Übung "Multikollinearität"

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

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

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

$$x_{1t} = 0.9 x_{1t-1} + \varepsilon_t \qquad x_{1_0} = 0 \qquad \varepsilon_t \sim IN(0, \sigma_\varepsilon^2) \quad \sigma_\varepsilon = 2$$

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

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 Beobachtugswerten 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 = alfa+beta(1)*x1+beta(2)*x2

' Öffnen eines Workfiles und Festlegen des Stichprobenumfangs
        workfile simdat Q 1970:1 1996:4

' Generieren einer autokorr. Variablen x1 mit Erwartungswert 0
        rndseed(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 x2, die je nach dem Wert von gamma
' mehr oder weniger stark mit x1 korreliert ist
        coef(1) gamma
        gamma(1) = 5
        rndseed(type=kn4) 136
        genr x2 = 1.5*x1+gamma(1)*nrnd

' Berechnung der Korrelation zwischen x1 und x2
        coef(1) rho
        rho(1) = @cor(x1,x2)
        plot x1 x2

' Generieren von y mit einem "wahren Modell" mit den Parametern
' alfa=2, beta(1)=4, beta(2)=3 und sigma=5
        rndseed 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, gamma und sigma zu variieren und die
' Auswirkungen auf die Schätzgleichung zu analysieren:
' - Mit kleinerem gamma verstärkt sich die Multikollinearität zwischen
'    x1 und x2. Verkleinern Sie gamma z.B. von 5 auf 0.5, 0.05 und 0.005 !
' - Mit grösserem sigma wird der Zusammenhang zwischen x1, x2 und y
'   von grösserem Störeinfluss überlagert. Vergrössern Sie sigma 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 GAMMA = 5, 0.5, 0.05 und 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 GAMMA = 0.05 und 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.