in "`Theta_lsqfit.m`". Benutzt wird die Funktion `lsqcurvefit` aus der Optimization Toolbox.

2.2.3 Belastungskennlinie, Innenwiderstand

Untersucht wird eine nicht-ideale Spannungsquelle (vergl. Kap. 2.1.1).

a) Warum eignet sich die Formel $R_i = U_0 / I_k$ (Verhältnis von Leerlaufspannung zu Kurzschlussstrom) in der Praxis kaum zur Bestimmung des Innenwiderstands?

b) Wie geht man stattdessen vor, um mit einer minimalen Zahl von Messungen die nötigen Informationen zur Bestimmung des Innenwiderstands zu bekommen? Wie wird der Innenwiderstand dann berechnet?

c) Wenn mehr Messungen vorliegen, ist die Bestimmung des Innenwiderstandes mittels Ausgleichsgerade die Methode der Wahl. Bestimmen Sie die Ausgleichsgerade zur folgenden Messreihe:

I = 0,2 / 0,3 / 0,5 / 0,7 / 0,8 / 1,0 A

U = 33 / 31 / 29 / 27 / 25 / 23 V

d) Wie errechnet sich hieraus der Innenwiderstand? Wie groß ist er?

e) Geben Sie ein Maß für die Unsicherheit des berechneten Wertes an.

f) Wie groß ist die Leerlaufspannung? Unsicherheit?

2.2.4 Leerlaufspannung und Innenwiderstand einer Spannungsquelle

An einer Spannungsquelle wird die Spannung bei verschiedenen Strömen wie folgt ermittelt:

Strom I [A]	4	12	16	28
Spannung U [V]	10	8,5	7,5	6

Bestimmen Sie anhand der Ausgleichsgeraden die Leerlaufspannung und den Innenwiderstand. (Empfehlung zur eigenen Übung: Handrechnung OHNE PC oder Taschenrechner!)

2.2.5 Kurzschlussstrom einer Solarzelle

Eine Solarzelle liefert bei den aufgeführten Bestrahlungsstärken die Kurzschlussströme

E_e [W/m²]	1000	750	500	250	100
I_{SC} [A]	2.5	1.8	1.3	0.6	0.2

a) Welche mittlere Empfindlichkeit in [A/(W/m²)] lässt sich daraus bestimmen?

b) Geben Sie ein Maß für die Unsicherheit dieses Wertes an. (siehe S. 39)

2.2.6 Ausgleichsgerade, Bestimmtheitsmaß, Unsicherheit der Steigung

Ein Messverstärker habe ein Übertragungsverhalten entsprechend der Kennlinie $U_a = 5 \cdot \sin(U_e)$, $0 \leq U_e \leq 1$. Bestimmen Sie U_a bei $U_e = \{0; 0{,}2; 0{,}4; 0{,}6; 0{,}8; 1\}\,V$.

a) Skizzieren Sie den Verlauf.

b) Bestimmen Sie die Ausgleichsgerade.

c) Vergleichen Sie mit einer geeigneten Berechnung in Matlab

(z. B. Funktion `polyfit`).

d) Berechnen Sie das Bestimmtheitsmaß, die statistische Standardabweichung der Steigung absolut und relativ als Maß der Unsicherheit sowie die Unsicherheit des Achsabschnitts.

e) Skizzieren Sie den Verlauf mit Ausgleichsgerade und zeichnen Sie auch die im Rahmen der berechneten Unsicherheiten maximal abweichenden Geraden ein.

2.2.7 Trendfestellung (Korrelation / Signifikanz)

Im Verlauf einer Woche wurden für die Tagesproduktion eines bestimmten Kontaktschalters jeweils folgende Mittelwerte der Schaltabstände registriert:

Messung Nr.	1	2	3	4	5	6
Schaltweg [µm]	35	32	29	27	31	30

Lässt sich aus dieser Messreihe bereits eine Aussage über einen Trend bezüglich der Veränderung der Schaltabstände dieser Produktion ableiten?

Hinweis: Ein Trend liegt vor, wenn für das Bestimmtheitsmaß $B > 0{,}5$ gilt.