

Das multiple Regressionsmodell

Kapitel 3

Ökonometrie I
Michael Hauser

Inhalt

- ▶ Das multiple Regressionsmodell
- ▶ Lineare Regression in Matrixnotation
- ▶ Beispiele: 2 Konsumfunktionen für Ö
- ▶ Annahmen des linearen Regressionsmodells
- ▶ OLS Lösung in Matrixnotation
- ▶ Quadratsummenzerlegung, R^2
- ▶ Erwartung und Varianz des OLS Schätzers (Unverzerrtheit)

Das multiple Regressionsmodell

Das multiple Regressionsmodell

Wir formulieren allgemeiner

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + u$$

u ist *unkorreliert* und $E(u) = 0$, $V(u) = \sigma^2$

y ... abhängige, zu erklärende, erklärte Variable, Regressand,

x ... unabhängige, erklärende Variable, Regressor,

u ... (unbeobachtbarer) Fehler, Störterm (ZV)

Andere Bezeichnungen für y , x sind:

y	x
Response- vorhergesagte	Stimulus, Kontroll-Variable Prädiktor-Variable

Das multiple Regressionsmodell

Da u eine Zufallsvariable ist, ist auch y als Funktion von u eine ZV.

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + u$$

Die x -Variablen können

- ▶ fix, dh fix gegebene Zahlen, oder
- ▶ stochastische, dh Zufallsvariable, bzw Realisationen von ZVn, sein. Wir weisen gesondert darauf hin, wenn wir sie als ZVn ansehen.

Das multiple Regressionsmodell: Anforderung

Damit das Modell sinnvoll ist, verlangen wir zusätzlich, dass auf der rechten Seite keine der x -Variablen doppelt auftritt.

Wären zB x_1 und x_2 identisch, könnte man zwischen dem Effekt von x_1 und x_2 (trivialerweise) nicht unterscheiden.

Allgemeiner verlangen wir, dass keine x -Variable als (exakte) Linearkombination der anderen x -Variablen darstellbar ist. [keine (exakte) Multikollinearität]

Jedes x enthält eine eigene Information, die in den anderen x nicht enthalten ist.

Das multiple Regressionsmodell

- ▶ β_0 ... Interzept oder Konstante

$$E(y|x_1 = \dots = x_k = 0) = \beta_0$$

Wenn alle $x_i = 0$ sind, so ist y im Durchschnitt β_0 .

- ▶ β_j ... Steigungen, $j = 1, \dots, k$

$$[x_j \rightarrow (x_j + 1)] \text{ cp} \quad \longrightarrow \quad [y \rightarrow (y + \beta_j)]$$

Eine Veränderung von x um Δx_j führt cp zu einer Veränderung in y um $\Delta y = \beta_j \Delta x_j$.

$$\Delta x_j \text{ cp} \quad \longrightarrow \quad \Delta y = \beta_j \Delta x_j$$

$$\partial y / \partial x_j = \beta_j$$

Die β -Koeffizienten sind Steigungen, die *ceteris paribus* (cp), dh bei Konstanthalten der anderen x Variablen, gelten.

Bsp: Konsumfunktion für Ö, 1977-2009

$$C_t = \beta_0 + \beta_1 Y_t^d + \beta_2 W_{t-1} + u_t, \quad t = 1, \dots, n$$

C ... privater Konsum, real

Y^d ... disponibles Einkommen, real

W_{t-1} ... Vermögen der Vorperiode, real

Die OLS Schätzung für Österreich 1977 – 2009 ($n = 34$) ergibt

$$\begin{array}{lll} \hat{C}_t = 7039.9 + 0.596 Y_t^d + 0.064 W_{t-1} & R^2 = 0.992 \\ (7040) \quad (0.098) \quad (0.020) & RSS = 69.4 \cdot 10^6 \end{array}$$

C ... privater Konsum, real, Mio €

Y^d ... disponibles Einkommen, real, Mio €

W ... Geldvermögen in Mio €, real

real ... zu Preisen 2005. (dats_01.wf1)

Bsp: Konsumfunktion für Ö, 1977-2009

- ▶ R^2 ist sehr hoch. Das deutet darauf hin, dass ein Trend (alles wächst), der sowohl die Konsum-, wie auch die Einkommens- und Vermögensentwicklung bestimmt, den Zusammenhang überlagert.
- ▶ $b_0 = 7039.9$ ist der autonome Konsum, wenn das Einkommen und das Vermögen null ist. [b_0 ist nicht signifikant von null verschieden.]
- ▶ $b_1 = 0.596$: Steigt das laufende Eink (zu konstanten Preisen gemessen) cp um €1 Mio, so erhöht sich der Konsum (real) um €0.596 Mio.
- ▶ $b_2 = 0.064$: Steigt das Vermögen (real) um 1 Mio 2005-€, dann erhöht sich der Konsum um 0.064 Mio.

Bsp: Konsumfunktion für Ö, 1977-2009

Ein dynamisches Modell ist

$$\widehat{C}_t = 432.2 + 0.331 Y_t^d + 0.640 C_{t-1} \quad R^2 = 0.997$$

(1266) (0.082) (0.089) $RSS = 43.2 \cdot 10^6$

C_{t-1} ... Konsum der Vorjahres

Eine Veränderung in Y^d in Periode t hat nicht nur auf den Konsum der aktuellen Periode einen Einfluss, sondern auch auf C_{t+1} , etc.

Das stochastische Modell

Wir unterscheiden zwischen einem

- ▶ theoretischen Modell:

Hier wird ein Zusammenhang zwischen theoretischen Größen postuliert, der für alle Zeitpunkte t bzw Beobachtungseinheiten i gilt. (Anforderung von *intrinsic and extrinsic closure*.)

- ▶ stochastischen Modell:

Es ist das theoretische Modell in quantifizierter Form. Es wird als wahr bzw korrekt spezifiziert angenommen. Es beschreibt die Struktur in der Grundgesamtheit aus der gezogen wird. Allerdings sind formale Modelle nur relativ 'einfache' Varianten der Theorie.

Das stochastische Modell

► empirischen Modell:

Das ist das Modell, das an die vorliegenden Daten angepasst wurde. Für die Parameter β_0, \dots, β_k etc haben wir mittels OLS Stichprobenwerte b_0, b_1, \dots, b_k ermittelt.

$$y = b_0 + b_1 x_1 + \dots + b_k x_k + \hat{u}$$

Unsere Aufgabe ist uA zu überprüfen, ob die empirische Variante den Anforderungen des stochastischen bzw des theoretischen Modells genügt.

Das Modell in Matrixnotation

Das Modell in Matrixnotation

Die Matrixnotation hilft komplexe Zusammenhänge kompakt anzugeben.

$$y_t = \beta_0 + \beta_1 x_{t,1} + \dots + \beta_k x_{t,k} + u_t$$

Wir schreiben das Modell für jede Beobachtungseinheit separat an.

$$y_1 = \beta_0 + \beta_1 x_{1,1} + \dots + \beta_k x_{1,k} + u_1$$

$$\vdots \quad \quad \quad \vdots$$

$$y_n = \beta_0 + \beta_1 x_{n,1} + \dots + \beta_k x_{n,k} + u_n$$

Das entspricht

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix} \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

Das Modell in Matrixnotation

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{bmatrix} 1 & x_{11} & \dots & x_{1k} \\ \vdots & \vdots & \vdots & \vdots \\ 1 & x_{n1} & \dots & x_{nk} \end{bmatrix} \begin{pmatrix} \beta_0 \\ \vdots \\ \beta_k \end{pmatrix} + \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix}$$

oder

$$\mathbf{y}_{n \times 1} = \mathbf{X}_{n \times (k+1)} \boldsymbol{\beta}_{(k+1) \times 1} + \mathbf{u}_{n \times 1}$$

Unsere Vektoren sind alle Spaltenvektoren.

Kurz:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

Das Modell in Matrixnotation

- ▶ Die abhängige Variable y wird n -mal beobachtet. Diese Werte werden in einem Spaltenvektor der Länge n zusammengefasst: $\mathbf{y} = \mathbf{y}_{n \times 1}$
- ▶ Die unabhängigen Variablen x_j bilden k Spalten der Matrix \mathbf{X} .
- ▶ Tritt im Modell eine Konstante auf, so hat die \mathbf{X} -Matrix als erste Spalte eine Spalte mit 1-en. Daher hat $\mathbf{X}_{n \times (k+1)}$ die Dimension $n \times (k + 1)$.
- ▶ Die Parameter β_0, \dots, β_k bilden einen Spaltenvektor der Länge $(k + 1)$. Er steht auf Grund der Definition der Matrixmultiplikation *rechts* von \mathbf{X} .

Exkurs: Matrixmultiplikation

Multiplizieren wir einen Zeilenvektor mit einem Spaltenvektor gleicher Länge, so erhalten wir eine reelle Zahl.

$$\mathbf{u}'\mathbf{u} = [\mathbf{u}']_{1 \times n} \mathbf{u}_{n \times 1} = (u_1 \dots u_n) \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} = \sum u_j^2 \in \mathbb{R}$$

Multiplizieren wir einen Spaltenvektor der Länge n mit einem Zeilenvektor gleicher Länge, erhalten wir eine symmetrische $(n \times n)$ -Matrix.

$$\mathbf{u}_{n \times 1} [\mathbf{u}']_{1 \times n} = \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} (u_1 \dots u_n) = \begin{bmatrix} u_1 u_1 & \dots & u_1 u_n \\ \vdots & \ddots & \vdots \\ u_n u_1 & \dots & u_n u_n \end{bmatrix}$$

Exkurs: Matrixmultiplikation

Zwei Vektoren \mathbf{x} und \mathbf{u} gleicher Länge sind *orthogonal*, wenn

$$\mathbf{x}'\mathbf{u} = \sum_{i=1}^n x_i u_i = 0$$

Annahmen des Regressionsmodells

Modellannahmen für Querschnittsdaten

Fassen wir die notwendigen Modellanforderungen, die wir intuitiv getroffen haben, zusammen: [Woo 3]

MLR.1 Lineares Modell

Wir kennen das wahre Modell und es ist linear in den Parametern.

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + u$$

Alle relevanten Variablen sind im Modell, keine ist zuviel.

MLR.2 Random Sampling (von Querschnittsdaten)

Wir führen ein Experiment durch und ziehen wiederholt und zufällig Vektoren der Länge $(k + 1)$

$$(y_i, x_{i,1}, \dots, x_{i,k}) \quad \text{mit} \quad i = 1, \dots, n$$

aus der Grundgesamtheit. Dh die Fehler verschiedener Ziehungen, u_i und u_ℓ , sind unabhängig, $i \neq \ell$.

Modellannahmen für Querschnittsdaten

MLR.3 Keine perfekte Kollinearität

Wir wollen, dass jede x Variable zumindest zT einen eigenen Beitrag liefert. Das ist gewährleistet, wenn

$$\text{Rank}(\mathbf{X}) = k + 1$$

Der Rang von \mathbf{X} ist $k + 1$.

MLR.4 Bedingtes Mittel von u ist null. $E(u|x_1, \dots, x_k) = 0$.

Diese Bedingung erfordert mehr als $E(u) = 0$. Sie besagt, dass x 's und u im theoretischen Modell unkorreliert sind. [x 's sind exogen. Das Modell ist abgeschlossen.]

MLR.5 Homoskedastizität

Die Varianz von u ist konstant.

Kommentar zu den Annahmen

MLR.2 Hier wird genauso zufällig gezogen, wie in der LV Einführung in die Statistik.

MLR.3 $\mathbf{X}_{n \times (k+1)}$ hat maximalen Rang, ie $(k + 1)$. Wir setzen natürlich voraus, dass $(k + 1) \ll n$. Es gibt deutlich mehr Beobachtungen als erklärende Variable.

Die Matrix $(\mathbf{X}'\mathbf{X})_{(k+1) \times (k+1)}$ hat ebenfalls Rang $(k + 1)$, also vollen Rang, und ist invertierbar. Damit existiert eine eindeutige Lösung für des Minimierungsproblem.

MLR.4 Ist diese Ann verletzt, spricht man von einem Endogenitätsproblem.

Annahmen im Regressionsmodell: Zusammenfassung

Man schreibt oft:

- ▶ Das Modell ist linear (in den Parametern)
- ▶ u unkorreliert, $E(u) = 0$, $V(u) = \sigma^2$

und meint alle Annahmen MLR.1 - MLR.5.

Annahmen für Zeitreihen, skip

Annahmen für Zeitreihen

ad Zeitreihen-Daten:

Die modifizierten Annahmen für Zeitreihen werden ebenfalls angegeben, damit wir auch mit Zeitreihen arbeiten dürfen. Sie werden aber erst in Ökonometrie 2 näher erläutert.

Der Unterschied zu Querschnittsdaten besteht einerseits darin, dass für ZR-Daten eine natürliche Reihenfolge besteht, auf die oben nicht eingegangen wird. Weiters muss man iA für ZR-Daten ein theoretisches Experiment konstruieren, da in der Praxis wiederholte Ziehungen nicht möglich sind.

Annahmen für Zeitreihen

Für Zeitreihen ist iA ein Experiment nur beschränkt möglich. Daher wird Ann 2 modifiziert in TS3 und TS5. [Woo 10]

TS.1 Lineares Modell

Wir kennen das wahre Modell und es ist linear in den Parametern.

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + u$$

Alle relevanten Variablen sind im Modell, keine ist zuviel. [TS.1 = MLR.1]

TS2 Keine perfekte Kollinearität

Wie oben erwähnt wollen wir, dass jede Variable einen eigenen Beitrag liefert, sonst können wir zwischen den Variablen nicht unterscheiden. [TS.2 = MLR.3]

Annahmen von OLS für Zeitreihen

TS.3 $E(u_t | \mathbf{X}) = 0$

u_t ist mit allen x -Werten, sowohl aus Periode t , wie auch aus den anderen Perioden $t \pm \ell$ unkorreliert.

Man sagt: Alle x -Variablen sind strikt exogen, $\ell \neq 0$.

[TS.3 \approx MLR.4]

TS.3' Eine schwächere Variante verlangt nur die kontemporäre Unkorreliertheit.

$E(u_t | \mathbf{x}_t) = 0$. Die x sind kontemporär exogen.

TS.4 Homoskedastizität

Die Varianz von u ist konstant. [TS.4 = MLR.5]

TS.5 Zeitliche Unkorreliertheit von u . $\text{Corr}(u_t, u_s) = 0$ für $t \neq s$. [TS.5 \approx MLR.2]

OLS Lösung in Matrixnotation

OLS Lösung in Matrixnotation

Das Modell ist

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$$

OLS minimiert die Fehlerquadratsumme.

Der Fehler $\hat{\mathbf{u}} = \hat{\mathbf{u}}(\mathbf{b})$ ist für gegebenes $\mathbf{b} = (b_0, \dots, b_k)'$

$$\hat{\mathbf{u}} = \mathbf{y} - \mathbf{X}\mathbf{b}$$

Wir suchen das \mathbf{b} , das die Fehlerquadratsumme minimiert

$$\sum \hat{u}_i^2 = \hat{\mathbf{u}}'\hat{\mathbf{u}} = (\mathbf{y} - \mathbf{X}\mathbf{b})'(\mathbf{y} - \mathbf{X}\mathbf{b}) \rightarrow \min$$

Die Lösung lässt sich kompakt schreiben als

$$\mathbf{b}_{LS} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

Zum Vergleich: Die Lösung für b_1 im bivariaten Modell

$$b_1 = \text{cov}(x, y) / \text{var}(x) = [1 / \text{var}(x)] \text{cov}(x, y)$$

OLS Lösung in Matrixnotation

Bem: Der Einschluss des Interzepts in das Modell bewirkt, dass y , wie auch alle x so interpretiert werden können, als ob sie mittelwert-bereinigt wären. (Vgl Frisch-Waugh)

Als Schätzer für die Fehlervarianz $\hat{\sigma}^2$ wird

$$\hat{\sigma}^2 = \frac{1}{n - (k + 1)} \sum \hat{u}_i^2$$

verwendet, die um Freiheitsgrade korrigierte Varianz. $\hat{\sigma}$ ist der **Standardfehler der Regression**.

Im bivariaten Modell (mit Interzept) ist die Anzahl der Freiheitsgrade für die Varianz der Residuen $n - 2$.

Beispiel: Lösung für $k = 2$

$$\mathbf{b} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$$

Im Detail für $k = 2$

$$\begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix} = \begin{bmatrix} n & \sum x_1 & \sum x_2 \\ \sum x_1 & \sum x_1^2 & \sum x_1 x_2 \\ \sum x_2 & \sum x_1 x_2 & \sum x_2^2 \end{bmatrix}^{-1} \begin{pmatrix} \sum y \\ \sum x_1 y \\ \sum x_2 y \end{pmatrix}$$

Quadratsummenzerlegung

Das geschätzte Modell lautet

$$\mathbf{y} = \mathbf{X}\mathbf{b} + \hat{\mathbf{u}}$$

Mit $\hat{\mathbf{y}} = \mathbf{X}\mathbf{b}$

$$\mathbf{y} = \hat{\mathbf{y}} + \hat{\mathbf{u}}$$

Auch hier sind die x orthogonal zu u , daher gilt die Quadratsummenzerlegung

$$\mathbf{y}'\mathbf{y} = \hat{\mathbf{y}}'\hat{\mathbf{y}} + \hat{\mathbf{u}}'\hat{\mathbf{u}}$$

Die 'Kovarianzterme' sind null. Dh

$$TSS = ESS + RSS$$

Es gilt: Alle \mathbf{x}_j , Spalten von \mathbf{X} , sind orthogonal zu $\hat{\mathbf{u}}$.

$$\mathbf{X}'\hat{\mathbf{u}} = \mathbf{0}_{(k+1) \times 1}$$

Bestimmtheitsmaß

Das Bestimmtheitsmaß ist einfach

$$R^2 = \frac{\hat{\mathbf{y}}' \hat{\mathbf{y}}}{\mathbf{y}' \mathbf{y}} = \frac{ESS}{TSS}$$

- ▶ Das Bestimmtheitsmaß ist das multiple Analogon zum bivariat definierten quadrierten Korrelationskoeffizienten.
- ▶ Es ist die quadrierte Korrelation zwischen y und \hat{y}

$$R^2 = \text{corr}(y, \hat{y})^2$$

\hat{y} ist die beste Linearkombination der x -Variablen (die y im OLS-Sinn approximiert).

- ▶ OLS sucht die Linearkombination der x , die R^2 maximiert.

Eigenschaften des OLS Schätzers

Erwartung und Varianz des OLS Schätzers

Gegeben ist das Modell

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + u$$

mit u unkorreliert, $E(u) = 0$, $V(u) = \sigma^2$

- ▶ Der Erwartungswert von b_j bzw \mathbf{b} :

$$E(b_j) = \beta_j \quad \text{bzw} \quad E(\mathbf{b}) = \boldsymbol{\beta}$$

Der Schätzer ist *unverzerrt* (auch erwartungstreu).

- ▶ Die Varianz-Kovarianzmatrix von \mathbf{b} gegeben \mathbf{X} :

$$V(\mathbf{b}|\mathbf{X}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}$$

- ▶ Die Varianz von b_j ist das zugehörige Diagonalelement:

$$V(b_j) = \sigma^2[(\mathbf{X}'\mathbf{X})^{-1}]_{jj}$$

Varianz-Kovarianzmatrix der Koeffizienten b_j

- ▶ Die Standardabweichung $\sqrt{\sigma^2[\mathbf{X}'\mathbf{X}]_{jj}^{-1}}$ in der Stichprobe

$$se(b_j) = \sqrt{\hat{\sigma}^2[\mathbf{X}'\mathbf{X}]_{jj}^{-1}}$$

heißt **Standardfehler** von b_j .

Beispiel: $k=2$

$$\text{var}(\mathbf{b}) = \begin{bmatrix} \text{var}(b_0) & \text{cov}(b_0, b_1) & \text{cov}(b_0, b_2) \\ \text{cov}(b_0, b_1) & \text{var}(b_1) & \text{cov}(b_1, b_2) \\ \text{cov}(b_0, b_2) & \text{cov}(b_1, b_2) & \text{var}(b_2) \end{bmatrix}$$

Übungen und Referenzen

Übungen

- 1 | Verwenden sie ihre selbst zusammengestellten Daten. \mathbf{X} ist die (5×2) -Matrix, die aus dem Vektor aus Einsen und dem x -Vektor besteht. \mathbf{y} die abhängige Variable.

Schätzen sie mit OLS die Parameter des Modells

$$y_i = \beta_0 + \beta_1 x_i + u_i$$

Berechnen sie dazu

- (a) $\mathbf{X}'\mathbf{X}$ (b) $(\mathbf{X}'\mathbf{X})^{-1}$
(c) $\mathbf{X}'\mathbf{y}$ (d) $\mathbf{b}_{OLS} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{y}$

$$\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_5 \end{pmatrix} \quad \mathbf{X} = \begin{pmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_5 \end{pmatrix}$$

Übungen

Hinweis: Inverse einer 2×2 -Matrix

Sei

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

dann ist die Inverse dazu

$$\mathbf{A}^{-1} = \frac{1}{(ad - bc)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

$$|A| = \det(A) = ad - bc.$$

Sie können zur Kontrolle ihre Daten in Eviews oder EXCEL eingeben und die Regression berechnen.

- 2 Überlegen sie sich, was passiert, wenn gleich viele Beobachtungen wie Variablen vorliegen. $n = k + 1$
Wie lautet die Lösung im Fall $n < k + 1$?

Übungen

3 Berechnen sie zu Bsp 1 die modellierte Reihe \hat{y} , und die Residuen \hat{u} .

(a) Berechnen sie \hat{y} .

$$\hat{y} = Xb$$

(b) Berechnen sie \hat{u} .

$$\hat{u} = y - Xb = y - \hat{y}$$

(c) Zeigen sie und formulieren sie in Worten

$$y = \hat{y} + \hat{u}$$

(d) Berechnen sie die Varianz der Residuen, $\hat{\sigma}^2$. ($n = 5, k = 1$)

$$\hat{\sigma}^2 = [1/(n - k - 1)]\hat{u}'\hat{u}$$

Zusatzbeispiel:

4 Berechnen sie zu Bsp 1 die Varianz-Kovarianz-Matrix für \mathbf{b} , $\text{var}(\mathbf{b})$.

(a) Berechnen sie

$$\hat{\sigma}^2[\mathbf{X}'\mathbf{X}]^{-1}$$

(b) Berechnen sie

$$\text{se}(b_1) = \sqrt{\hat{\sigma}^2[\mathbf{X}'\mathbf{X}]_{22}^{-1}}$$

- 5 Vergleichen sie unser Modell und unsere Schätzung für Okun's Law mit der Graphik auf https://en.wikipedia.org/wiki/Okun's_law.
Kommentieren sie die Unterschiede bez Land, Beobachtungsperiode, Anzahl der Beobachtungen, un-/abhängige Variable, Koeffizienten, R^2 , Interpretation, etc. Schätzen sie für Ö die funktionale Form, die für die USA gewählt wurde.
- Bem:** Wenn sie $y = \beta_0 + \beta_1 x + u$ schätzen, erhalten sie für
 $b_1 = \text{corr}(x, y) s_y / s_x$.
- Wenn sie *hingegen* $x = \alpha_0 + \alpha_1 y + v$ schätzen, erhalten sie für
 $a_1 = \text{corr}(x, y) s_x / s_y$.
- $a_1 \neq b_1$ i.A.!
- Verwenden sie dazu `okuns_law_K2Ue1_R.txt`, oder den letzten Teil von `okuns_law_K4T1Ue2_R.txt`.

- 6 Wählen sie das Datenfile `twoyear.wf1` bzw. `twoyear_R.txt`. Darin befinden sich uA die Variablen

`lwage` ... $\log(\text{wage})$, `jc` ... Jahre in einem 2-year College,

`univ` ... Jahre in einem 4-year College,

`exper` ... Berufserfahrung in Monaten

Sie schätzen

$$\log(\text{wage}) = \beta_0 + \beta_1 jc + \beta_2 univ + \beta_3 exper + u$$

Interpretieren sie die Parameter, etc.

- 7 Schätzen sie eine der beiden Konsumfunktionen von den Folien für eine andere Periode und vergleichen sie die Parameter, etc. Das Datenfile für die Folien ist `dat_01.wf1`. Verwenden sie aber das aktuellere `dat_01_22.wf1` bzw. `dat_01_22_R.txt` mit etwas anderen Variablenbezeichnungen.

Für `dat_01.wf1`:

```
Quick...PCR C PYR (WLN(-1)*100/PCD(-1))
```

bzw `Quick...PCR C PYR PCR(-1)`

Für `dat_01_22.wf1`:

```
Quick...CPR C YHR (WLN(-1)*100/PCD(-1))
```

bzw `Quick...CPR C YHR CPR(-1)`

mit `Smpl 1977 2007` und `Smpl 1977 2022`

Referenzen

Hackl 3 , 4

Wooldridge 3

EViews

EViews Befehle

Varianz-Kovarianz-Matrix der geschätzten Koeffizienten:

Sie haben eben ein Modell mit OLS (`Quick, ...`) geschätzt, oder haben eine `equation` geöffnet:

`View` → `Covariance`

Okun's Modell kann ohne vorhergehende explizite Variablentransformation einfach geschätzt werden als:

`Quick` → `d(ur)` `c` `dlog(bipr)`