Das klassische Regressionsmodell: Ein Beispiel

Kapitel 2

Ökonometrie I
Michael Hauser

Inhalt

- Ein Beispiel für das klassische, bivariate Regressionsmodell: Okun's Gesetz
- Das bivariate, lineare Regressionsmodell
- OLS, Zerlegung der Fehlerquadratsumme, Bestimmtheitsmaß R^2
- EViews - Anleitung

Beispiel: Okun’s Gesetz

Okun’s Gesetz: Ökonomische Theorie

Eine ökonomische Theorie:
Zwischen dem ’output gap’ und der ’Abweichung der aktuellen Arbeitslosenrate von der natürlichen’, $(u - u^*)$, besteht eine *positive* Beziehung: $c > 0$.

$$\frac{Y^* - Y}{Y^*} = c(u - u^*)$$

Der output gap, $(Y^* - Y)$, bezeichnet die Abweichung des aktuellen Outputs, Y, vom potenziellen, Y^*, zum Vollbeschäftigungsniveau.
u^* sei die natürliche Arbeitslosenrate.

Bem:
Zwischen der Arbeitslosenrate und dem GDP besteht eine *negative* Beziehung.

Okun's Gesetz: Deterministisches Modell

Deterministisches Modell:
Da der potenzielle Output und die natürliche Arbeitslosenrate, u^*, nicht direkt beobachtet werden können, verwenden wir *Approximationen*. Wir nehmen an

- die natürliche Arbeitslosenrate u^* ist konstant (ändert sich mit der Zeit nicht), und
- die Wachstumsrate des potenziellen Outputs ist konstant.

Dann findet man (nach weiteren technischen Approximationen) einen *negativen* Zusammenhang zwischen der

- Wachstumsrate des Outputs, ρ, und der
- Veränderung der Arbeitslosenrate, $\Delta u = u_t - u_{t-1} = DUR$.

Deterministisches Modell

Modell:
Wir *interpretieren* die Beziehung aus der Sicht der aggregierten Güternachfrage. Diese bestimmt die Arbeitslosenrate. $\beta < 0$.

$$\Delta u_t = \alpha + \beta\, \rho_t$$

Die Wachstumsrate des Outputs erklärt die Veränderung der Arbeitslosenrate.

Wir nehmen an diese Beziehung ist

- deterministisch und ändert sich nicht über die Zeit (ist stabil), und
- vollständig spezifiziert - dh es fehlen keine relevanten Erklärungsgrößen.

Bem: Ein angebotsorientierter Ökonom würde von $\rho_t = \gamma + \delta \Delta u_t$ mit $\delta < 0$ ausgehen.

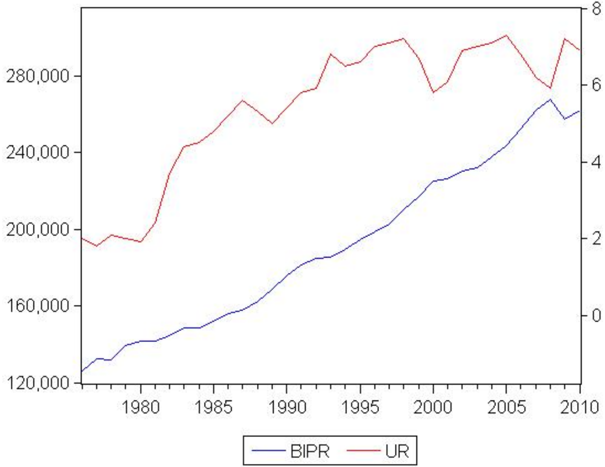
Okun's Gesetz

Da sich dieser Zusammenhang empirisch als relative stabil über die Zeit herausgestellt hat, wird er als **'stylized fact'** der Makroökonomie, als (eine Variante von) **Okun's law** bezeichnet.

Daten:

- u messen wir in % der unselbständig Beschäftigten, Jahresdurchschnitt.
- ρ als Wachstumsrate des realen *BIP* in %, jährliches Wachstum.

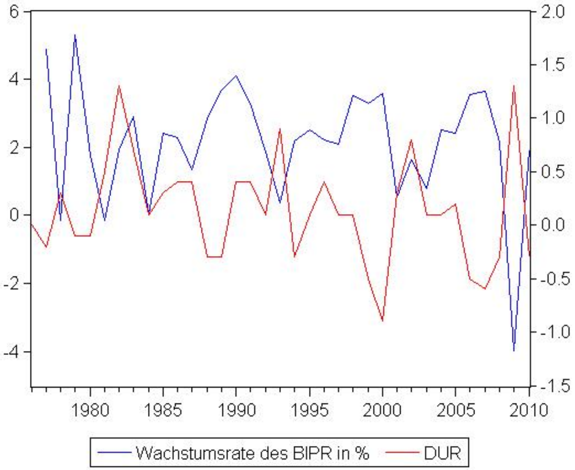
Plot der Daten *BIPR*, *UR* für 1976-2010, AT



BIPR in Mio Euro, zu konstanten Preisen 2005,
UR in % der unselbständig Beschäftigten, AT Rechnung

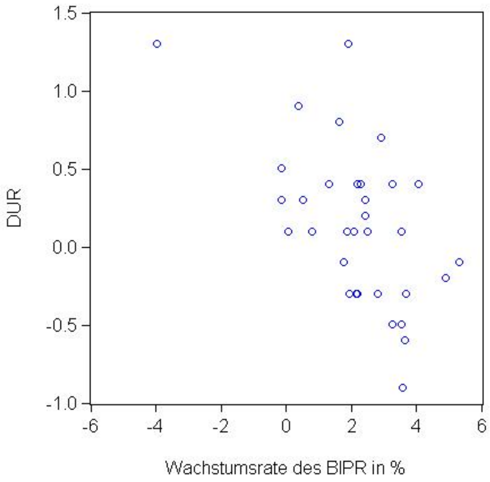
Bem: Beide Reihen steigen mit der Zeit (positiver Zusammenhang)!

Zeitreihenplot: ρ, Δu



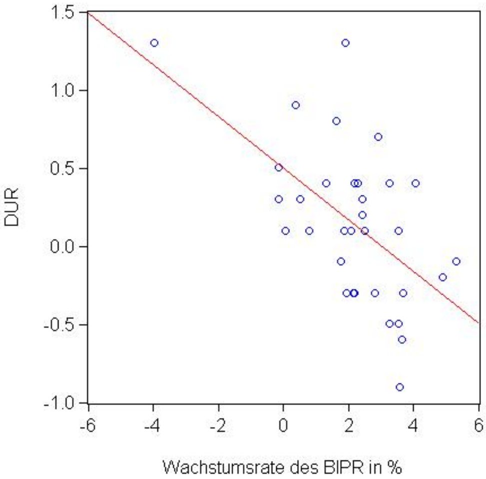
Negativer Zusammenhang. Ist Δu hoch, ist zugleich ρ niedrig, und umgekehrt.

Streudiagramm: $\rho \times \Delta u$



Negativer Zusammenhang. Steigt ρ, tendiert Δu zu fallen. Wie liegt die Regressionsgerade?

Streudiagramm: $\rho \times \Delta u$ mit Regressionsgerade



Negativer Zusammenhang. Steigt ρ, fällt Δu im Durchschnitt.

Stochastisches Modell

Stochastisches Modell:
Das Streudiagramm zeigt keine exakte lineare Beziehung zwischen Wachstumsrate und Δu, DUR. Der unterstellte Zusammenhang ist aber gut zu erkennen.
Auch scheinen keine dritten systematischen Effekte vorzuliegen, die die einfache Beziehung in Frage stellen.

Daher fügen wir dem deterministischen Modell eine stochastische Störung, ϵ, hinzu, die die Abweichungen erfassen soll.

$$\Delta u_t = \alpha + \beta\, \rho_t + \epsilon_t$$

Die stochastische Störung, ϵ

Wir verlangen von ihr, dass sie *unsystematisch* ist. Dh, sie

- hat einen *Durchschnitt von null*. Dh die Lage der Geraden soll der deterministische Teil des Modells bestimmen.
- hat eine *konstante Varianz*. Eine zunehmende Varianz würde den Einfluss dritter Variablen widerspiegeln.
- *zufällig* ist. Eine 'Nicht-Zufälligkeit' würde ebenfalls den Einfluss dritter Variablen anzeigen.
  Sie soll auch mit den Variablen auf der rechten Seite des Modells (hier ρ_t) nicht korrelieren.

Manchmal wird der Störterm auch als *Summe aller dritter Einflüsse betrachtet, die einzeln nicht relevant sind*, sodass ϵ nach dem ZGWS als unabhängig und normal verteilt mit Mittel 0 angenommen werden kann.

OLS-Schätzung

Modell:

$$\Delta u_t = \alpha + \beta\,\rho_t + \epsilon_t, \quad t = 1, \ldots, n$$

Wir berechnen nun mittels OLS die Gerade, die 'am besten' die Punktewolke approximiert.

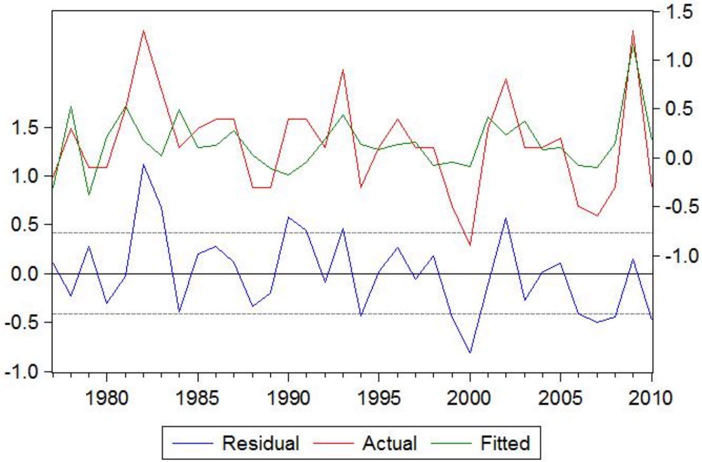
Die OLS **Schätzung** für Österreich 1977 – 2010 ergibt, $n = 34$

$$\begin{array}{rllr} \widehat{\Delta u_t} = & 0.500 - 0.165\,\rho_t & \qquad R^2 & = 0.322 \\ & (0.116)\ \ (0.042) & \qquad RSS & = 5.575 \end{array}$$

R^2 ... multiples Bestimmtheitsmaß,
RSS ... Fehlerquadratsumme, residual sum of squares,
in Klammern die Standardfehler der geschätzten Koeffizienten.

OLS ... ordinary least squares, MKQ ... Methode der kleinsten Quadrate

Fit graphisch: $\widehat{\Delta u_t} = 0.500 - 0.165\rho_t$



Die Fitted Linie soll der Actual bestmöglich reproduzieren. ZB 1982 ist eine große Abweichung, ein großes Residuum erkennbar.

Fit graphisch

- beobachtet: Δu_t
- fitted: $\widehat{\Delta u}_t$
- Residuum: $\Delta u_t - \widehat{\Delta u}_t = \widehat{\epsilon}_t$

$$\Delta u_t = \underbrace{0.500 - 0.165\rho_t}_{} + \widehat{\epsilon}_t$$

$$\Delta u_t = \widehat{\Delta u}_t + \widehat{\epsilon}_t$$

Interpretation der Schätzung

$$\widehat{\Delta u_t} = 0.500 - 0.165\,\rho_t$$

- Steigt ρ_t auf $(\rho_t + 1)$, so sinkt (im Durchschnitt) die 'Veränderung der Arbeitslosenrate' um 0.165.
  Dh: Steigt das Wirtschaftswachstum um 1 Prozentpunkt, so ist die Veränderung der Arbeitslosenrate um 0.165 (%-Punkte) kleiner als zuvor.
- Ist $\rho = 0$, so ist $\Delta u = 0.5$ pa.
  Dh: Liegt kein Wirtschaftswachstum vor, so steigt die Arbeitslosenrate (im Durchschnitt) jedes Jahr um 0.5, einen halben Prozentpunkt.

Interpretation der Schätzung

- Damit die Arbeitslosenrate konstant bleibt, ie

$$\Delta u = 0$$

$\Delta u = 0 = 0.5 - 0.165\,\rho_0$, benötigen wir ein Wachstum ρ_0 von

$$\rho_0 = 0.5/0.165 = 3.03\%\text{ pa}$$

Das bivariate, lineare Regressionsmodell

Das bivariate, lineare Regressionsmodell allgemein

Wir regressieren y auf x

$$y_t = \beta_0 + \beta_1 \, x_t + u_t$$

$y \ldots$ **abhängige** oder **zu erklärende** Variable, **Regressand**,
$x \ldots$ **unabhängige** oder **erklärende** Variable, **Regressor**,
$u \ldots$ **Fehler**, **Störterm.**

$\beta_0 \ldots$ **Interzept** oder Konstante,
$\beta_1 \ldots$ **Steigung.**

Interpretation:

- Wenn $x_t = 0$, so ist im Durchschnitt $y_t = \beta_0$.
- $dy/dx = \beta_1$ im Durchschnitt. Steigt x um 1 Einheit, so steigt y durchschnittlich um β_1.

Methode der kleinsten Quadrate, MKQ
Ordinary least squares, OLS

Methode der kleinsten Quadrate, OLS

Gegeben sind n Beobachtungen (Paare) (x_t, y_t), $t = 1, \ldots, n$. Wir suchen *Schätzwerte* für β_0 und β_1, bezeichnet mit b_0 und b_1, sodass die Anpassung des Modells (durch eine Gerade) an die Daten (Punktewolke) 'optimal' ist.

Dazu minimieren wir die **Fehlerquadratsumme**, die Summe der quadrierten Residuen.
Die **Residuen**, $\widehat{u}_t$, sind für jede beliebige Wahl von b_0 und b_1

$$y_t - (b_0 + b_1\, x_t) = y_t - \widehat{y}_t = \widehat{u}_t$$

- $\widehat{y}_t$ ist der Teil von y, der von x *erklärt* wird.
- $\widehat{u}_t$ ist der Teil von y, der von x *nicht erklärt* wird.

$$\widehat{y}_t = b_0 + b_1\, x_t \qquad \text{bzw} \qquad y_t = \widehat{y}_t + \widehat{u}_t$$

Methode der kleinsten Quadrate, OLS

Die Fehlerquadratsumme, RSS (residual sum of squares), ist

$$\sum_{t=1}^{n} \widehat{u}_t^2 = S(b_0, b_1)$$

Das Minimierungsproblem ist ein Optimierungsproblem für eine quadratische Funktion in 2 Variablen $S(b_0, b_1)$ ohne Nebenbedingungen. Variabel sind b_0 und b_1. (Die Daten sind fix gegeben.)

Dazu berechnet man die beiden partiellen Ableitungen von $S(b_0, b_1)$ bez b_0 und b_1 und setzt sie null (Bedingungen 1.Ordnung).

$$\frac{\partial S(b_0, b_1)}{\partial b_0} = 0 \quad \text{und} \quad \frac{\partial S(b_0, b_1)}{\partial b_1} = 0$$

Methode der kleinsten Quadrate, OLS

Das Ergebnis ist:

$$b_0 = \overline{y} - b_1\,\overline{x} \qquad \text{und} \qquad b_1 = \frac{\sum(x_t - \overline{x})(y_t - \overline{y})}{\sum(x_t - \overline{x})^2}$$

b_0: Die Lösung für β_0, b_0, besagt:
Nimmt man auf beiden Seiten der Gleichung das arithmetische Mittel, so müssen beide Seiten übereinstimmen,

$$\overline{y} = b_0 + b_1\overline{x} + \overline{u}$$

Die Lösung ergibt sich für $\overline{u} = 0$.

$\overline{x}, \overline{y}$ ... Mittelwert von x, bzw y

Methode der kleinsten Quadrate, OLS

b_1: (ohne Berücksichtigung von Freiheitsgraden)

$$b_1 = \frac{\mathrm{cov}(x,y)}{s_x^2} = \mathrm{corr}(x,y)\frac{s_y}{s_x}$$

Da die Standardabweichungen s_x und s_y immer positiv sind, *bestimmt die Korrelation zwischen x und y das Vorzeichen von b_1.*

Vergleichen sie dazu oben

- die Form der Punktewolke,
- die Steigung der Regressionsgeraden im Streudiagramm,
- das Vorzeichen des b_1-Koeffizienten, und
- den Korrelationskoeffizienten von ρ und Δu. ($\rightarrow$ `Descriptive Statistics`)

Normalgleichungen

Aus den Bedingungen 1.Ordnung ergibt sich für

b_0: $\frac{\partial S(b_0,b_1)}{\partial b_0} = 0$: $\sum 2(y_t - b_0 - b_1 x_t)(-1) = 0$,

b_1: $\frac{\partial S(b_0,b_1)}{\partial b_1} = 0$: $\sum 2(y_t - b_0 - b_1 x_t)(-x_t) = 0$,

und damit die Normalgleichungen

$$\sum \widehat{u}_t = 0$$

$$\sum \widehat{u}_t x_t = 0$$

Die OLS-Lösung hat die Eigenschaft, dass

- *das Mittel von* $\widehat{u}$ - wie gewünscht - *null* ist, und
- x und $\widehat{u}$ *orthogonal (unkorreliert)* sind.

Normalgleichungen

Orthogonal bedeutet, dass

- die Korrelation zwischen $\widehat{u}$ und x null ist, und
- die Vektoren $\boldsymbol{x}$ und $\widehat{\boldsymbol{u}}$ - geometrisch interpretiert - in einem rechten Winkel aufeinander stehen.

Zerlegung der TSS, total sum of squares, S_{yy}

Statt Varianzen verwenden wir Summen, bzw (Abweichungs-)Quadratsummen, sum of squares, SS, in den Darstellungen. Freiheitsgrade werden erst bei Tests berücksichtigt.

$$S_{yy} = \sum(y_t - \overline{y})^2, \quad S_{xx} = \sum(x_t - \overline{x})^2, \quad S_{xy} = \sum(x_t - \overline{x})(y_t - \overline{y})$$

$$b_1 = \frac{S_{xy}}{S_{xx}}$$

Unser geschätztes Modell (hat einen Interzept und) lautet $y_t = b_0 + b_1 x_t + \widehat{u}_t$, bzw

$$y_t = \widehat{y}_t + \widehat{u}_t$$

Zerlegung der TSS, total sum of squares, S_{yy}

$$y_t = \widehat{y}_t + \widehat{u}_t$$

Für die zugehörigen sum of squares,

- **TSS**, *total* SS (für y),
- **ESS**, *erklärte* SS (für $\widehat{y}$) und
- **RSS**, *residual* SS (für $\widehat{u}$) gilt

$$TSS = ESS + RSS$$

Die SS von y, TSS, wird *zerlegt in* die durch das Modell erklärte SS, ESS, *und* der SS der Residuen, RSS, die nicht erklärt wird.

Dies gilt, weil x und $\widehat{u}$ orthogonal bzw unkorreliert sind.

Analog zu: $V(X+Y) = V(X) + V(Y)$ wenn $\mathrm{Corr}(X, Y) = 0$.

Aus x und u unkorr, folgt $(b_0 + b_1 x)$ und u unkorr, und damit $\hat{y}$ und u unkorr.

Das Bestimmtheitsmaß, R^2

Das Bestimmtheitsmaß gibt den Anteil der erklärten Varianz (SS) an der Gesamtvarianz (SS) von y an.

$$R^2 = \frac{ESS}{TSS} = \frac{TSS - RSS}{TSS} = 1 - \frac{RSS}{TSS}$$

Oder: R^2 ist 1 minus dem Anteil der nicht-erklärten SS.

$$0 \leq R^2 \leq 1$$

- $R^2 = 0$: x kann gar nichts von der Varianz von y erklären.
- $R^2 = 1$: Es besteht eine exakter linearer Zusammenhang zwischen x und y. Alle Residuen sind null!
- $R^2 = r^2_{y\widehat{y}}$
  Nur im bivariaten Modell ist das Bestimmtheitsmaß gleich dem Quadrat von r_{xy}, $R^2 = r^2_{xy}$.

Bsp: Okun's law, EViews Output

Dependent Variable: UR-UR(-1)
Method: Least Squares
Date: 03/13/11 Time: 22:17
Sample (adjusted): 1977 2010
Included observations: 34 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	0.500758	0.116154	4.311149	0.0001
RHO	-0.165253	0.042385	-3.898842	0.0005

R-squared	0.322048	Mean dependent var	0.144118
Adjusted R-squared	0.300862	S.D. dependent var	0.499206
S.E. of regression	0.417409	Akaike info criterion	1.147520
Sum squared resid	5.575359	Schwarz criterion	1.237306
Log likelihood	-17.50785	Hannan-Quinn criter.	1.178140
F-statistic	15.20097	Durbin-Watson stat	1.748845
Prob(F-statistic)	0.000465		

Umsetzung in EViews

OLS Schätzung

Im Workfile zuerst die Variablen DUR und RHO erzeugen:

`Generate DUR = u - u(-1)` und

`Generate rho = log(BIPR)-log(BIPR(-1))`

OLS Schätzung im Workfile

→ `Quick` → `Estimate equation`: Modell eingeben

```
dur c rho
```

`Enter`.

`c` steht für die Konstante, Interzept.

`Coefficient` ... die geschätzten Paramter

`R-squared` ... R^2

`sum of squared res` ... *RSS*

Abspeichern des geschätzten Modells durch → `Name`.

Graphiken

Graphischer Vergleich von Beobachtungen und Erklärung durch das Modell:
Im Estimation output (des Objekts equation)
$\rightarrow$ `View` $\rightarrow$ `Actual, Fitted, Residual` $\rightarrow$ `Actual, Fitted, Residual Graph`.

Graphische Darstellung von 2 oder mehreren Reihen:
Im Workfile: Auswahl der Reihen durch Str und li Maustaste. Rechte Maustaste: $\rightarrow$ `Open` $\rightarrow$ `as Group` $\rightarrow$ `View` $\rightarrow$ `Graph`:
Zeitreihenplot: $\rightarrow$ `Line&Symbol (Single graph)`
Streudiagramm: $\rightarrow$ `Scatter` (evt mit `Fit lines` - Option als `Regression line`).

Univariate und multivariate deskriptive Statistiken

Univariate deskriptive Statistiken:
Reihe anklicken $\rightarrow$ `View` $\rightarrow$ `Descriptive ...` $\rightarrow$ `Histogram ...`

Multivariate deskriptive Statistiken:
Daten anklicken $\rightarrow$ `Open` $\rightarrow$ `as Group` $\rightarrow$

- alle univarite Statistiken:
  `Descriptive ...` $\rightarrow$ `Common Sample`
- Kovarianzen, Korrelationen, etc:
  `Covariance Analysis`

Referenzen

Referenzen

Hackl, 2.1-2, 5.3.1

Wooldridge, Chapter 2.1-2.3, App 2A

Übungen

Übungen

1. Wählen sie das Datenfile `okuns_law.wf1` bzw. `okuns_law_R.txt`. Reproduzieren sie alle transformierte Variable, Graphiken und die Schätzung auf den Folien.

Übungen

2 I Wählen sie (zufällig) je 5 Werte für die Variablen X und Y, $x_1, \ldots, x_5, y_1, \ldots, y_5$. Wir interpretieren die Werte als Stichprobe im Umfang $n = 5$ für das Paar (X, Y), $(x_1, y_1), \ldots, (x_5, y_5)$.
(a) Berechnen sie das (arithmetische) Mittel/Mittelwert für X und Y.
(b) Berechnen sie die Stichprobenvarianzen zu X und Y.
(c) Berechnen sie die Stichproben-Kovarianz und den Korrelationskoeffizienten.
(d) Berechnen sie den Standardfehler zu $\overline{x}$.
(e) Geben sie ein Konfidenzintervall zu $\overline{x}$ an.
(f) Testen sie die Hypothese $\mu_X = 2$.

Übungen

3 I (a) Skizzieren sie zu Bsp 2f die Testverteilung.
(b) Markieren sie darin die Realisation der Teststatistik.
(c) Zeichnen sie den Kritischen Bereich (Ablehnungsbereich) ein.
(d) Schraffieren sie die zum p-Wert gehörige Fläche und schätzen sie dessen Wert ab.

Übungen

4 Zeichnen sie je eine Punktewolke (Streudiagramm) mit ca. 20 Punkten. Eine die
(a) sich gut durch einen linearen Funktion beschreiben lässt.
(b) einen negativen Zusammenhang zwischen den beiden Variablen anzeigt.
(c) sich nur schlecht durch eine lineare Funktion beschreiben lässt.
Versuchen sie die Korrelationskoeffizienten in (a)–(c) anzugeben, ohne sie zu berechnen.

Übungen

5 I (a) Zeichnen sie in die Streudiagramme aus (4a)–(4c) Geraden ein, die die Wolken gut beschreiben.
(b) Interpretieren sie die Geraden als lineare Funktion $y = a + b\,x + u$.
u enthält die Ungenauigkeit der linearen Approximation.
Wählen sie je 2 Punkte mit kleinen und großen u-Werten aus.
(c) Wählen sie Skalen auf beiden Achsen und versuchen sie die Steigung b und den Interzept a abzulesen.