

Übung "Multikollinearität"

Diese Übung demonstriert die Auswirkungen von Multikollinearität anhand künstlich generierter Daten. Es ist der folgende datengenerierende Prozess vorgegeben:

$$y_t = \alpha + \beta_1 x_{1t} + \beta_2 x_{2t} + u_t \quad u_t \sim IN(0, \sigma^2)$$

$$\alpha = 2 \quad \beta_1 = 4 \quad \beta_2 = 3 \quad \sigma = 5, 10, 20$$

$$x_{1t} = 0.9x_{1t-1} + \varepsilon_t \quad x_{10} = 0 \quad \varepsilon_t \sim IN(0, \sigma_\varepsilon^2) \quad \sigma_\varepsilon = 2$$

$$x_{2t} = 1.5x_{1t} + \eta_t \quad \eta_t \sim IN(0, \gamma^2) \quad \gamma = 5, 0.5, 0.05$$

Im nachfolgenden Eviews-Programm (multicoll.prg) kann man durch Wahl von σ ($sigma = 5, 10, 20$) und γ ($gamma = 5, 0.5, 0.05$) unterschiedliche Stichproben für x_{1t} , x_{2t} , y_t mit je 108 Beobachtungswerten erzeugen:

- Vergrößert man $sigma$, so wird der systematische Einfluss von x_{1t} , x_{2t} auf y_t von stärkeren Störeinflüssen u_t überlagert.
- Verkleinert man $gamma$, so verstärkt sich die Korrelation zwischen x_{1t} und x_{2t} (Multikollinearität).

Beides wirkt sich ungünstig auf die Präzision der Parameterschätzungen aus: Grosse Standardfehler der Schätzungen für α , β_1 und β_2 .

Die Übung dient zur Illustration.

Eviews-Programm (multicoll.prg)

- ' multicoll.prg: Regression mit künstlich generierten Daten - Multikollinearitätsproblem
- ' Schätzung einer Gleichung $y = \alpha + \beta(1) \cdot x_1 + \beta(2) \cdot x_2$

- ' Öffnen eines Workfiles und Festlegen des Stichprobenumfangs

```
workfile simdat Q 1970:1 1996:4
```

- ' Generieren einer autokorr. Variablen x_1 mit Erwartungswert 0

```
rdseed(type=kn4) 416
genr x1 = 0
smpl 1970:2 1996:4
genr x1 = 0.9*x1(-1)+2*nrnd
smpl 1970:1 1996:4
```

- ' Generieren einer Variablen x_2 , die je nach dem Wert von γ

- ' mehr oder weniger stark mit x_1 korreliert ist

```
coef(1) gamma
gamma(1) = 5
rdseed(type=kn4) 136
genr x2 = 1.5*x1+gamma(1)*nrnd
```

- ' Berechnung der Korrelation zwischen x_1 und x_2

```
coef(1) rho
rho(1) = @cor(x1,x2)
plot x1 x2
```

- ' Generieren von y mit einem "wahren Modell" mit den Parametern

- ' $\alpha=2$, $\beta(1)=4$, $\beta(2)=3$ und $\sigma=5$

```
rdseed 552
coef(1) sigma
sigma(1) = 5
genr u = sigma(1)*nrnd
genr y = 2+4*x1+3*x2+u
```

- ' Schätzung der Gleichung

```
coef(1) alfa
coef(2) beta
equation test.ls y = alfa(1)+beta(1)*x1+beta(2)*x2
```

- ' Die Aufgabe besteht darin, γ und σ zu variieren und die

- ' Auswirkungen auf die Schätzgleichung zu analysieren:

- ' - Mit kleinerem γ verstärkt sich die Multikollinearität zwischen
- ' x_1 und x_2 . Verkleinern Sie γ z.B. von 5 auf 0.5, 0.05 und 0.005 !
- ' - Mit grösserem σ wird der Zusammenhang zwischen x_1 , x_2 und y
- ' von grösserem Störeinfluss überlagert. Vergrössern Sie σ z.B.
- ' von 5 auf 10 und 20 !

Vorgehen in EViews:

1. Öffnen sie über FILE / OPEN das Programm multicoll.prg.
2. Lassen sie die Parameter GAMMA und SIGMA zunächst auf den Werten, die im Programm vorgegeben sind und lassen Sie das Programm durch Anklicken von RUN laufen.
3. Schauen Sie sich die Resultate an, indem Sie im WORKFILE den Koeffizienten RHO und die Gleichung TEST anklicken.
4. Lassen Sie die Fenster für RHO und TEST geöffnet und verändern Sie nun GAMMA (5, 0.5, 0.05, 0.005) und dann SIGMA (5, 10, 20) (es gibt 12 Kombinationen).

Betrachten Sie die Auswirkungen auf die Standardfehler der Parameterschätzungen bzw. die t-Werte!

Fragen:

1. Testen Sie die falsche Hypothese $\beta(1) = 0$ und die (etwas weniger) falsche Hypothese $\beta(1) = \beta(2)$ in den 9 Situationen, die sich durch die Parameterwahl $\text{GAMMA} = 5, 0.5, 0.05$ und $\text{SIGMA} = 5, 10, 20$ ergeben. Erstellen Sie eine Tabelle mit den marginalen Signifikanzniveaus und kommentieren Sie die Ergebnisse!
2. Überlegen Sie sich, wie in einer Situation mit ausgeprägter Multikollinearität (zu verfahren ist, um präzisere Parameterschätzungen zu erhalten! Vergleichen Sie für $\text{GAMMA} = 0.05$ und $\text{SIGMA} = 5$ die unrestringierte Schätzung mit
 - a) einer Schätzung mit der (etwas falschen) Restriktion $\beta(1) = \beta(2)$.
 - b) einer Schätzung mit der (drastisch falschen) Restriktion $\beta(2) = 0$
 - c) einer Schätzung mit der (richtigen) Restriktion $\beta(2) = 3$.