Funktionale Formen

Kapitel 7

Ökonometrie I

Michael Hauser

Inhalt

- Änderung von Skalen, Messeinheiten
- Aggregation über die Zeit
- Univariate, bivariate, multiple, simultane Regressionsmodelle, VARs (vector autoregressions)
- Lineares vs nicht-lineares Modell
- Deterministische Trends: linear, quadratisch, exponentiell; geometrisches Wachstum
- Stochastische Trends (random walk ohne/mit Drift, Differenzenbildung) weggelassen
- Funktionale Formen: loglog, semilog
- Dummy Variable

Änderung von Skalen, Messeinheiten

Skalen, Messeinheiten

Betrachten wir eine Konsumgleichung für Jahresdaten

$$C = \beta_0 + \beta_1 Y^d + \beta_2 r + u$$

$C \ldots$ priv Konsum in Mrd € zu konst Preisen,
$Y^d \ldots$ dispon Eink in Mrd € zu konst Preisen,
$r \ldots$ realer Zinssatz in %.

Für Ö 2005 (zu konstanten Preisen 2005) verwenden wir folgende Zahlenwerte

Y^d in Mrd €	C in Mrd €	r in %
202.72	131.43	2.0

Die *Messeinheit von* β_1 (Konsumneigung) ist:
(Mrd € Konsum) pro (1 Mrd € Einkommen) = Konsum / Eink

Änderung von Skalen, Messeinheiten: Bsp 1

Angenommen wir messen nun das disponible Einkommen in Mio € und die Zinsen in (% / 100).

$[Y^d \cdot 1000]$ in Mio €	C in Mrd €	$[r/100]$
202720	131.43	0.02

Wir schätzen nun mit den neu skalierten Variablen

$$C = \gamma_0 + \gamma_1 [Y^d \cdot 1000] + \gamma_2 [r/100] + u$$

Beide Modellvarianten liefern die gleiche Anpassung.

$$\gamma_0 = \beta_0, \qquad \gamma_1 = \beta_1/1000, \qquad \gamma_2 = \beta_2 \cdot 100$$

Bzw

$$C = \beta_0 + (\beta_1/1000)[Y^d \cdot 1000] + (\beta_2 \cdot 100)[r/100] + u$$

Änderung von Skalen, Messeinheiten: Bsp 1

- Die Konstante bleibt unverändert,
- γ_1 ist um den Faktor (1/1000) kleiner als β_1, da die Werte von $[Y^d \cdot 1000]$ um den Faktor 1000 größer sind als Y^d.
- Bei γ_2 und β_2 ist es umgekehrt. γ_2 ist um den Faktor 100 größer als β_2, da die Werte von $[r/100]$ um den Faktor (1/100) kleiner sind als r.

Änderung von Skalen, Messeinheiten: Bsp 2

Werden hingegen nur die Einheiten, in denen C (der anhängigen Variablen) gemessen wird, geändert zB

$$C \ \rightarrow \ C \cdot 1000$$

so entspricht dies einer Multiplikation der gesamten Gleichung mit 1000, wobei die Daten für Y^d und r unverändert bleiben.

$$[C \cdot 1000] = (\beta_0 \cdot 1000) + (\beta_1 \cdot 1000)\, Y^d + (\beta_2 \cdot 1000) r + [1000 \cdot u]$$

- Alle Koeffizienten werden mit 1000 multipliziert.
- Es verändert sich auch die Skala des Fehlers, und damit auch die Fehlervarianz ($\times 1000^2$).

Aggregation über die Zeit

Monats- und Jahresdaten

Wir modellieren die Dynamik der (nominellen) Renditen des S&P500, ρ, für die Periode 1967M03-1993M08 univariat (autoregressiv). Es liegen Monatsdaten im Datenfile `sp500.wf1` vor.

$$\rho_t = 0.005 + 0.279\rho_{t-1} + \widehat{u}_t$$

Die *Rendite in stetiger Zeit* bez Preis P_t ist definiert als die Differenz der Logarithmen: $\rho_t = \log(P_t) - \log(P_{t-1})$.

Die langfristige (GGW-)Rendite, ρ_∞, erhält man über $\rho_\infty = \rho_t = \rho_{t-1}$

$$\rho_\infty = 0.005 + 0.279\rho_\infty$$

ρ_∞ ergibt sich als $\rho_\infty = 0.005/(1 - 0.279) = 0.018$.

Monats- und Jahresdaten

Die durchschnittliche *monatliche* Rendite beträgt 0.018, 1.8%.

Die durchschnittliche *Jahresrendite* - in der wir üblicherweise denken - ergibt sich als die aggregierte durchschnittliche monatliche Rendite. Das ist das 12-fache der monatlichen.

$$\rho_\infty\text{-annualisiert} = 0.018 \cdot 12 = 0.21 = 21\%$$

Dimension des Modells

Dimension des Modells

Univariates Modell, autoregressives Modell der Ordnung 1:

Die einfachste Variante eines bivariaten Modells ist

$$y_t = \beta_0 + \beta_1 y_{t-1} + u_t$$

Wir regressieren eine Variable auf ihre eigene Vergangenheit. Wir verlangen dabei $|\beta_1| < 1$.

Bivariates Regressionsmodell:

Wir haben mit dem bivariaten Modell mit einer Erklärenden begonnen

$$y = \beta_0 + \beta_1 x + u$$

Multiples Regressionsmodell:

Die nächstliegende Erweiterung ist das multiple Modell mit k Erklärenden

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

Dimension des Modells

System von Gleichungen:

Eine weitere Verallgemeinerung in einer linearen Welt sind simultane Systeme von Gleichungen.
ZB mit 2 endogenen Variablen y_1 und y_2:

$$y_1 = \beta_0 + \beta_1 y_2 + \beta_2 z_1 + u$$
$$y_2 = \gamma_0 + \gamma_1 y_1 + \gamma_2 z_2 + v$$

In der Regel sind ökonomische Beziehungen in einem System eingebettet.
Die Frage ist stets, ob ein einfaches Modell (vgl Occam's razor) eine ausreichend gute Approximation in einer komplexen Welt darstellt.

Dimension des Modells

Bsp: Ein einfaches Makromodell mit Konsum-, Investitionsfunktion und Identität

$$
\begin{aligned}
C_t &= \beta_0 + \beta_1 Y_t + \beta_2 C_{t-1} + u \\
I_t &= \gamma_0 + \gamma_1 Y_{t-1} + \gamma_2 r_t + v \\
Y_t &= C_t + I_t + G_t
\end{aligned}
$$

C ... private consumption
I ... gross investments
r ... interest rate
G ... government expenditures

Dimension des Modells

Vektor autoregressive Modelle (in Standardform):

Die direkte Verallgemeinerung des univariaten autoregressiven Modells auf 2 endogene Variablen ist

$$
\begin{aligned}
y_{1,t} &= \mu_1 + \phi_{11}\, y_{1,t-1} + \phi_{12}\, y_{2,t-1} + u_{1,t} \\
y_{2,t} &= \mu_2 + \phi_{21}\, y_{1,t-1} + \phi_{22}\, y_{2,t-1} + u_{2,t}
\end{aligned}
$$

Es werden dynamische Beziehungen zwischen den Variablen modelliert. $u_{1,t}$ und $u_{2,t}$ dürfen kontemporär korrelieren.

Lineare und nicht-lineare Modelle

Lineare Modelle sind linear in den Parametern

Wir verstehen (hier) unter einem linearen Modell, ein Modell, das **linear in den Parametern** ist. ZB

$$\log(y) = \beta_0 + \beta_1 \log(x) + u$$

ist linear in den Parametern und kann wie gewohnt mit OLS geschätzt werden. Wir müssen nur die Variablen y und x geeignet transformierten und das Modell passend anschreiben

$$\tilde{y} = \beta_0 + \beta_1 \tilde{x} + u$$

Lineare Modelle

Die Cobb-Douglas Produktionsfunktion

$$Q = \gamma K^{\alpha} L^{\beta}$$

ist nicht linear in den Parametern, kann aber durch Transformation in ein Modell, das linear in den Parametern ist, übergeführt werden. Wir logarithmieren

$$\log(Q) = \tilde{\gamma} + \alpha \log(K) + \beta \log(L) + u$$

Nicht-lineare Modelle

Es gibt mehr nicht-lineare Modelle als lineare. Ein Bsp ist die *logistsche Wachstumsfunktion*.

$$y(t) = \frac{y^*}{1 + \exp\left(-\alpha(t - t_0)\right)}$$

$y^* \ldots$ Sättigungsniveau (wenn $t \to \infty$), $\alpha \ldots$ beeinflusst die Wachstumsrate. Nach t abgeleitet, $\dot{y} = dy/dt$, erhält man die *Diffusionsgleichung*

$$\dot{y} = (\alpha/y^*)(y^* - y)y = \alpha y - (\alpha/y^*)y^2$$

Sie wird als lineares Modell in diskreter Zeit geschätzt.

Bsp sind die Ausbreitung einer Innovationen, aber auch einer Pandemie. Zur Empirie von Pandemien s zB Reuther und Reuther(2023).

Reuther, G. und Reuther, R. (2023): Hauptsache Panik. Ein neuer Blick auf Pandemien in Europa. Leipzig.

Nicht-lineare Modelle

In der VWL approximiert man in der Regel *nicht-lineare Funktionen* lokal durch eine Taylorreihe erster Ordnung. Dadurch erhält man lineare Approximationen, die als lineare Modelle angeschrieben werden.

Nicht-lineare Modelle

Typische nicht-lineare Modelle sind *Logit- und Probit-Modelle*. Sie dienen zur Modellierung von binären Variablen. Es wird die Wahrscheinlichkeit des Eintretens durch dritte Variable erklärt.

Bsp:

- Kaufentscheidung bei dauerhaften Gütern
  - kaufe einen Ultra-HD-Fernseher ja/nein-

$$\mathrm{P}(\mathrm{Kauf}_i) = f(\mathrm{Einkommen}_i, \ldots)$$

- Einstellungszusage für einen Langzeitarbeitslosen
  - wird eingestellt/nicht eingestellt

$$\mathrm{P}(\mathrm{Einstellung}_i) = f(\mathrm{Förderung}_i, \ldots)$$

Deterministische Trends

- Linearer Trend
- Quadratischer Trend
- Exponentielles Wachstum

Trends; linearer, quadratischer Trend

Ein *Trend* ist ein monotones Steigen/Fallen über die Zeit *ohne* ökonomische Erklärung.

- Linearer Trend:
  Das Modell eines linearen Trends ist

$$y_t = \beta_0 + \beta_1 t + u_t$$

- Quadratischer Trend:

$$y_t = \beta_0 + \beta_1 t + \beta_2 t^2 + u_t$$

  Er besitzt ein Minimum/Maximum bei $t_0 = -\beta_1/(2\beta_2)$.

Exponentieller Trend

- Exponentieller Trend:

$$y_t = \gamma \exp(\rho t)\, v_t$$

Die Funktion ist nicht-linear. Durch Logarithmieren erhalten wir aber

$$\log(y_t) = \tilde{\gamma} + \rho t + u_t$$

$\tilde{\gamma} = \log(\gamma)$, $u_t = \log(v_t)$.

ρ ist die *durchschnittliche Wachstumsrate*.

$$\rho = \mathrm{E}[\log(y_t) - \log(y_{t-1})] = \mathrm{E}[\Delta \log(y_t)]$$

$\log(y_{t-1}) = \tilde{\gamma} + \rho(t-1)$ (ohne Störterm)
$\log(y_t) - \log(y_{t-1}) = \rho$ (ohne Störterm)

Exponentieller Trend

- Exponentielles Wachstum im Niveau entspricht einem linearen Wachstum in den logarithmierten Daten.
- Dieses Verhalten findet man in allen Reihen, die real und/oder nominell wachsen (Inflation).

Im Zuge der Schätzung von

$$\log(y_t) = \tilde{\gamma} + \rho t + u_t$$

unterstellen wir, dass $u \sim \mathrm{N}(0, \sigma^2)$. Wollen wir auf das Niveau y

$$\mathrm{E}(y_t) = \gamma \exp(\rho t)\, \mathrm{E}[\exp(u_t)]$$

schließen, benötigen wir $\mathrm{E}[\exp(u_t)]$.

Exponentieller Trend, lognormal Verteilung

Lognormal-Verteilung:
Ist u normal verteilt mit $\mathrm{E}(u) = 0, \mathrm{V}(u) = \sigma^2$, so ist v, $v = \exp(u)$, *lognormal verteilt* mit

$$\mathrm{E}(v) = \exp(\sigma^2/2)$$

- Für kleine σ^2-Werte ist $\mathrm{E}(v) \approx 1$.
- Für große kann man die statistische Beziehung *nicht* ignorieren. Das ist zB für die Prognose relevant.

Bem: $y_t = \gamma \exp(\rho t)\, v_t$ und $\log(y_t) = \tilde{\gamma} + \rho t + u_t$ mit $v_t = \exp(u_t)$.

Loglog-, Semilog-Form

Elastizität

Die *Elastizität* misst *die prozentuelle Veränderung einer Funktion* $f(x)$ *bei prozentueller Veränderung von* x.

$$Ef(x) = \frac{df(x)/f(x)}{dx/x} = \frac{x}{f(x)} f'(x) = \frac{d \log f(x)}{d \log x}$$

Das zugehörige Modell bei konstanter Elastizität ist

$$\log(y) = \beta_0 + \beta_1 \log(x) + u$$

β_1 ist die Elastizität.

Die *Steigung* einer Funktion $f(x)$ ist hingegen *die Veränderung von* $f(x)$ *bei Veränderung von* x. ($x \to x + 1$).
Das zugehörige Modell bei konstanter Steigung ist

$$y = \beta_0 + \beta_1 x + u$$

Loglog-Form

Das standard Bsp für die Loglog-Form ist die Cobb-Douglas Produktionsfunktion

$$Q = \gamma K^{\alpha} L^{\beta} v$$

Logarithmieren ergibt die *Loglog-Form*

$$\log(Q) = \tilde{\gamma} + \alpha \log K + \beta \log(L) + u$$

α und β sind die Kapital- bzw Arbeits-Elastizitäten des Outputs. Die Funktion ist offensichtlich linear in den Parametern.

Schließt man von $\log(Q)$ auf Q, ist gegebenenfalls $\mathrm{E}(v) > 1$ zu berücksichtigen.

Semilog-Form

Es gibt 2 Varianten der Semi-Elastizität.

A Die *Semi-Elastizität* misst *die prozentuelle Veränderung von $f(x)$ bei einer (nicht prozentuellen) Veränderung in x*.

$$Sf_A(x) = \frac{d \log f(x)}{dx} = \frac{1}{f(x)} f'(x)$$

Das zugehörige Regressionsmodell bei konstanter Semi-Elastizität lautet (*Semilog-Form A*)

$$\log(y) = \beta_0 + \beta_1 x + u$$

β_1 ist die Semi-Elastizität (Typ A) von y bezüglich x.

Die Lösung für y ist

$$y = \tilde{\beta}_0 \exp(\beta_1 x)\, v \qquad \text{mit} \qquad v = \exp(u)$$

Sie ist vergleichbar mit einem expo Wachstum in *der Zeit* x.

Semilog-Form

B *Semi-Elastizität* wird auch für *Veränderungen von $f(x)$ bei prozentueller Veränderung von x* verwendet.

$$Sf_B(x) = \frac{df(x)}{d\log(x)} = \frac{df(x)}{dx}x = f'(x)\,x$$

Das zugehörige Regressionsmodell mit konstanter Semi-Elastizität lautet (*Semilog-Form B*)

$$y = \beta_0 + \beta_1 \log(x) + u$$

β_1 ist die Semi-Elastizität (Typ B) von y bezüglich x.

In Worten:

A: Eine Veränderung von x um 1 (eine Einheit) erhöht y um β_1%.

B: Eine Veränderung von x um 1% erhöht y um β_1 Einheiten.

Dummy Variable

Dummy Variable

Dummy Variable sind ein mächtiges aber zugleich einfaches Instrument um Aussagen zu präzisieren und um zu differenzieren.
Technisch gesehen ist eine Dummy eine *binäre Variable*. Sie kann 2 Werte annehmen:

$$0 \qquad \text{oder} \qquad 1$$

Mit der ’1’ markieren wir ausgewählte Beobachtungseinheiten.

Wir behandeln hier Dummies nur als erklärende Variable. (Will man ein binäres Verhalten - kaufe keine/eine Waschmaschine - modellieren, verwendet man Logit oder Probit Modelle.)

Bsp 1: Dummy Variable und Interzept

Bsp 1: Sie wollen den Einkommensunterschied zwischen Frauen und Männern untersuchen.

Sie markieren zB die Frauen mit 1. Jeder Mann erhält eine 0.

$$W_i = \beta_0 + \beta_1 D_i + u_i$$

$W \ldots$ Lohnsatz,

$$D = \begin{cases} 0 & \text{für männlich (M)} \\ 1 & \text{für weiblich (F)} \end{cases}$$

Bsp 1: Dummy Variable und Interzept

Die zugehörige $\boldsymbol{X}$-Matrix hat die Form

i	**1**	D	
1	1	0	M
2	1	0	M
3	1	1	F
...	...	...	...

Die Gleichungen für die ersten Beobachtungen lauten (ohne Fehler)

$$
\begin{aligned}
W_1 &= \beta_0 + \beta_1 \cdot 0 = \beta_0 \\
W_2 &= \beta_0 + \beta_1 \cdot 0 = \beta_0 \\
W_3 &= \beta_0 + \beta_1 \cdot 1 = \beta_0 + \beta_1
\end{aligned}
$$

Bsp 1: Dummy Variable und Interzept

- β_0 gibt daher den Durchschnittslohn der Männer, $\overline{W}_M$, an.
- $(\beta_0 + \beta_1)$ ist der Durchschnittslohn der Frauen, $\overline{W}_F$.
- β_1 ist die durchschnittliche *Lohndifferenz* von Frauen zu Männern.

$$\beta_1 = \overline{W}_F - \overline{W}_M$$

Allgemein:

- Die Konstante modelliert die nicht-markierten Einheiten (0),
- die Koeffizienten vor den Dummies, die durchschnittliche Abweichung zu den nicht-markierten Einheiten (0).

Bsp 1: 2 Dummy Variable, kein Interzept

Alternativ zu 'Interzept und Dummy' können wir 2 Dummies verwenden:

- *DM* gibt den Männern eine 1 und den Frauen eine 0.
- *DF* gibt den Frauen eine 1 und den Männern eine 0.

$$W_i = \beta_1 DM_i + \beta_2 DF_i + u_i$$

Die zugehörige $\boldsymbol{X}$-Matrix hat die Form

i	*DM*	*DF*	
1	1	0	M
2	1	0	M
3	0	1	F
...	...	...	...

Bsp 1: 2 Dummy Variablen, kein Interzept

Die Gleichungen für die ersten Beobachtungen lauten (ohne Fehler)

$$\begin{aligned} W_1 &= \beta_1 \cdot 1 + \beta_2 \cdot 0 &= \beta_1 \\ W_2 &= \beta_1 \cdot 1 + \beta_2 \cdot 0 &= \beta_1 \\ W_3 &= \beta_1 \cdot 0 + \beta_2 \cdot 1 &= \beta_2 \end{aligned}$$

β_1 zeigt nun den Durchschnittslohn der Männer, $\overline{W}_M$, und
β_2 ist der Durchschnittslohn der Frauen, $\overline{W}_F$.

Dummy Variablen Falle, Dummy Variable Trap

Man könnte geneigt sein, eine *Konstante, eine Dummy für Frauen und zusätzlich eine Dummy für Männer* in das Regressionsmodell aufzunehmen. Damit ist das Modell aber überbestimmt, bzw die zugehörigen Parameter sind *nicht* identifiziert. Technisch erkennt man dies daran, dass der Rang der $\boldsymbol{X}$-Matrix nicht $(k+1)$ ist, bzw $(\boldsymbol{X}'\boldsymbol{X})$ nicht invertierbar (singulär) ist.

Andererseits wissen wir aus dem Fall:
'Männer-Dummy und Frauen-Dummy aber kein Interzept', dass

$$\beta_1 + \beta_2 = \overline{W}_M + \overline{W}_F = \overline{W}$$

Daher macht es keinen Sinn den durchschnittlichen Lohnsatz nochmals mittels Interzept im Modell berechnen zu wollen.

Dummy Variablen Falle

Regel:

- 1 Männer- und 1 Frauen-Dummy, oder
- Konstante und Männer-Dummy, oder
- Konstante und Frauen-Dummy,
- aber NICHT
  Konstante *und* Männer- *und* Frauen-Dummy.

Bsp 2: Variante ohne Interzept

Bsp 2: Sie könnten auch daran interessiert sein die Anzahl der Arztbesuche p.a., A, nach Alterskohorten zu unterscheiden.
Sie verwenden 10 Dummies (oder 9 plus Interzept).
$D0 \ldots$ Alter zw 0 und 10 Jahren,
$\ldots$,
$D9 \ldots$ Alter zw 90 und 100, und darüber.

$$A_i = \gamma_0 D0_i + \ldots + \gamma_9 D9_i + u_i$$

Bsp 2: Variante ohne Interzept

Es könnte sich herausstellen, dass D2 bis D5 ähnliche Koeffizienten haben, aber alle von null verschieden sind.
Eine einfache Methode mit der man einen Eindruck über die Signifikanz der Unterschiede bekommt ist, die Dummy einer mittleren Kohorte, zB D4, wegzulassen und stattdessen den Interzept zu verwenden.
Die Koeffizienten zeigen nun den Unterschied zu dieser Kohorte an. Die automatisch berechneten *t*-Tests zeigen deren Signifikanzen an.

$$A_i = \beta_0 + \beta_1 D0_i + \ldots + \beta_4 D3_i + \beta_5 D5_i + \ldots + \beta_9 D9_i + u_i$$

Bsp 2: Variante ohne Interzept

Wenn sie auf die Gleichheit der Parameter vor $D2$ bis $D5$ testen wollen, bietet sich ein F-Test an.

$$H_0 : \gamma_2 = \ldots = \gamma_5$$

$$H_1 : \gamma_i \neq \gamma_j \qquad \text{mit} \qquad i \neq j, \; i, j \in \{2, \ldots, 5\}$$

Das Modell unter H_0 lautet

$$A_i = \gamma_0 D0_i + \gamma_1 D1_i + \gamma[D2_i + \ldots + D5_i] + \gamma_6 D6_i + \ldots + \gamma_9 D9_i + u_i$$

Bsp 1: Erweiterung des Modells

Die Dummy Variablen können natürlich mit anderen Dummies, anderen erklärenden Variablen und funktionalen Formen kombiniert werden. ZB

$$W_i = \beta_0 + \beta_1 DF_i + \beta_2 B_i + u_i$$

$B \ldots$ höchste abgeschlossene Bildung

Dummy Variable zum Eliminieren von Ausreißer

Singuläre Ereignisse stören oft einen sonst stabilen Zusammenhang. Solche Ereignisse erkennt man (in der Regel) an den Residuen als besonders große Werte.

Zum Eliminieren konstruiert man eine Dummy die genau für diese Beobachtung eine 1 und sonst 0 hat.

In der Regressionsgleichung nimmt der Koeffizient vor der Dummy den gesamten Effekt dieser einen Beobachtung auf. Damit beeinflusst dieser 'Ausreisser' die Schätzung der anderen Parameter nicht.

Referenzen

Hackl (16-22), 14.1, 9.2
Wooldridge 2.4, 10.5, 7.2

EViews Befehle

Erzeugen einer Dummy-Variablen:
Zum Markieren einzelner Beobachtungen:
Reihe `dum = 0` erzeugen. Im Daten-Fenster von `dum` die einzelnen Werte editieren und wenn gewünscht 1 setzen.

Zum Markieren der Männer, Variable `dm`, unter Verwendung einer bestehenden Variablen `sex`, die mit M='1' und F='2' kodiert ist:
Reihe `dm = 2 - sex` anlegen.