universität
innsbruck

Multikollinearität

Grundlagen der Ökonometrie

herbert.stocker@uibk.ac.at

www.hsto.info/econometrics

Multikollinearität

2 Fälle:

1. **Perfekte Multikollinearität:** es existiert eine exakte lineare Abhängigkeit zwischen Regressoren.

Multikollinearität

2 Fälle:

1. **Perfekte Multikollinearität:** es existiert eine exakte lineare Abhängigkeit zwischen Regressoren.
   - Verletzt Gauss-Markov Annahme A2: es existiert keine eindeutige Lösung für OLS Schätzfunktion!
   - Beispiel: *Dummyvariablenfalle*: $y_i = \widehat{\beta}_1 + \widehat{\beta}_2 m_i + \widehat{\beta}_3 w_i + \varepsilon_i \quad (1 = m_i + m_i)$
   - einfach zu erkennen, tatsächlich kaum ein Problem ...

Multikollinearität

2 Fälle:

1. **Perfekte Multikollinearität:** es existiert eine exakte lineare Abhängigkeit zwischen Regressoren.
   - Verletzt Gauss-Markov Annahme A2: es existiert keine eindeutige Lösung für OLS Schätzfunktion!
   - Beispiel: *Dummyvariablenfalle*: $y_i = \widehat{\beta}_1 + \widehat{\beta}_2 m_i + \widehat{\beta}_3 w_i + \varepsilon_i \quad (1 = m_i + m_i)$
   - einfach zu erkennen, tatsächlich kaum ein Problem . . .

2. **(hohe) Multikollinearität:** Regressoren sind untereinander hoch, aber nicht perfekt, korreliert

Multikollinearität

2 Fälle:

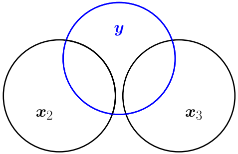
1. **Perfekte Multikollinearität:** es existiert eine exakte lineare Abhängigkeit zwischen Regressoren.
   - Verletzt Gauss-Markov Annahme A2: es existiert keine eindeutige Lösung für OLS Schätzfunktion!
   - Beispiel: *Dummyvariablenfalle*: $y_i = \widehat{\beta}_1 + \widehat{\beta}_2 m_i + \widehat{\beta}_3 w_i + \varepsilon_i \quad (1 = m_i + m_i)$
   - einfach zu erkennen, tatsächlich kaum ein Problem . . .

2. **(hohe) Multikollinearität:** Regressoren sind untereinander hoch, aber nicht perfekt, korreliert
   - verletzt keine Gauss-Markov Annahme!
     $\Rightarrow$ OLS erwartungstreu, effizient & konsistent!!!
   - trotzdem, Schätzungen sind ungenau (große Standardfehler)!
   - ähnelt Problemen mit kleiner Stichprobe
     $\rightarrow$ wenig (unabhängige) Information in den Daten!

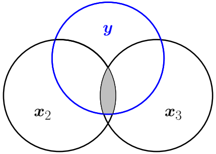
Multikollinearität

Venn-Diagramme:

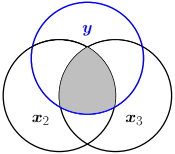
a) Keine Multikollinearität:



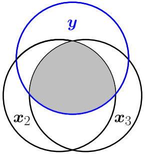
b) Niedrige Multikollinearität:



c) Hohe Multikollinearität:



d) Sehr hohe Multikollinearität:



Multikollinearität

Konsequenzen der Multikollinearität:

- OLS Schätzfunktionen sind erwartungstreu, effizient und konsistent, *aber ungenau*!!!

$$\mathrm{var}(\widehat{\beta}_h) = \frac{\sigma^2}{(1 - R_h^2) \sum_{i=1}^{n} (x_{ih} - \overline{x}_{\cdot h})^2}$$

mit $x_{ih} = \hat{\alpha}_1 + \hat{\alpha}_2 x_{i2} + \hat{\alpha}_{h-1} x_{ih-1} + \hat{\alpha}_{h+1} x_{ih+1} + \ldots \hat{\alpha}_k x_{ik} + \hat{\varepsilon}_i^* \quad \Rightarrow \quad R_h^2$

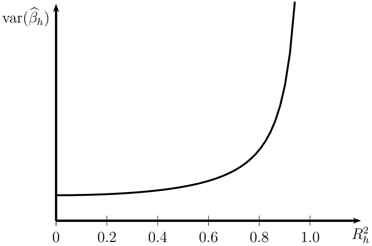
Multikollinearität

Konsequenzen der Multikollinearität:

- OLS Schätzfunktionen sind erwartungstreu, effizient und konsistent, *aber ungenau*!!!

$$\mathrm{var}(\widehat{\beta}_h) = \frac{\sigma^2}{(1 - R_h^2)\sum_{i=1}^{n}(x_{ih} - \overline{x}_{\cdot h})^2}$$

mit $x_{ih} = \hat{\alpha}_1 + \hat{\alpha}_2 x_{i2} + \hat{\alpha}_{h-1} x_{ih-1} + \hat{\alpha}_{h+1} x_{ih+1} + \ldots \hat{\alpha}_k x_{ik} + \hat{\varepsilon}_i^* \quad \Rightarrow \quad R_h^2$



- wenn $R_h^2 = 0$: $\mathrm{var}(\widehat{\beta}_h)$ $\rightarrow$ bivariate Regression
- wenn $R_h^2 > 0.9$: ähhh ...
- wenn $R_h^2 \rightarrow 1$: $\mathrm{var}(\widehat{\beta}_h) \rightarrow \infty$!

Multikollinearität

Achtung: Multikollinearität generell nur diskutieren, wenn besondere Gründe dafür sprechen!

Erkennung von Multikollinearität:

- Hohes R^2 und wenige signifikante Koeffizienten.
- Schätzergebnisse schwanken stark in Abhängigkeit von der gewählten Spezifikation und/oder bei Ausschluss einzelner Beobachtungen ($\rightarrow$ Interpretation wird häufig zum Kaffeesudlesen . . .)
- Hohe paarweise Korrelationen zwischen Regressoren; besser
- **Variance Inflation Factor** (VIF):

$$\text{VIF}_h = \frac{1}{1 - R_h^2}$$

mit R_h^2 Bestimmtheitsmaß einer Hilfsregression (s.o.);
(VIF > 10, bzw. $R_h^2 > 0.9$ Hinweis auf Probleme)

Multikollinearität

Maßnahmen bei Multikollinearität:

- Zusätzliche Daten erheben: selten eine Option, und hilft wenig.

Multikollinearität

Maßnahmen bei Multikollinearität:

- Zusätzliche Daten erheben: selten eine Option, und hilft wenig.
- Externe Information einbeziehen: auch schwierig

Multikollinearität

Maßnahmen bei Multikollinearität:

- Zusätzliche Daten erheben: selten eine Option, und hilft wenig.
- Externe Information einbeziehen: auch schwierig
- Variablen ‘weglassen’: . . . den Teufel mit dem Beelzebub austreiben

Multikollinearität

Maßnahmen bei Multikollinearität:

- Zusätzliche Daten erheben: selten eine Option, und hilft wenig.
- Externe Information einbeziehen: auch schwierig
- Variablen ‘weglassen’: . . . den Teufel mit dem Beelzebub austreiben
- Spezielle Schätzverfahren (LASSO, Ridge Regression, . . .): immer ein trade-off, selten eine Lösung

Multikollinearität

Maßnahmen bei Multikollinearität:

- Zusätzliche Daten erheben: selten eine Option, und hilft wenig.
- Externe Information einbeziehen: auch schwierig
- Variablen ‘weglassen’: . . . den Teufel mit dem Beelzebub austreiben
- Spezielle Schätzverfahren (LASSO, Ridge Regression, . . .): immer ein trade-off, selten eine Lösung
- **Sich damit abfinden** – manchmal nicht die schlechteste Lösung . . .

Multikollinearität

Maßnahmen bei Multikollinearität:

- Zusätzliche Daten erheben: selten eine Option, und hilft wenig.
- Externe Information einbeziehen: auch schwierig
- Variablen ‘weglassen’: . . . den Teufel mit dem Beelzebub austreiben
- Spezielle Schätzverfahren (LASSO, Ridge Regression, . . .): immer ein trade-off, selten eine Lösung
- **Sich damit abfinden** – manchmal nicht die schlechteste Lösung . . .

Multikollinearität

Maßnahmen bei Multikollinearität:

- Zusätzliche Daten erheben: selten eine Option, und hilft wenig.
- Externe Information einbeziehen: auch schwierig
- Variablen ‘weglassen’: . . . den Teufel mit dem Beelzebub austreiben
- Spezielle Schätzverfahren (LASSO, Ridge Regression, . . .): immer ein trade-off, selten eine Lösung
- **Sich damit abfinden** – manchmal nicht die schlechteste Lösung . . .

Thanx . . .