Methoden der Ökonometrie

Fragenkatalog - Stand: 4.12.2019[1]

Dieser Katalog umfasst 60 Seiten und besteht aus 68 Aufgaben mit insgesamt 695 Punkten.

Aufgabenübersicht

[1]Benutzerhinweise: Dieser Katalog darf für den individuellen Gebrauch und für Unterrichtszwecke, jedoch nicht für den kommerziellen Gebrauch gedruckt und reproduziert werden. Dozenten dürfen die Aufgaben online an Hochschulen, entweder zum individuellen Nutzen oder für dozentenbegleiteten Unterricht nutzen. Darüber hinaus wird ihnen gestattet, Kopien einer begrenzten Anzahl von Seiten für Studenten (nicht als Teil eines Lehrbuchs) zu erstellen, wobei folgendes Urheberzeichen auf jeder Kopie erscheinen muss:

Ich danke Kathrin Kagerer und Stefan Rameseder fur ihre Zuarbeiten, wichtigen Korrekturen und substantiellen Verbesserungsvorschläge sehr herzlich.

1. Aufgabe (11 Punkte) (*Lineare Gleichungssysteme und Matrizen*)

Gegeben sind zwei Gleichungen der Form $a_1x_1 + a_2x_2 = b$:

$$\text{I: } x_1 - x_2 = 1, \quad \text{II: } x_1 + 3x_2 = 9.$$

(a) (1 Punkt) Lösen Sie das lineare Gleichungssystem.

(b) (2 Punkte) Die Lösungen jeder Gleichung stellen eine Gerade dar. Berechnen Sie die Parametrisierung der Geraden in Abhängigkeit der Koeffizienten a_1, a_2 und b, d. h. eine Form

$$x_2 = mx_1 + t$$

und zeichnen Sie beide in ein (etwas größeres) Koordinatensystem. Stimmt Ihre Lösung von vorher mit dem Schnittpunkt überein?

(c) (1 Punkt) Schreiben Sie das Problem von oben als Matrixgleichung $\boldsymbol{Ax} = \boldsymbol{b}$. Welche Dimensionen haben $\boldsymbol{A}, \boldsymbol{b}$?

(d) (2 Punkte) Berechnen Sie die Determinate der Matrix $\boldsymbol{A}$ und damit die Inverse $\boldsymbol{A}^{-1}$ von $\boldsymbol{A}$. Berechnen Sie somit wieder die Lösung des Gleichungssystems.

(e) (1 Punkt) Sind die Spalten bzw. die Zeilen der Matrix $\boldsymbol{A}$ linear unabhängig? Begründung.

(f) (2 Punkte) Zeichnen Sie den Vektor $\boldsymbol{b}$ in Ihr Koordinatensystem. Machen Sie sich klar, dass die Matrix $\boldsymbol{A}^{-1}$ diesen Punkt **linear** (in diesem Koordinatensystem) auf die Lösung Ihres Gleichungssystems abbildet: Schreiben Sie dazu $\boldsymbol{b}$ als Linearkombination zweier Einheitsvektoren

$$\boldsymbol{b} = \lambda \boldsymbol{e}_1 + \mu \boldsymbol{e}_2 = \lambda \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \mu \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{mit } \lambda, \mu \in \mathbb{R},$$

berechnen **und** zeichnen Sie die Einzelschritte von $\boldsymbol{A}^{-1}\boldsymbol{b} = \lambda \boldsymbol{A}^{-1}\boldsymbol{e}_1 + \mu \boldsymbol{A}^{-1}\boldsymbol{e}_2$.

(g) (2 Punkte) Sehen Sie nach, was der Kern Ker und das Bild Im einer Matrix sind. Berechnen Sie $\mathrm{Ker}(\boldsymbol{A})$ und $\mathrm{Im}(\boldsymbol{A})$. Zeichnen Sie beides ein. Was ist der Rang einer Matrix bzw. $\mathrm{rk}(\boldsymbol{A})$?

Fazit: Es gibt eine 1:1 Korrespondenz zwischen dem Lösen von Gleichungssystemen und der Invertierbarkeit von linearen Abbildungen.

Lineare Abbildungen, also Funktionen $f : \mathbb{R}^n \to \mathbb{R}^m$ mit $f(\lambda \boldsymbol{v} + \mu \boldsymbol{w}) = \lambda f(\boldsymbol{v}) + \mu f(\boldsymbol{w})$ für beliebige $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$ und $\lambda, \mu \in \mathbb{R}$, können - nach Festlegung einer Basis, in diesem Kurs der Standardbasis - <u>eindeutig</u> durch Matrizen $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ dargestellt werden:

$$\text{Für } \boldsymbol{x} \in \mathbb{R}^n \text{ gilt: } f(\boldsymbol{x}) = \boldsymbol{Ax} \in \mathbb{R}^m$$

Nur quadratische Matrizen, d. h. im Falle $m = n$, können invertiert werden bzw. besitzen eine eindeutige Lösung. Im ökonometrischen Kontext $\boldsymbol{y} = \boldsymbol{X\beta}$ mit $\boldsymbol{X} \in \mathbb{R}^{m \times n}$ ist aber generell $m >> n$ (Anzahl der Beobachtungen (Gleichungen) >> Anzahl der Regressoren (unbekannten Parametern), das Gleichungssystem $\boldsymbol{y} = \boldsymbol{X\beta}$ hat also keine Lösung mehr und eine Lösung muss durch Wahl einer Verlustfunktion „geschätzt“ werden.

2. Aufgabe (10 Punkte) (*Eigenschaften von Matrizen*)

Untersuchen Sie die Matrizen $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$ und $\boldsymbol{D}$ hinsichtlich der folgenden Punkte. Die Matrizen seien dabei $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 0 \\ 4 & 5 & 0 \\ 0 & 0 & 9 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 2 & -4 \\ 0 & 2 \end{pmatrix}$, $\boldsymbol{C} = \begin{pmatrix} 1 & 4 \\ 2 & 7 \\ 3 & 6 \end{pmatrix}$, $\boldsymbol{D} = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$.

(a) (1 Punkt) Geben Sie die Dimensionen der Matrizen an.

(b) (1 Punkt) Welche dieser Matrizen sind quadratisch, symmetrisch, Diagonalmatrizen, Dreiecksmatrizen?

(c) (1 Punkt) Transponieren Sie alle Matrizen und geben Sie die Dimensionen von $\boldsymbol{A}^T$, $\boldsymbol{B}^T$, $\boldsymbol{C}^T$ und $\boldsymbol{D}^T$ an.

(d) (2 Punkte) Invertieren Sie die quadratischen Matrizen, bei denen dies möglich ist.

(e) (1 Punkt) Welche Matrizen lassen sich miteinander addieren, welche multiplizieren? Berechnen Sie ein Produkt.

(f) (1 Punkt) Zeigen Sie, dass Matrizen in Bezug auf Multiplikation im Allgemeinen nicht kommutativ sind,

(g) (3 Punkte) Berechnen Sie $\mathrm{Ker}(\boldsymbol{D})$, $\mathrm{Im}(\boldsymbol{D})$ und $\mathrm{rk}(\boldsymbol{D})$.

3. Aufgabe (4 Punkte) (*Positiv-definite und idempotente Matrizen*)

Eine $(k \times k)$-Matrix $\mathbf{M}$ heißt idempotent, falls $\mathbf{M} \cdot \mathbf{M} = \mathbf{M}$ gilt.

Eine $(k \times k)$-Matrix $\mathbf{M}$ heißt $\begin{cases} \text{positiv definit} \\ \text{positiv semi-definit} \\ \text{negativ definit} \\ \text{negativ semi-definit} \end{cases}$, wenn $\begin{cases} \mathbf{x}^T\mathbf{M}\mathbf{x} > 0 \\ \mathbf{x}^T\mathbf{M}\mathbf{x} \geq 0 \\ \mathbf{x}^T\mathbf{M}\mathbf{x} < 0 \\ \mathbf{x}^T\mathbf{M}\mathbf{x} \leq 0 \end{cases}$ für beliebige $(k \times 1)$-Vektoren $\mathbf{x} \neq \mathbf{0}$ gilt und indefinit, wenn keiner der Fälle für alle $(k \times 1)$-Vektoren $\mathbf{x} \neq \mathbf{0}$ gilt.

(a) (1 Punkt) Überprüfen Sie die Matrizen $\mathbf{B} = \begin{pmatrix} 2 & -4 \\ 0 & 2 \end{pmatrix}$ und $\mathbf{D} = \begin{pmatrix} 0.5 & 0.5 \\ 0.5 & 0.5 \end{pmatrix}$ auf Idempotenz.

(b) (2 Punkte) Zeigen Sie, dass die Matrix $\mathbf{B}$ positiv semi-definit ist.

(c) (1 Punkt) Ist die Matrix $\mathbf{E} = -\mathbf{B}$ positiv (semi-)definit, negativ (semi-)definit oder indefinit (keine Berechnung, nur Begründung)?

4. Aufgabe (6 Punkte) (*Rang einer Matrix*)

Nun seien die beiden Matrizen $\mathbf{X}_1$ und $\mathbf{X}_2$ gegeben,

$$\mathbf{X}_1 = \begin{pmatrix} 1 & 1 & -7 \\ 1 & 3 & -5 \\ 1 & 2 & -8 \\ 1 & 4 & -5 \end{pmatrix}, \qquad \mathbf{X}_2 = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 2 \\ 1 & 0 & 1 & 3 \\ 1 & 1 & 0 & 4 \\ 1 & 0 & 1 & 5 \end{pmatrix}.$$

(a) (3 Punkte) Bestimmen Sie den Rang von $\mathbf{X}_1$ und $\mathbf{X}_2$. Haben die Matrizen vollen Rang?

(b) (2 Punkte) Bestimmen Sie den Rang von $\mathbf{X}_1^T\mathbf{X}_1$ und $\mathbf{X}_2^T\mathbf{X}_2$. Was stellen Sie jeweils fest?

(c) (1 Punkt) Begründen Sie, ob $\mathbf{X}_1^T\mathbf{X}_1$ und $\mathbf{X}_2^T\mathbf{X}_2$ invertierbar sind. Sie brauchen die Matrizen dazu nicht zu invertieren.

5. Aufgabe (3 Punkte) (*Zum Skalarprodukt*)

Zeigen Sie die folgenden Eigenschaften des Skalarprodukts, indem Sie die Definition des Skalarprodukts nutzen. Seien $a, b \in \mathbb{R}$ und $\mathbf{x}$, $\mathbf{x}_1$, $\mathbf{x}_2$ und $\mathbf{y}$ $(n \times 1)$-Vektoren:

(a) (1 Punkt) $< a \cdot \boldsymbol{x} \, , \, b \cdot \boldsymbol{y} > = \; a \cdot b \cdot < \boldsymbol{x} \, , \, \boldsymbol{y} >$,

(b) (1 Punkt) $< \boldsymbol{x}_1 + \boldsymbol{x}_2 \, , \, \mathbf{y} > = < \boldsymbol{x}_1 \, , \, \boldsymbol{y} > + < \boldsymbol{x}_2 \, , \, \boldsymbol{y} >$.

(c) (1 Punkt) $< \boldsymbol{x} \, , \, \boldsymbol{x} > = \; ||\boldsymbol{x}||^2$.

6. Aufgabe (4 Punkte) (*Eigenschaften von Vektoren*)

(a) (2 Punkte) Bestimmen Sie alle Vektoren in

$$\left\{ \mathbf{y} \in \mathbb{R}^3 \,\middle|\, \mathbf{y} = \begin{pmatrix} -1 & 2 & y \end{pmatrix}^T, \; y \in \mathbb{R} \right\},$$

die die Länge 3 haben.

(b) (2 Punkte) Welche(r) dieser Vektoren liegt/liegen für

$$\mathbf{X} = \begin{pmatrix} 4 \\ -8 \\ 8 \end{pmatrix}$$

in $\delta(\mathbf{X})$? Stellen Sie diese Vektoren als Linearkombinationen der $\boldsymbol{x}_i$ dar.

7. Aufgabe (8 Punkte) (*Wahrscheinlichkeits- und Dichtefunktion*)

(a) Bestimmen Sie für die reellen Zufallsvariablen X und Y mit den Verteilungsfunktionen $F_X(x)$, bzw. $F_Y(y)$ jeweils die Wahrscheinlichkeits- bzw. Dichtefunktion $f_X(x)$, bzw. $f_Y(y)$ und skizzieren Sie sowohl $F_X(x)$, bzw. $F_Y(y)$ als auch $f_X(x)$, bzw. $f_Y(y)$.

i. (2 Punkte) $F_X(x) = \frac{1}{1+\exp(-\frac{x-a}{b})}$ (Skizze mit $a = 0$, $b = 1$),

ii. (2 Punkte) $F_Y(y) = \begin{cases} 0 & \text{für } y < 3 \\ 0.7 & \text{für } 3 \leq y < 8 \\ 1 & \text{für } y \geq 8 \end{cases}$.

(b) (2 Punkte) Zeigen Sie, dass $f(x_1, x_2) = f(x_1)\, f(x_2)$ aus $F(x_1, x_2) = F(x_1, \infty)\, F(\infty, x_2)$ folgt. (Hinweis: $F(x_1) = F(x_1, \infty)$.)

(c) (2 Punkte) Bestimmen Sie für $\rho = 0$ die marginale Dichte von x_1 aus der gemeinsamen Dichtefunktion der Normalverteilung

$$f(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\left(\frac{x_1-\mu_1}{\sigma_1}\right)^2 - 2\rho\frac{x_1-\mu_1}{\sigma_1}\frac{x_2-\mu_2}{\sigma_2} + \left(\frac{x_2-\mu_2}{\sigma_2}\right)^2\right]\right\}.$$

(Hinweis: $f(x_1) = \int_{-\infty}^{\infty} f(x_1, z_2)\, d\, z_2$.)

8. Aufgabe (4 Punkte) (*Von der Dichte zur Verteilung und Erwartungswert*)

Betrachten Sie die Zufallsvariable X mit Dichtefunktion

$$f(x) = \begin{cases} 6\,x - 6\,x^2 & \text{für } 0 \leq x \leq 1 \\ 0 & \text{sonst} \end{cases}.$$

(a) (2 Punkte) Bestimmen und skizzieren Sie $F(x)$.
Verwenden Sie dabei $F(x) = \int_{-\infty}^{x} f(z)\, dz$.

(b) (2 Punkte) Bestimmen und skizzieren Sie $\mathrm{E}[X]$.
Verwenden Sie dabei $\mathrm{E}[X] = \int_{-\infty}^{\infty} x\, f(x)\, dx$.

9. Aufgabe (7 Punkte) (*Eigenschaften der Standardnormalverteilungsdichte*)

(a) (2 Punkte) Using $\phi(x) = (2\,\pi)^{-1/2} \exp\left(-\frac{1}{2}x^2\right)$ for the density $\phi(x)$ of the standard normal distribution, show that the derivative of $\phi(x)$ is the function $-x\,\phi(x)$, and that the second derivative is $(x^2 - 1)\,\phi(x)$.

(b) Use these facts to show

i. (2 Punkte) that the expectation of a standard normal random variable is 0, and

ii. (3 Punkte) that its variance is 1.

(Hinweis: Mit der Regel von l'Hospital gilt
$\frac{1}{\sqrt{2\,\pi}} \lim_{x\to\pm\infty} \frac{x}{\exp\left(\frac{1}{2}x^2\right)} = \frac{1}{\sqrt{2\,\pi}} \lim_{x\to\pm\infty} \frac{1}{x\,\exp\left(\frac{1}{2}x^2\right)} = 0$.)
[Quelle: Davidson / MacKinnon (2004), Exercise 1.6.]

Allgemeine Bemerkung:

Gleichungen der Form wie in a), in denen Funktionen mit Hilfe ihrer Ableitung gesucht sind, nennt man Differentialgleichungen (kurz: DGL). Die erste Gleichung $f'(x) = -xf(x)$ nennt man eine gewöhnliche (die Funktion hängt von einer Variablen ab) Differentialgleichung 1. Ordnung (nur die erste Ableitung kommt vor), die zweite Gleichung $f''(x) = (x^2 - 1)f(x)$ gewöhnliche Differentialgleichung 2. Ordnung.

10. Aufgabe (8 Punkte) (*Momente: Erwartungswert, Varianz, Kovarianz*)

Einige wichtige Rechenregeln für die Zufallsvariablen X, Y, Z, die Konstanten a, b, c und die Funktion h sind hier zu finden. Einen für die Aufgaben nötiger Auszug finden Sie im Folgenden:

- $\mathrm{E}[a + bX + cY] = a + b\,\mathrm{E}[X] + c\,\mathrm{E}[Y]$,
- $\mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}[X])^2] = \mathrm{E}[X^2] - \mathrm{E}[X]^2$,
- $\mathrm{Var}(a + bX) = \mathrm{Var}(bX) = b^2\mathrm{Var}(X)$,
- $\mathrm{Cov}(X, Y) = \mathrm{E}[(X - \mathrm{E}[X])(Y - \mathrm{E}[Y])] = \mathrm{E}[X\,Y] - \mathrm{E}[X]\,\mathrm{E}[Y]$,
- $\mathrm{E}[h(Y)\,X|Y] = h(Y)\,\mathrm{E}[X|Y]$,
- $\mathrm{E}[\mathrm{E}[X|Y]] = \mathrm{E}[X]$ (Gesetz der iterierten Erwartungen),
- $\mathrm{Var}(X|Y) = \mathrm{E}[(X - E[X|Y])^2|Y] = \mathrm{E}[X^2|Y] - \mathrm{E}[X|Y]^2$,
- $\mathrm{Cov}(X, Z|Y) = \mathrm{E}[(X - \mathrm{E}[X|Y])(Z - \mathrm{E}[Z|Y])|Y] = \mathrm{E}[X\,Z|Y] - \mathrm{E}[X|Y]\,\mathrm{E}[Z|Y]$.

(a) Zeigen Sie den Verschiebungssatz sowohl im unbedingten als auch im bedingten Fall, d.h. zeigen Sie, dass

i. (1 Punkt) $\mathrm{Var}(X) = \mathrm{E}[X^2] - \mathrm{E}[X]^2$ aus $\mathrm{Var}(X) = \mathrm{E}[(X - \mathrm{E}[X])^2]$ und

ii. (1 Punkt) $\mathrm{Var}(X|Y) = \mathrm{E}[X^2|Y] - \mathrm{E}[X|Y]^2$ aus $\mathrm{Var}(X|Y) = \mathrm{E}[(X - \mathrm{E}[X|Y])^2|Y]$ folgt.

(b) Zeigen Sie mit Hilfe obiger Rechenregeln:

i. (1 Punkt) $\mathrm{Cov}(b\,X,\, c\,Y) = b\,c\,\mathrm{Cov}(X,\, Y)$

ii. (2 Punkte) $\mathrm{Var}(b\,X + c\,Y) = b^2\,\mathrm{Var}(X) + c^2\,\mathrm{Var}(Y) + 2\,b\,c\,\mathrm{Cov}(X, Y)$

(c) (1 Punkt) Für die Zufallsvariable X gelte nun $\mathrm{E}[X] = 0$.
Zeigen Sie, dass dann gilt: $\mathrm{Var}(X) = \mathrm{E}[X^2]$ und $\mathrm{Cov}(X, Y) = \mathrm{E}[X\,Y]$.

(d) (2 Punkte) Zeigen Sie, dass aus $\mathrm{E}[X|Y] = 0$ auch $\mathrm{E}[X] = 0$ und $\mathrm{Cov}(X, Y) = 0$ folgt.

11. Aufgabe (3 Punkte) (*Interpretation des bedingten Erwartungswerts*)

Suppose that at a large university, college grade point average, GPA, SAT score, SAT, are related by the conditional expectation $\mathrm{E}[GPA|SAT] = 0.7 + 0.002\,SAT$.

(a) (1 Punkt) Find the expected GPA when $SAT = 800$. Find $\mathrm{E}[GPA|SAT = 1400]$. Comment on the difference.

(b) (1 Punkt) If the average SAT in the university is 1100, what is the average GPA?
(Hint: Use the Law of Iterated Expectations.)

(c) (1 Punkt) If a student's SAT score is 1100, does this mean that he or she will have the GPA found in part (b)? Explain.

12. Aufgabe (3 Punkte) (*Bedingter Erwartungswert und stochastische Unabhängigkeit*)

(a) (1 Punkt) If two random variables X_1 and X_2 are statistically independent, show that $\mathrm{E}[X_1|X_2] = \mathrm{E}[X_1]$.
(Hinweis: Rechnen Sie mit Dichten!)
[Quelle: Davidson / MacKinnon (2004), Exercise 1.9.]

(b) (2 Punkte) Finden Sie nun ein (empirisches) Beispiel für die beiden diskreten Zufallsvariablen X und Y, so dass $\mathrm{E}[X] = 0$, aber $\mathrm{E}[X|Y] \neq 0$ für alle Y gilt.

13. Aufgabe (3 Punkte) (*Multivariate Verteilungen*)

Gegeben sei der Zufallsvektor $\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N(\boldsymbol{\mu}, \boldsymbol{\Omega})$. Außerdem sei $\mathbf{x}^* = \mathbf{A}\mathbf{x}$ gegeben, wobei $\mathbf{A}$ eine nicht-stochastische (2×2)-Matrix sei.

(a) (1 Punkt) Geben Sie die Verteilung von $\mathbf{x}^*$ an.

(b) (1 Punkt) Im Folgenden sei $\mathbf{x}^* = \begin{pmatrix} x_1 - x_2 \\ x_1 + x_2 \end{pmatrix}$, $\boldsymbol{\mu} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ und $\boldsymbol{\Omega} = \begin{pmatrix} 1 & 0.7 \\ 0.7 & 2 \end{pmatrix}$.
Bestimmen Sie die Matrix $\mathbf{A}$, so dass $\mathbf{x}^* = \mathbf{A}\mathbf{x}$ gilt.

(c) (1 Punkt) Geben Sie die Verteilung von $\mathbf{x}^*$ nun mit konkreten Werten an.

14. Aufgabe (3 Punkte) (*Multivariate Verteilungen im Modell*)

Für eine Zufallsstichprobe sei das Modell

$$y_t = \beta_1 + \beta_2 x_t + u_t$$

mit $x_t \sim N(\mu_x, \sigma_x^2)$ und $u_t \sim N(0, \sigma^2)$ gegeben. Dabei seien x_t und u_t unabhängig (woraus $\mathrm{Cov}(x_t, u_t) = 0$ folgt).
Damit lässt sich zeigen, dass y_t normalverteilt ist (vgl. z.B. Davidson/MacKinnon (2004), S. 130ff) und dass y_t und x_t gemeinsam multivariat normalverteilt sind (vgl. z.B. Davidson/MacKinnon (2004), S. 132), also

$$\begin{pmatrix} y_t \\ x_t \end{pmatrix} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$$

gilt. Bestimmen Sie

(a) (1 Punkt) den Erwartungswertvektor $\boldsymbol{\mu} = \begin{pmatrix} \mu_y \\ \mu_x \end{pmatrix}$ und

(b) (2 Punkte) die Kovarianzmatrix $\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_y^2 & \sigma_{yx} \\ \sigma_{yx} & \sigma_x^2 \end{pmatrix}$.

15. Aufgabe (7 Punkte) (*Folgenkonvergenz*)

(a) (2 Punkte) Beweisen Sie mit der Definition, dass die Folge $(a_n)_{n\in\mathbb{N}} := \frac{1}{5n^2}$ konvergiert.

(b) (2 Punkte) Beweisen Sie mit der Definition, dass die Folge $(b_n)_{n\in\mathbb{N}} := (-1)^n$ divergiert.

(c) (1 Punkt) Konvergiert $(b_n + b_{n+1})_{n\in\mathbb{N}}$? Begründung.

(d) (2 Punkte) Zeigen Sie nun, dass man eine beliebige, konvergente Folge $(c_n)_{n\in\mathbb{N}}$ als Summe von Folgen $(x_n)_{n\in\mathbb{N}} + (y_n)_{n\in\mathbb{N}}$ schreiben kann, wobei mindestens eine divergiert.

16. Aufgabe (5 Punkte) (*Funktionenkonvergenz*)

Sei $f_n(x) := \frac{\sin(nx)}{\sqrt{n}}$ eine Funktionenfolge mit $x \in \mathbb{R}$.

(a) (2 Punkte) Zeigen Sie, dass $f_n(x)$ für alle $x \in \mathbb{R}$ punktweise konvergiert.

(b) (3 Punkte) Zeigen Sie, dass die Ableitung des Grenzwerts $f'(x)$ ungleich dem Grenzwerts der Ableitung der Funktionenfolge $\lim\limits_{n\to\infty} f'_n(x)$ ist.

Bemerkung: Obwohl Funktionenfolge und Grenzwert stetig sind und die Ableitung vom Grenzwert existiert, existieren die Ableitungen der Folgenglieder an keinem Punkt.

17. Aufgabe (3 Punkte) (*Konvergenz von korrelierten Zufallsvariablen*)

Es gelte $X \sim N(0, \sigma_x^2)$ und $Y \sim N(0, \sigma_y^2)$ mit $Corr(X, Y) = 0.5$. X und Y seien außerdem gemeinsam normalverteilt. Es sei Z_n eine Folge von Zufallsvariablen

$$Z_n = \frac{n + 1000}{n} X - \frac{3n^2 + 50}{n^2} Y, \quad n = 1, 2, \ldots.$$

Bestimmen Sie die Verteilung der Zufallsvariable, gegen die die Folge von Zufallsvariablen Z_n für n gegen unendlich in Wahrscheinlichkeit konvergiert.

18. Aufgabe (7 Punkte) (*Konvergenz von Zufallsvariablen*)

Untersuchen Sie das Konvergenzverhalten der Folgen A_n bis F_n von Zufallsvariablen, d.h. bestimmen Sie jeweils (falls möglich) den Wahrscheinlichkeitslimes und (oder nur) die Grenzverteilung. Dabei seien X, Y und Z voneinander unabhängige $N(0, 1)$-, $N(2, 9)$-, bzw. $N(0, 5)$-verteilte Zufallsvariablen und die Folge von Zufallsvariablen Y_i, $i = 1, \ldots, n$, sei $NID(2, 9)$-verteilt.

(a) (1 Punkt) $A_n = X + \frac{Y}{n}$,

(b) (1 Punkt) $B_n = X^2 + \frac{Y^2}{\sqrt{n}}$,

(c) (1 Punkt) $C_n = (-1)^n \cdot Z$,

(d) (1 Punkt) $D_n = (-1)^n \cdot Y$,

(e) (1 Punkt) $E_n = \frac{1}{n} \cdot (-1)^n \cdot Z$,

(f) (2 Punkte) $F_n = \frac{1}{n} \cdot X^2 + \frac{1}{9} \cdot (Y - 2)^2 + \left(\frac{1}{\sqrt{5}} \cdot Z\right)^2$.

19. Aufgabe (9 Punkte) (*Summe von zwei st. ab., normalverteilten Zufallsvariablen*)

Betrachten Sie die beiden Zufallsvariablen X_1 und X_2. Dabei sei $X_1 \sim N(0,1)$ und $X_2 = Z \cdot X_1$, wobei Z eine von X_1 unabhängige Zufallsvariable mit $\mathrm{P}(Z=1) = \frac{1}{2}$ und $\mathrm{P}(Z=-1) = \frac{1}{2}$ sei.

(a) (3 Punkte) Bestimmen Sie die Verteilung von X_2.

(b) (3 Punkte) Bestimmen Sie $\mathrm{Cov}(X_1, X_2)$.

(c) (3 Punkte) Geben Sie die Verteilung von $X = X_1 + X_2$ an.

Bemerkung:
[Teilaufgaben (a) und (b) nach Davidson / MacKinnon (2004), Exercise 1.14.]

20. Aufgabe (3 Punkte) (*Von der Standardnormal- zur χ^2-Verteilung*)

Es sei $z_i \sim NID(0,1)$ mit $i = 1, \ldots, m$. Dann ist $y = \sum_{i=1}^{m} z_i^2 \sim \chi^2(m)$.
Zeigen Sie, dass gilt

(a) (1 Punkt) $\mathrm{E}(y) = m$,

(b) (2 Punkte) $\mathrm{Var}(y) = 2m$.
Hinweis: Um die Varianz von y bestimmen zu können, benötigen Sie das Ergebnis $\mathrm{E}(z^4) = 3$ für standardnormalverteilte Zufallsvariablen z.
Dies kann dadurch gezeigt werden (hier nicht gefordert), dass die ersten vier Ableitungen der Dichtefunktion der Standardnormalverteilung, $\phi(z) = \frac{1}{\sqrt{2\pi}} \exp(-\frac{1}{2}z^2)$, bestimmt werden und damit anschließend $\mathrm{E}(z) = \int_{-\infty}^{\infty} z\phi(z)\mathrm{d}z$ und analog $\mathrm{E}(z^2)$, $\mathrm{E}(z^3)$ und $\mathrm{E}(z^4)$.

21. Aufgabe (9 Punkte) (*Basics mit R*)

Wiederholen Sie die Grundlagen zu `R` aus der Einführung.

(a) (3 Punkte) Plotten Sie die Wahrscheinlichkeits- bzw. Dichte- und Verteilungsfunktionen aus der Aufgabe zur Wahrscheinlichkeits- und Dichtefunktion. Verwenden Sie dabei für Teil (b) $a = 1$ und $b = 2$.

(b) (2 Punkte) Ziehen Sie für die Zufallsvariablen $x \sim \text{uniform}[0,5]$ und $u \sim \mathrm{N}(0,9)$ jeweils 100 Realisationen. Verwenden Sie dabei einen Randomseed von 42.

(c) (1 Punkt) Erzeugen Sie anschließend die Zufallsvariable $y = -2 + 1.4\,x + u$ und plotten Sie x gegen y.

(d) (2 Punkte) Schätzen Sie nun das Modell $y = \beta_1 + \beta_2\,x + u$ und zeichnen Sie die Regressionsgerade blau in Ihre Graphik ein.

(e) (1 Punkt) Programmieren Sie in `R` einen fairen Würfel.
Hinweis: `?sample`.

22. Aufgabe (8 Punkte) (*Visualisierung der Verteilungen*)

Plotten Sie die Dichten folgender Zufallsvariablen für einem geeigneten Bereich in R. Verwenden Sie dazu für jede Teilaufgabe ein gemeinsames Graphikfenster für die beiden Dichten.

Hinweise:

- Die Dichten aus Teil (a) können Sie z.B. mit folgendem R-Programmcode plotten

```
plot(function(x) dnorm(x, mean=0, sd=sqrt(1)), from=-4, to=8, lwd=2,
      ylab="Dichte", main = "Normalverteilung")
plot(function(x) dnorm(x, mean=2, sd=sqrt(5)), from=-4, to=8, lwd=2,
      col="blue", lty=2, add=TRUE)
legend("topright", inset=0.02, legend=c("N(0,1)", "N(2,5)"),
      col=c("black", "blue"), lwd=c(2,2), lty=c(1,2))
```

- Falls Sie die Befehle für die Verteilungen nicht wissen, versuchen Sie `?distributions`.

(a) (2 Punkte) $A_1 \sim N(0,1)$ und $A_2 \sim N(2,5)$ (Programmcode erläutern!),

(b) (2 Punkte) $B_1 \sim \chi^2(4)$ und $B_2 \sim \chi^2(20)$,

(c) (2 Punkte) $C_1 \sim t(30)$ und $C_2 \sim t(2)$,

(d) (2 Punkte) $D_1 \sim F(4,30)$ und $D_2 \sim F(4,2)$,

Verwenden Sie auch andere Verteilungsparameter, um sich die hier aufgeführten Zusammenhänge zwischen den unterschiedlichen Verteilungen klar zu machen.

23. Aufgabe (2 Punkte) (*Fehlspezifikation*)

Betrachten Sie das Modell

$$y_t = \beta_1 + \beta_2 x_t + \beta_3 z_t + v_t, \quad \mathrm{E}[v_t|x_t, z_t] = 0. \tag{1}$$

(a) (1 Punkt) Geben Sie $\mathrm{E}[y_t|x_t]$ an.

(b) (1 Punkt) Angenommen, Sie wollen folgendes Modell schätzen:

$$y_t = \beta_1 + \beta_2 x_t + u_t. \tag{2}$$

Wie sieht u_t in Termen von Modell (1) aus? Bestimmen Sie damit $\mathrm{E}[u_t|x_t]$.

24. Aufgabe (16 Punkte) (*DGP und Spezifikation*)

Im Folgenden bezeichnen y_t Gewinnrenditen einer beliebigen Anlageform. In der Literatur finden Sie dazu folgendes Modell

$$y_t \sim \theta_1 \cdot t(k) + \theta_2 \quad \forall t = 1, ..., n, \quad \theta_1, \theta_2 \in \mathbb{R} \tag{3}$$

wobei $t(k)$ eine t-verteilte Zufallsvariable mit $k \in \mathbb{N} \cup \{\infty\}$ Freiheitsgraden ist.

(a) (3 Punkte) Ist obiges Modell ein parametrisches Modell, ein semi-parametrisches Modell oder ein nicht-parametrisches Modell? Begründung. Nennen Sie auch Beispiele für die anderen Modelle. Gehen die Parameter linear ein, wenn θ_1 vorgegeben wird?

(b) (2 Punkte) Was ist ein DGP? Erklären Sie dieses Konzept mit eigenen Worten anhand eines Beispiels.

(c) (2 Punkte) Wann heißt ein Modell korrekt spezifiziert? Erklären Sie die Definition mit eigenen Worten anhand eines Beispiels.

(d) (2 Punkte) Was wird im Modell (3) spezifiziert? Begründung. Vergleichen Sie diese Spezifikation mit der des (bedingten) Erwartungswerts. Welche ist allgemeiner?

(e) (3 Punkte) Berechnen Sie den Erwartungswert $\mathrm{E}[y_t]$ und die Varianz $\mathrm{Var}\,(y_t)$. Welche Annahmen benötigen Sie für die Berechnung?

(f) (1 Punkt) Ist obiges Modell vollständig spezifiziert? Falls nicht, was müsste man ergänzen?

(g) (3 Punkte) Der wahre DGP sei nun

$$f_Y(y_t; \mu_0, \sigma_0^2) = \frac{1}{\sigma_0}\phi\left(\frac{y_t - \mu_0}{\sigma_0}\right) = \frac{1}{\sqrt{2\pi\sigma_0^2}} \exp\left(-\frac{1}{2}\frac{(y_t - \mu_0)^2}{\sigma_0^2}\right) \quad \text{mit } (\sigma_0, \mu_0) \in \mathbb{R}_+ \times \mathbb{R} \Longleftrightarrow y_t \sim N(\mu_0, \sigma_0^2)$$

Zeigen Sie, dass das Modell (3) korrekt spezifiert ist.

25. Aufgabe (19 Punkte) (*Eigenschaften von Schätzern und Schätzungen*)

Ihr lustiger Kommilitone entwickelt eine neue Schätzvorschrift für den Steigungsparameter β_1 folgender Regression:

$$y_t = \beta_1 x_t + u_t, \quad u_t|x_t \sim IID(0, \sigma^2), \quad t = 1, \ldots, n. \tag{4}$$

Hierfür **sortiert** er die Daten der x-Größe nach. Anschließend berechnet er folgenden Schätzer für β_1:

$$\tilde{\beta}_1 = \frac{y_n - y_1}{x_n - x_1} = \frac{1}{x_n - x_1} \cdot y_n + \frac{-1}{x_n - x_1} \cdot y_1. \tag{5}$$

Hinweis: $\mathbf{X} = (x_1, ..., x_n)^T$.

(a) (2 Punkte) Was ist ein Schätzer, was eine Schätzung? Fertigen Sie eine Skizze an, aus der die Idee dieses Schätzers (5) ersichtlich wird.

(b) (2 Punkte) Finden Sie eine Matrix $\boldsymbol{A}(\boldsymbol{X})$ nur abhängig von Werten in $\boldsymbol{X}$, so dass Sie $\tilde{\beta}_1(\boldsymbol{y}, \boldsymbol{X}) = \boldsymbol{A}(\boldsymbol{X})\boldsymbol{y}$ schreiben können.
Hinweis 1: Achten Sie auf die Dimensionen, insbesondere die von $\boldsymbol{A}$.
Hinweis 2: Man nennt einen solchen Schätzer **linear**, wenn **A nur** von $\boldsymbol{X}$ abhängt.

(c) (5 Punkte) Berechnen Sie $\mathrm{E}[\tilde{\beta}_1 \mid \boldsymbol{X}]$. Gilt $\mathrm{E}[\tilde{\beta}_1 \mid \boldsymbol{X}] = \mathrm{E}[\tilde{\beta}_1]$? Ist $\tilde{\beta}_1$ unverzerrt/erwartungstreu? Welche Annahmen benötigen Sie hierzu?
Hinweis: Gehen Sie davon aus, dass das Modell korrekt spezifiert ist, es also einen Parameter $\beta_{1,0}$ gibt, so dass folgendes gilt:

$$\mathrm{E}[y \mid \boldsymbol{X}] = \beta_{1,0}\, \mathrm{E}[x \mid \boldsymbol{X}] = \beta_{1,0} x$$

(d) (3 Punkte) Berechnen Sie $\mathrm{Var}\left(\tilde{\beta}_1 \mid \mathbf{X}\right)$.

(e) (1 Punkt) Bestimmen Sie $\mathrm{Var}\left(\hat{\beta}_1 \mid \boldsymbol{X}\right)$ im Modell (4), wobei $\hat{\beta}_1$ der KQ-Schätzer von β_1 ist.
Hinweis: $\mathrm{Var}\left(\hat{\beta}_1 \mid \mathbf{X}\right) = \sigma^2(\boldsymbol{X}^T\boldsymbol{X})^{-1}$.

(f) (4 Punkte) Welchen Schätzer wählen Sie, wenn Sie die Wahl zwischen dem KQ-Schätzer und dem Schätzer (5) haben und auch der KQ-Schätzer erwartungstreu ist?
Hinweis: Vergleichen Sie dazu beide Varianzen.

(g) (2 Punkte) Fertigen Sie eine neue Skizze an, in die Sie die Regressionsgeraden für $\hat{\beta}_1$ und $\tilde{\beta}_1$ einzeichnen. Falls Ihnen der Kleinst-Quadrate-Schätzer nicht mehr geläufig ist, so können Sie hier oder in der angegebenen Literatur Ihr Wissen auffrischen.

26. Aufgabe (31 Punkte) (*Einführung in Tests*)

Im Zuge des Erneuerbare-Energien-Gesetz (EEG), das im Jahr 2000 das bis dahin gültige Stromeinspeisungsgesetz von 1991 ablöste und die Einspeisevergütung von regenerativen Energien neu bewertete, soll nun die installierte Leistung von Wasserkraft bis 2000 und ab 2001 betrachtet werden. Eine Übersicht finden Sie in der Tabelle im Anhang.

Im Folgenden sollen Sie nun Schritt für Schritt einen statistischen Test entwickeln, um eine Aussage über folgende Hypothese zu treffen:

„Durch das EEG veränderte sich die Volatilität der angeschlossenen Wasserkraftleistung.“

Betrachten Sie vorerst folgende Abstrahierung: Gegeben seien $T_I, T_{II} \in \mathbb{N}$ mit $T_I < T_{II}$ und

$$\begin{aligned} &y_t \sim NID(0, \sigma_I^2) \text{ für } t = 1, ..., T_I,\\ &y_t \sim NID(0, \sigma_{II}^2) \text{ für } t = T_I + 1, ..., T_{II} \text{ und}\\ &y_t \text{ stochastisch unabhängig für } t = 1, ..., T_{II}. \end{aligned}$$

Man bezeichne mit $\boldsymbol{y}^T = \begin{pmatrix} y_1 & \dots & y_{T_I} \end{pmatrix}^T \in \mathbb{R}^{T_I}$ und $\boldsymbol{y}'^T = \begin{pmatrix} y_{T_I+1} & \dots & y_{T_{II}} \end{pmatrix}^T \in \mathbb{R}^{T_{II}-T_I}$ die zwei Teilstichproben.

Allgemeine Hinweise:

- Falls $\sigma_I^2 \neq \sigma_{II}^2$ liegt ein einfacher Fall von Heteroskedastie vor.
- Der zu entwickelnde Test ist ein einfacher Strukturbruchtest.

(a) (1 Punkt) Formulieren Sie oben genannte Hypothese als statistischen Hypothesentest.

(b) (3 Punkte) Geben Sie die Verteilungen von $X := \sum_{i=1}^{T_I} y_i$ bzw. $X' := \frac{1}{T_I}\sum_{i=1}^{T_I} y_i$ und $Y := \sum_{j=T_I+1}^{T_{II}} y_j$ bzw. $Y' := \frac{1}{T_{II}-T_I}\sum_{j=T_I+1}^{T_{II}} y_j$ an. Begründen Sie Ihr Vorgehen.

(c) (2 Punkte) Um die Verteilung Ihrer Teststatistik herleiten zu können, müssen Sie zeigen, dass stetige Funktionen stochastisch unabhängiger Zufallsvariablen wieder stochastisch unabhängig sind. Überlegen Sie sich also in eigenen Worten, dass für (univariate) unabhängige X und Y die Zufallsvariablen $X' := g(X)$ und $Y' := h(Y)$ für g, h stetig wieder stochastisch unabhängig sind. Was heißt dies für $\mathrm{E}[u_t^2 u_s^2]$, wobei $u_t \sim IID(0, \sigma_u^2)$.
Hinweis: Dieses Ergebnis wird als „Blockungslemma“ geführt; der Beweis basiert auf der Unabhängigkeit der zugrundeliegenden Sigma-Algebren und wird hier nicht erwartet.

(d) (3 Punkte) Falls man die Varianzen σ_I^2 und σ_{II}^2 wüsste, wie müsste man $\sum_{i=1}^{T_I} y_i^2$ vormultiplizieren, um eine bekannte Verteilung zu bekommen? Welche ist das?

(e) (1 Punkt) Für die Teststatistik benötigen Sie einen Schätzer für die Varianz. Wie sieht der unkorrigierte Schätzer für die Varianzen in unserem Fall aus?

(f) (5 Punkte) Leiten Sie nun die Verteilung der Teststatistik unter H_0 her. Nutzen Sie dabei die Varianzschätzer und ihre bisherigen Ergebnisse.
Hinweise: Beachten Sie die stochastische Unabhängigkeit mit Hilfe der Teilaufgabe (c).

(g) Im Anhang finden Sie eine Tabelle mit Werten für die installierte Wasserkraftleistung nach Jahren, die unter anderem auch für die beiden Stichproben $\boldsymbol{y}$ und $\boldsymbol{y}'$ jeweils mittelwertbereinigt angegeben sind.

i. (1 Punkt) Welche Werte weisen T_I und T_{II} auf?

ii. (2 Punkte) Berechnen Sie die Teststatistik auf Basis Ihrer Nullhypothese.

iii. (3 Punkte) Zeichnen Sie mit `R` die Verteilung der Teststatistik unter H_0. Markieren Sie in der Graphik den Wert Ihrer Teststatistik.
Hinweise:
- Benutzen Sie den Befehl der vorangegangen Aufgabe.
- Befehl für das Einzeichnen der Teststatistik: `?abline` oder `?points`.

iv. (3 Punkte) Erläutern Sie anhand der vorher erstellten Graphik den Zusammenhang zwischen Signifikanzniveau und kritischen Werten.

v. (3 Punkte) Wählen Sie zwei gängige Signifkanzniveaus, berechnen Sie die kritischen Werte, einerseits exakt mit `R`, andererseits approximativ anhand einer Tabelle und geben Sie die Ablehnungsbereiche, die Nicht-Ablehnungsbereiche und die Testentscheidungen an.

vi. (2 Punkte) Wie lautet Ihre Entscheidung? Würden Sie sagen, dass aufgrund Ihrer statistischen Testentscheidung sich die Volatilität veränderte? Begründung.

vii. (2 Punkte) Sehen Sie sich nochmal den Datensatz an und überlegen Sie eine Möglichkeit, wie man besser mit dieser Problemstellung umgehen könnte.

Anhang:

Jahr	Wasserkraft (MW)	Bereinigt
1992	3.550	0.004
1993	3.509	-0.037
1994	3.563	0.017
1995	3.595	0.049
1996	3.510	-0.036
1997	3.525	-0.021
1998	3.601	0.055
1999	3.523	-0.023
2000	3.538	-0.008
2001	3.538	-0.510
2002	3.785	-0.263
2003	3.934	-0.114
2004	3.819	-0.229
2005	4.115	0.067
2006	4.083	0.035
2007	4.169	0.121
2008	4.138	0.090
2009	4.151	0.103
2010	4.395	0.347
2011	4.401	0.353

27. Aufgabe (7 Punkte) (*Modellschreibweisen und der Kleinst-Quadrate-Schätzer als Momentenschätzer*)
Gegeben sei das Modell

$$y_t = \beta_1 + \beta_2\, x_t + u_t = \mathbf{X}_t\boldsymbol{\beta} + u_t, \quad t = 1, \ldots, n,$$

wobei $\mathrm{E}[u_t|x_t] = 0$ gelte.
Außerdem sei $\mathrm{E}[\mathbf{X}_t^T\mathbf{X}_t]$ invertierbar und die Elemente von $\mathrm{E}[\mathbf{X}_t^T\mathbf{X}_t]$ sowie $\mathrm{E}[\mathbf{X}_t^T y_t]$ lassen sich gemäß einem Gesetz der Großen Zahlen (LLN) durch die entsprechenden arithmetischen Mittel aus der Stichprobe schätzen. Dann gilt

$$\widehat{\mathrm{E}}[\mathbf{X}_t^T\mathbf{X}_t] = \begin{pmatrix} \widehat{\mathrm{E}}[1] & \widehat{\mathrm{E}}[x_t] \\ \widehat{\mathrm{E}}[x_t] & \widehat{\mathrm{E}}[x_t^2] \end{pmatrix} = \begin{pmatrix} 1 & \frac{1}{n}\sum_{t=1}^n x_t \\ \frac{1}{n}\sum_{t=1}^n x_t & \frac{1}{n}\sum_{t=1}^n x_t^2 \end{pmatrix}$$

und

$$\widehat{\mathrm{E}}[\mathbf{X}_t^T y_t] = \begin{pmatrix} \widehat{\mathrm{E}}[y_t] \\ \widehat{\mathrm{E}}[x_t y_t] \end{pmatrix} = \begin{pmatrix} \frac{1}{n}\sum_{t=1}^n y_t \\ \frac{1}{n}\sum_{t=1}^n x_t y_t \end{pmatrix},$$

wobei die Matrix $\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 \\ \vdots \\ \mathbf{X}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_1 & \cdots & \boldsymbol{x}_k \end{pmatrix}$ und der Vektor $\mathbf{y} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ gegeben sind.

(a) (1 Punkt) Interpretieren Sie mit eigenen Worten die Objekte $\boldsymbol{X}_t$ für $t = 1, ..., n$ und $\boldsymbol{x}_j$ für $j = 1, ..., k$. Wie sieht in einem Modell mit Konstante die erste Spalte von $\boldsymbol{X}$ aus?

(b) (1 Punkt) Berechnen Sie $\mathbf{X}^T\mathbf{X}$ als $\sum_{t=1}^n \boldsymbol{X}_t^T\boldsymbol{X}_t$.

(c) (2 Punkte) Zeigen Sie $\mathbf{X}^T\mathbf{X} = n\,\widehat{\mathrm{E}}[\mathbf{X}_t^T\mathbf{X}_t]$ und $\mathbf{X}^T\mathbf{y} = n\,\widehat{\mathrm{E}}[\mathbf{X}_t^T y_t]$ für den gegebenen Fall.

(d) (2 Punkte) Leiten Sie den KQ-Schätzer $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}$ als Momentenschätzer ab.

(e) (1 Punkt) Welche Annahme bezüglich der Regressormatrix $\mathbf{X}$ benötigen Sie in Teil (d) außerdem? Was bedeutet diese hier für die Realisationen von $x_1, \ldots, x_n$?

28. Aufgabe (3 Punkte) (*Partitionierte Matrizen*)
Betrachten Sie die beiden Matrizen

$$\mathbf{X}_1 = \begin{pmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_{k_1} \end{pmatrix} \quad \text{und} \quad \mathbf{X}_2 = \begin{pmatrix} \mathbf{x}_{k_1+1} & \mathbf{x}_{k_1+2} & \cdots & \mathbf{x}_k \end{pmatrix},$$

wobei die $(n \times 1)$-Vektoren $\mathbf{x}_1, \ldots, \mathbf{x}_{k_1}$ bzw. $\mathbf{x}_{k_1+1}, \ldots, \mathbf{x}_k$ die Spalten der beiden Matrizen $\mathbf{X}_1$ und $\mathbf{X}_2$ sind.

(a) (2 Punkte) Bilden Sie $\mathbf{X}_1^T\mathbf{X}_2$, indem Sie die Einträge von $\mathbf{X}_1^T\mathbf{X}_2$ jeweils als Skalarprodukt aus geeigneten Spalten von $\mathbf{X}_1$ und $\mathbf{X}_2$ darstellen.

(b) (1 Punkt) Wann ist $\mathbf{X}_1^T\mathbf{X}_2 = \mathbf{0}$? Ist die Bedingung dafür eine genau-dann-wenn-Bedingung? Begründung.

29. Aufgabe (13 Punkte) (*Geometrische Eigenschaften einer Kleinst-Quadrate-Schätzung*)
Betrachten Sie das Modell

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

wobei die $(n \times k)$-Matrix $\mathbf{X}$ vollen Rang k habe.
Der OLS-Schätzer für $\boldsymbol{\beta}$ lautet dann

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y}.$$

Als Schätzungen für $\mathbf{y}$ und $\mathbf{u}$ ergeben sich damit

$$\hat{\mathbf{y}} = \mathbf{X}\hat{\boldsymbol{\beta}} = \underbrace{\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T}_{=:\mathbf{P}_\mathbf{X}}\mathbf{y} = \mathbf{P}_\mathbf{X}\mathbf{y},$$
$$\hat{\mathbf{u}} = \mathbf{y} - \hat{\mathbf{y}} = \mathbf{y} - \mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{y} - \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{y} = (\underbrace{\mathbf{I}_n - \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T}_{=:\mathbf{M}_\mathbf{X}})\mathbf{y} = \mathbf{M}_\mathbf{X}\mathbf{y}.$$

Obige Matrizen $\mathbf{P}_\mathbf{X} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ und $\mathbf{M}_\mathbf{X} = \mathbf{I} - \mathbf{P}_\mathbf{X} = \mathbf{I} - \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T$ werden als Projektionsmatrizen bezeichnet (s. Skript).

(a) (2 Punkte) Was bedeutet der Rang einer Matrix in Bezug auf $\mathbf{X}$ aus obigem Modell?

(b) (1 Punkt) Geben Sie die Dimensionen der Matrizen $\mathbf{P}_\mathbf{X}$ und $\mathbf{M}_\mathbf{X}$ an.

(c) (3 Punkte) Zeigen Sie, dass $\mathbf{P}_\mathbf{X}$ und $\mathbf{M}_\mathbf{X}$ symmetrisch und idempotent sind.

(d) (3 Punkte) Verwenden Sie die Ergebnisse aus Teilaufgabe (c), um zu zeigen, dass die Matrizen $\mathbf{P}_\mathbf{X}$ und $\mathbf{M}_\mathbf{X}$ positiv semi-definit sind.

(e) Zeigen Sie außerdem, dass für $\mathbf{P}_\mathbf{X}$ und $\mathbf{M}_\mathbf{X}$ gilt:

i. (1 Punkt) $\mathbf{P}_\mathbf{X}\mathbf{X} = \mathbf{X}$,

ii. (1 Punkt) $\mathbf{M}_\mathbf{X}\mathbf{X} = \mathbf{0}$,

iii. (1 Punkt) $\mathbf{P}_\mathbf{X}\mathbf{M}_\mathbf{X} = \mathbf{M}_\mathbf{X}\mathbf{P}_\mathbf{X} = \mathbf{0}$,

iv. (1 Punkt) $\hat{\mathbf{u}} = \mathbf{M}_\mathbf{X}\mathbf{y} = \mathbf{M}_\mathbf{X}\mathbf{u}$.

30. Aufgabe (3 Punkte) (*Verteilung der quadrierten Residuen*)
Betrachten Sie das Modell

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad \mathbf{u} \sim N(\mathbf{0}, \sigma^2\mathbf{I}),$$

wobei die $(n \times k)$-Matrix $\mathbf{X}$ vollen Rang k habe. Zeigen Sie

$$\frac{\hat{\mathbf{u}}^T\hat{\mathbf{u}}}{\sigma^2} \sim \chi^2(n-k).$$

Gehen Sie dabei in folgenden Schritten vor:

(a) (1 Punkt) Bestimmen Sie zunächst die Verteilung von $\frac{\mathbf{u}}{\sigma}$.

(b) (1 Punkt) Wenden Sie nun entsprechenden Satz im Skript mit $\mathbf{z} := \frac{\mathbf{u}}{\sigma}$ und $\mathbf{P} := \mathbf{M}_\mathbf{X}$ an.

(c) (1 Punkt) Fassen Sie anschließend Ihre Ergebnisse zu $\frac{\hat{\mathbf{u}}^T\hat{\mathbf{u}}}{\sigma^2} \sim \chi^2(n-k)$ zusammen.

31. Aufgabe (26 Punkte) (*Projektionsmatrizen und Skalarprodukt in einer ersten Schätzung*)
Betrachten Sie das Modell

$$\log(price) = \beta_1 + \beta_2 area + \beta_3 age + \beta_4 rooms + u,$$

das die Einflussgrößen auf Hauspreise in North Andover, Massachusetts, untersucht. Dabei bezeichne *price* den Kaufpreis in US$, *area* die Grundfläche des Hauses in Quadrat-Fuß, *age* das Alter des Hauses in Jahren und *rooms* die Anzahl der Zimmer im Haus.

(a) (2 Punkte) Importieren Sie den Datensatz hprice3.txt in `R`. Überprüfen Sie mit dem Befehl `names`, welche Variablen im Datensatz hprice3.txt enthalten sind. Wie viele Beobachtungen liegen vor? Schätzen Sie die Parameter aus obigem Modell mit OLS (`summary` verwenden). [Anmerkung zum Datensatz: hprice3.txt enthält einen Auszug aus dem Datensatz hprice3.wf1 aus Wooldridge (2009).]

(b) (4 Punkte) Wenden Sie den `plot`-Befehl auf Ihr `lm`-Objekt an. Sie erhalten damit 5 Graphen, welche Sie in der Konsole weiterklicken können. Interpretieren Sie insbesondere den *QQ*-Plot (`plot(lm-object, which=2)`) und den *Scale−Location*-Plot (`plot(lm-object, which=3)`) in Anbetracht der Annahmen B4 und B2b.

(c) (5 Punkte) Erläutern und interpretieren Sie **alle** Größen des OLS-Outputs, gehen Sie insbesondere auf den Parameter der Grundfläche ein und interpretieren Sie diesen in Termen von Quadratmeter (vgl. Erklärung eines R Output).

(d) (4 Punkte) Plotten Sie *area* gegen $\log(price)$. Zeichnen Sie außerdem die geschätzte Regressionsgerade ein. Wählen Sie dazu für *age* und *rooms* geeignete Werte (*age_star* und *rooms_star*) aus und begründen Sie Ihre Wahl.
Gehen Sie dabei in folgenden Schritten vor:

- Skizzieren Sie sich zunächst die Graphik, die Sie in `R` plotten möchten.
- Plotten Sie diese Graphik nun in `R`.
  [Hinweis: Sie können beispielsweise die folgenden Programmzeilen verwenden:

```
laerchennadelgruen      <- rgb(174/255, 167/255, 0)
intc        <- c(1,age_star,rooms_star) %*% ols$coef[c(1,3,4),1]
slpe        <- ols\$coef[2,1]
plot(area, log(price), main="Plot von area vs. log(price)", lwd=2, col="grey")
abline(c(intc, slpe), col=laerchennadelgruen, lwd=2)
legend("topleft", inset=0.02,
      legend=c("Beobachtungen", paste("bed. OLS-Gerade für age=",
         round(age_star,2), ", rooms=", round(rooms_star,2), sep="")),
      lty=c(-1,1), lwd=c(2,2), pch=c(1,-1),
      col=c("grey", laerchennadelgruen))
```

  Dabei bezeichne `ols` die `summary`-Ergebnisse aus Teil (a) und `age_star` $= age^*$ bzw. `rooms_star` $= rooms^*$ die von Ihnen ausgewählten Werte.
  Machen Sie sich klar, was die einzelnen Befehle (außer `rgb`) bedeuten. Lesen Sie dies ggf. nach oder finden Sie die Bedeutung durch Ausprobieren heraus.]

Für die nächsten Aufgaben ist es sinnvoll, sich mit dem Skalarprodukt und der Norm eines Vektors vertraut zu machen.
Gegeben seien nun das allgemeine Modell

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

sowie die Residuen $\hat{\mathbf{u}}$ und die gefitteten Werte $\hat{\mathbf{y}}$ aus einer OLS-Schätzung des Modells.

(e) (2 Punkte) Bestimmen Sie $< \hat{\mathbf{y}} \, , \, \hat{\mathbf{u}} >$.
[Hinweis: Verwenden Sie die Projektionsmatrizen $\mathbf{P_X}$ und $\mathbf{M_X}$.]

(f) (1 Punkt) Zeigen Sie $< \hat{\mathbf{y}} \, , \, \mathbf{y} > = \|\hat{\mathbf{y}}\|^2$.

(g) (2 Punkte) Zeigen Sie, dass $\frac{\|\hat{\mathbf{y}}\|}{\|\mathbf{y}\|} = \cos(\theta)$ gilt, wobei θ den Winkel zwischen $\hat{\mathbf{y}}$ und $\mathbf{y}$ angibt.
[Hinweis: Schreiben Sie $< \hat{\mathbf{y}} \, , \, \mathbf{y} >$ in Abhängigkeit von θ und verwenden Sie Teil (b).]

(h) (2 Punkte) Erläutern Sie anhand einer geeigneten Graphik (für $n = 2$ Beobachtungen), dass $\cos(\theta)$ aus Teil (c) nicht negativ sein kann. Was bedeutet dies für θ?

(i) (2 Punkte) Welche Spezialfälle ergeben sich für $\cos(\theta) = 1$ und $\cos(\theta) = 0$? Zeichnen Sie auch hierfür jeweils eine Graphik, die die Vektoren $\hat{\mathbf{y}}$ und $\hat{\mathbf{u}}$ enthält.

(j) (2 Punkte) Bestimmen Sie in `R` sowohl $\cos(\theta)$ als auch den Winkel θ für die OLS-Schätzung von

$$\log(price) = \beta_1 + \beta_2 area + \beta_3 age + \beta_4 rooms + u$$

und interpretieren Sie das Ergebnis.

32. Aufgabe (11 Punkte) (*FWL und Projektionsmatrizen*)

Gegeben sei das folgende Modell

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{u}.$$

Dabei seien $\mathbf{y}$ und $\mathbf{u}$ jeweils $(n \times 1)$-Vektoren. $\mathbf{X}$ sei eine $(n \times k)$-Matrix, $\mathbf{X}_1$ eine $(n \times k_1)$-Matrix und $\mathbf{X}_2$ eine $(n \times k_2)$-Matrix mit $k = k_1 + k_2$.

Die Matrix $\mathbf{X}$ ist also partitioniert in

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{X}_2 \end{pmatrix}.$$

Weiter seien $\mathbf{P_X}$, $\mathbf{P}_{\mathbf{X}_1} = \mathbf{P}_1$ und $\mathbf{P}_{\mathbf{X}_2} = \mathbf{P}_2$ bzw. $\mathbf{M_X}$, $\mathbf{M}_{\mathbf{X}_1} = \mathbf{M}_1$ und $\mathbf{M}_{\mathbf{X}_2} = \mathbf{M}_2$ die Projektionsmatrizen, die in die von $\mathbf{X}$, $\mathbf{X}_1$ und $\mathbf{X}_2$ aufgespannten Räume bzw. in die jeweils orthogonalen Unterräume projizieren.

Zeigen Sie sowohl mittels Berechnung als auch mittels Argumentation:

(a) (2 Punkte) $\mathbf{P_X}\mathbf{X}_1 = \mathbf{X}_1$
(Hinweis: Schreiben Sie $\mathbf{P_X}\mathbf{X}$ als $\mathbf{P_X}\begin{pmatrix} \mathbf{X}_1 & \mathbf{X}_2 \end{pmatrix}$),

(b) (2 Punkte) $\mathbf{P_X}\mathbf{P}_1 = \mathbf{P}_1\mathbf{P_X} = \mathbf{P}_1$,

(c) (2 Punkte) $\mathbf{M_X}\mathbf{M}_1 = \mathbf{M}_1\mathbf{M_X} = \mathbf{M_X}$.

Bestimmen Sie mittels Berechnung die folgenden Produkte und gehen Sie dabei auch auf mögliche Spezialfälle ein (z.B. wann ist die Matrix oder Teile davon $\mathbf{0}$):

(d) (1 Punkt) $\mathbf{P}_1\mathbf{X}_2$,

(e) (1 Punkt) $\mathbf{M}_1\mathbf{X}_2$,

(f) (1 Punkt) $\mathbf{P}_1\mathbf{X}$,

(g) (1 Punkt) $\mathbf{M}_1\mathbf{X}$,

(h) (1 Punkt) $\mathbf{M_X}\mathbf{X}_1$.

33. Aufgabe (23 Punkte) (*FWL mit verschiedenen Schätzern*)

Betrachten Sie folgendes Regressionsmodell

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta}_1 + \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{u} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u},$$

wobei $\mathbf{y}$ ein $(n \times 1)$-Vektor, $\mathbf{X}$ eine $(n \times k)$-Matrix, $\mathbf{X}_1$ eine $(n \times k_1)$-Matrix und $\mathbf{X}_2$ eine $(n \times k_2)$-Matrix mit $k = k_1 + k_2$ ist. Die Vektoren $\hat{\boldsymbol{\beta}}$, bzw. $\hat{\boldsymbol{\beta}}_1$ und $\hat{\boldsymbol{\beta}}_2$ bezeichnen die OLS-Schätzer für diese Regression. Die Matrix $\mathbf{P_X}$ (bzw. $\mathbf{P}_1$) projiziert orthogonal auf den von $\mathbf{X}$ (bzw. $\mathbf{X}_1$) aufgespannten Unterraum und $\mathbf{M_X} = \mathbf{I} - \mathbf{P_X}$ (bzw. $\mathbf{M}_1 = \mathbf{I} - \mathbf{P}_1$).

(a) Für welche der folgenden Regressionen ist die OLS-Schätzung für $\boldsymbol{\beta}_2$ numerisch identisch mit der Schätzung obigen Modells? Rufen Sie sich dabei die Aussage des FWL-Theorems in Erinnerung.
Hinweis:
   - Ermitteln Sie die jeweiligen Schätzer für $\boldsymbol{\beta}_2$ und vergleichen Sie diese mit $\hat{\boldsymbol{\beta}}_2$ aus obigem Modell.
   - Allgemeine Schreibweise des KQ-Schätzers $\hat{\boldsymbol{\alpha}}$ des Parameters $\boldsymbol{\alpha}$ im Modell:

   $$\boldsymbol{B}\boldsymbol{z} = \boldsymbol{A}\boldsymbol{\alpha} + \epsilon : \hat{\boldsymbol{\alpha}} = (\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T\boldsymbol{B}\boldsymbol{z}$$

   - Verbindung der Modelle: $\boldsymbol{y} = \hat{\boldsymbol{y}} + \hat{\boldsymbol{u}} = \boldsymbol{X}_1\hat{\boldsymbol{\beta}}_1 + \boldsymbol{X}_2\hat{\boldsymbol{\beta}}_2 + \hat{\boldsymbol{u}}$.

   i. (2 Punkte) $\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{u}$.
   Wann gilt $\hat{\boldsymbol{\beta}}_2^{(\mathrm{i.})} = \hat{\boldsymbol{\beta}}_2$?

   ii. (2 Punkte) $\mathbf{P}_2\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{u}$.
   Bestimmen Sie außerdem den Residuenvektor $\hat{\mathbf{u}}^{(\mathrm{ii.})}$.

   iii. (3 Punkte) $\mathbf{y} = \mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{u}$
   (FWL-Theorem, vgl. Davidson/MacKinnon, S. 68f).
   Warum ist $\hat{\boldsymbol{\beta}}_2^{(\mathrm{iii.})}$ gleich dem Schätzer für $\boldsymbol{\beta}_2$ aus dem Modell $\mathbf{M}_1\mathbf{y} = \mathbf{M}_1\mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{M}_1\mathbf{u}$?

   iv. (1 Punkt) $\mathbf{M}_2\mathbf{y} = \mathbf{X}_2\boldsymbol{\beta}_2 + \mathbf{u}$.
   Liegen die gefitteten Werte $\hat{\mathbf{y}}^{(\mathrm{iv.})}$ in $\delta(\mathbf{X})$?

(b) Im Folgenden liege das Modell

$$\log(wage) = \beta_1 + \beta_2\ hours + \beta_3\ exper + \beta_4\ age + \beta_5\ educ + u \tag{6}$$

vor. Verwenden Sie für Ihre Schätzungen `R` und den Datensatz wage2.txt.

   i. (1 Punkt) Schätzen Sie das Modell mit OLS und speichern Sie die Residuen und die gefitteten Werte.

   ii. (1 Punkt) Regressieren Sie die Residuen auf die gefitteten Werte (ohne Konstante). Regressieren Sie umgekehrt die gefitteten Werte auf die Residuen (ohne Konstante). Was fällt Ihnen an den Ergebnissen auf?

   iii. (2 Punkte) Stellen Sie die Schätzungen aus Teilaufgabe (b ii) mit Hilfe der Projektionsmatrizen $\mathbf{P_X}$ und $\mathbf{M_X}$ dar und erklären Sie dadurch die Ergebnisse aus Teilaufgabe (b ii).

iv. (1 Punkt) Führen Sie nun die Schätzung der Modelle

$$\log(wage) = \alpha_1 + \alpha_2 \ hours + \alpha_3 \ exper + \alpha_4 \ age + v \tag{7}$$

und

$$educ = \gamma_1 + \gamma_2 \ hours + \gamma_3 \ exper + \gamma_4 \ age + w \tag{8}$$

durch und speichern Sie dabei jeweils die Residuen.

v. (2 Punkte) Stellen Sie die Residuen der Modelle (7) und (8) mit Hilfe von Projektionsmatrizen dar. Partitionieren Sie die gesamte Regressormatrix $\mathbf{X}$ dazu geeignet.

vi. (2 Punkte) Erläutern Sie anhand der Regression der Residuen von (7) auf die von (8) die Bedeutung des Frisch-Waugh-Lovell-Theorems.

(c) (3 Punkte) Belegen Sie nun Ihre Ergebnisse aus Teil (a) empirisch. Partitionieren Sie dazu die Matrix $\mathbf{X}$ wieder so, dass $\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 & \mathbf{X}_2 \end{pmatrix}$, wobei die Matrix $\mathbf{X}_1$ die Konstante, *hours*, *exper* und *age* enthalte und die Matrix $\mathbf{X}_2$ lediglich *educ*.

(d) (3 Punkte) Wandeln Sie die Dauer der Ausbildung *educ* von Jahren in Semester um (*educ_s*).

Schreiben Sie Modell (6) in Matrixschreibweise ($\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}$) und finden Sie die Matrix $\mathbf{A}$, die genau der genannten Transformation entspricht.

Wie ändert sich der Vektor $\hat{\boldsymbol{\beta}}$ im Vergleich zu Teil (b i)? Überprüfen Sie Ihre Vermutung indem Sie Modell (6) mit der neuen Variable *educ_s* schätzen und interpretieren Sie die Zeile zu *educ_s* im Output.

34. Aufgabe (2 Punkte) (*Projektionsmatrizen und Bestimmtheitsmaß*)

Gehen Sie von folgendem Modell aus

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

Dabei enthalte $\mathbf{X}$ eine Konstante.

(a) (1 Punkt) Zeigen Sie, dass sich das Bestimmtheitsmaß $R^2 = \frac{SSE}{SST} = \frac{\sum_{t=1}^n (\hat{y}_t - \bar{y})^2}{\sum_{t=1}^n (y_t - \bar{y})^2}$ schreiben lässt als

$$R^2 = \frac{\|\mathbf{P_X M_\iota y}\|^2}{\|\mathbf{M_\iota y}\|^2}.$$

(b) (1 Punkt) Zeigen Sie außerdem, dass sich $R^2 = 1 - \frac{SSR}{SST} = 1 - \frac{\sum_{t=1}^n \hat{u}_t^2}{\sum_{t=1}^n (y_t - \bar{y})^2}$ auch schreiben lässt als

$$R^2 = 1 - \frac{\|\mathbf{M_X y}\|^2}{\|\mathbf{M_\iota y}\|^2}.$$

35. Aufgabe (16 Punkte) (*Einfache Probleme mit R: Dummies, Korrelationen, ...*)

Verwenden Sie für diese Aufgabe den Datensatz hprice3.txt (Auszug von hprice3.wf1 aus Wooldridge (2009)). Die darin enthaltenen Variablen sind *price* (Kaufpreis des Hauses in US$), *age* (Alter des Hauses in Jahren), *area* (Grundfläche des Hauses in Quadrat-Fuß) und *rooms* (Anzahl der Räume im Haus).

(a) (1 Punkt) Bilden Sie aus der Variable *rooms* die Dummyvariable *manyr*, die für Häuser mit mindestens 7 Räumen den Wert 1 annimmt und ansonsten 0.
Hinweis: Logische Operatoren in `R`.

(b) (1 Punkt) Schätzen Sie das Modell

$$price_t = \beta_1 + \beta_2\, age_t + \beta_3\, age_t^2 + \beta_4\, area_t + \beta_5\, manyr_t + \beta_6\, area_t \cdot manyr_t + u_t \qquad (9)$$

mit KQ und geben Sie den `summary`-Output an.
Speichern Sie den Vektor der geschätzten Parameter als Objekt `betas` ab.
(Hinweis: Funktionen aus mehreren Regressoren können in `R` in der `lm`-Schätzung nicht direkt eingegeben werden (z.B. "`+`" wird zur Trennung von Regressoren verwendet), sondern müssen innerhalb von `I()` übergeben werden, vgl. `?formula` und `?AsIs`.)

(c) (2 Punkte) Welche Annahmen müssen bezüglich Modell (9) gelten, damit der KQ-Schätzer für $\boldsymbol{\beta}^T = \begin{pmatrix}\beta_1 & \ldots & \beta_6\end{pmatrix}$ erwartungstreu ist?

(d) (1 Punkt) Wie müsste die Stichprobe aussehen, damit $\hat{\boldsymbol{\beta}}$ nicht berechnet werden kann?
Hinweis: Denken Sie an `manyr`.

(e) (1 Punkt) Was bedeutet es im gegebenen Beispiel, wenn $area_t$, age_t und $manyr_t$ vorherbestimmt sind?

(f) (2 Punkte) Interpretieren Sie für Ihre Schätzung den Einfluss des Alters auf den Hauspreis. Bei welchem Alter eines Hauses ist der Preis am geringsten / höchsten?
(Hinweis zur Interpretation bei Variablen mit quadratischem Einfluss: Hier kann für $\hat{\beta}_2$ und $\hat{\beta}_3$ keine ceteris-paribus-Interpretation durchgeführt werden. Um den Einfluss des Alters zu bestimmen, muss der (geschätzte) bedingte Erwartungswert von *price* also zunächst nach *age* abgeleitet werden.)

(g) (3 Punkte) Interpretieren Sie den Einfluss der Grundfläche des Hauses und von `manyr`.

(h) (2 Punkte) Zeichnen Sie in `R` die geschätzten Regressionskurven für Modell (9) für $age^* = 0$ in einen *area*-*price*-Scatterplot ein.

(i) (3 Punkte) Zeichnen Sie in `R` die geschätzten Regressionskurven für Modell (9) für $area^* =$ `mean(area)` in einen *age*-*price*-Scatterplot ein.
Zeichnen Sie auch Ihr Ergebnis aus Teil (f) ein.
(Hinweis: Um die Parabel $a\,x^2 + b\,x + c$ für $x \in [-10; 10]$ zu plotten, können Sie die Befehlszeile

```
plot(function(x) a * x^2 + b*x + c, from = -10, to=10)
```

(für bereits definierte `a`, `b`, `c`) verwenden. Wenn Sie die Parabel in einen bereits bestehenden Plot einfügen möchten, müssen Sie noch die Option `add=TRUE` ergänzen.)

36. Aufgabe (27 Punkte) (*Asymptotische Hypothestentests im Regressionsmodell*)

Die Datei bss-data.zip enthält Daten zur Erklärung des Hauspreises in der Bretagne und das Readme file eine kurze Beschreibung der Variablen.

Die Daten wurden im Artikel "C. Bontemps, M. Simioni, Y. Surry. 2008. Semiparametric Hedonic Price Models: Assessing the Effects of Agricultural Nonpoint Source Pollution. *Journal of Applied Econometrics* **23**: 825-842." untersucht.

Dabei wurden die den (logarithmierten) Hauspreis erklärenden Variablen in drei Kategorien unterteilt: Charakteristika des Hauses (Alter, Zustand, Anzahl der Zimmer, Grundstücksgröße), Charakteristika des Bezirks (im Département Ille et Villaine oder nicht, Anteil der leerstehenden Häuser, Bevölkerung, Durchschnittseinkommen), umweltbezogene Charakteristika des Bezirks (Anteil an Weideland, Nitrogenkonzentration). Das Hauptinteresse des Artikels lag bei den umweltbezogenen Charakteristika. Importieren Sie die Daten mit folgenden Programmzeilen:

```
daten <- read.table("bss-data.txt", header=FALSE,
            col.names=c("price_log", "age", "repair", "rooms", "lot",
            "county", "vacant", "pop", "avinc", "tmead", "nitro"))
```

(a) (2 Punkte) Erstellen Sie analog zu S. 835 des Artikels eine Tabelle mit deskriptiven Statistiken und vergleichen Sie Ihre mit der des Artikels. Stimmen die Ergebnisse für die Preis- und Einkommensvariablen überein?

Hinweis: Sie können entweder alle Größen einzeln erzeugen oder das Paket prettyR verwenden, welches die Funktion `describe` enthält, bei der Sie selbst auswählen können, welche deskriptiven Statistiken für welche Variablen ausgegeben werden sollen.

(b) (1 Punkt) Schätzen Sie das Modell

$$\begin{aligned}\log(price) = &\beta_1 + \beta_2 age + \beta_3 repair + \beta_4 rooms + \beta_5 lot + \\ &+ \beta_6 county + \beta_7 vacant + \beta_8 pop + \beta_9 avinc + \\ &+ \beta_{10} tmead + \beta_{11} nitro + u\end{aligned}$$

mit KQ. Verwenden Sie dazu die Einheiten wie im ursprünglichen Datensatz.

(c) (1 Punkt) Wenden Sie den `plot`-Befehl auf Ihr `lm`-Objekt an. Erklären Sie hierbei kurz den Quantile-Quantile und den Scale-Location Graphen.

(d) (1 Punkt) Interpretieren Sie $\hat{\beta}_3$, $\hat{\beta}_5$ und $\hat{\beta}_{10}$.

(e) (1 Punkt) Wie lautet die Verteilung von $\hat{\boldsymbol{\beta}} = (\hat{\beta}_1, \ldots, \hat{\beta}_{11})^T$?

(f) (1 Punkt) Wie schätzen Sie die Anpassungsgüte des verwendeten Modells ein?

(g) (1 Punkt) Erstellen und kommentieren Sie das Histogramm der Residuen dieser Schätzung.

(h) (1 Punkt) Überprüfen Sie mit Hilfe eines (asymptotischen) t-Tests, ob die Bevölkerung des Bezirks (*pop*) signifikanten Einfluss auf den Hauspreis hat.

(i) (2 Punkte) Verwenden Sie Ihren Output aus Teil (b), um das Hypothesenpaar

$$\mathrm{H}_0: \beta_3 = 0.4, \quad \mathrm{H}_1: \beta_3 \neq 0.4$$

mit einem (asymptotischen) t-Test zu überprüfen.

Welches Modell müssen Sie schätzen, um das Ergebnis des Tests direkt vom ensprechenden KQ-Output ablesen zu können?

(j) (2 Punkte) Überprüfen Sie die Hypothese
"*Häuser in gutem Zustand kosten ceteris paribus durchschnittlich genauso viel wie 120 Jahre jüngere in schlechtem Zustand.*"
mit Hilfe eines (asymptotischen) t-Tests. Geben Sie dazu das Modell an, das Sie schätzen müssen, um das Ergebnis an Hand eines KQ-Outputs unmittelbar ablesen zu können.

(k) (1 Punkt) Welche zusätzliche Information bräuchten Sie, um den Test aus Teil (j) mit Hilfe des KQ-Outputs aus Teil (b) durchführen zu können?

In den folgenden Teilaufgaben soll das Hypothesenpaar

$$\mathrm{H}_0: \beta_3 = 3\beta_4,\ \beta_7 = -0.02,\ \beta_9 = 0.04$$
$$\mathrm{H}_1: \beta_3 \neq 3\beta_4 \text{ und/oder } \beta_7 \neq -0.02 \text{ und/oder } \beta_9 \neq 0.04$$

überprüft werden (keine inhaltliche Interpretation, nur Training).

(l) (2 Punkte) Geben Sie das unter H_0 gültige Modell an und schätzen Sie dieses mit KQ.

(m) (1 Punkt) Bestimmen Sie die Residuenquadratsummen des Null- und Alternativhypothesenmodells mit Hilfe der `R`-Outputs aus Teil (b) und (l).

(n) (2 Punkte) Führen Sie nun einen geeigneten (asymptotischen) Test zur Überprüfung des obigen Hypothesenpaars durch.
Hinweis: $F = \frac{(SSR_{H_0} - SSR_{H_1})/q}{SSR_{H_1}/(n-k)} \overset{a}{\sim} F_{q,n-k}$.

(o) (1 Punkt) Bestimmen Sie mit Hilfe von `R` den p-value für den Test aus Teil (n) und interpretieren Sie diesen.
Hinweis: Wenn Sie die Teststatistik in `R` berechnen wollen, können Sie SSR mittels der Funktion `deviance`, angewandt auf das `lm`-Objekt, ausgeben lassen.

(p) (1 Punkt) Formulieren Sie die Nullhypothese in der Form $\mathrm{H}_0: \mathbf{R}\boldsymbol{\beta} = \mathbf{r}$, geben Sie $\mathbf{R}$ und $\mathbf{r}$ also explizit an.

(q) (1 Punkt) Installieren Sie das Zusatzpaket car und erläutern Sie die in der zugehörigen pdf-Dokumentation beschriebene Funktion `linearHypothesis`.

(r) (2 Punkte) Führen Sie nun den F-Test unter Verwendung des Befehls `linearHypothesis` durch. Verwenden Sie zunächst die Methode, bei der Sie die Matrix $\mathbf{R}$ und den Vektor $\mathbf{r}$ explizit eingeben müssen.

(s) (1 Punkt) Wiederholen Sie den Test, indem Sie nun die Restriktionen als Linearkombination der Parameter eingeben.

(t) (1 Punkt) Führen Sie zum Hypothesenpaar aus Teil (j) einen (asymptotischen) F-Test durch und vergleichen Sie die Teststatistiken und die Ergebnisse der beiden Tests.

(u) (1 Punkt) Überprüfen Sie mit Hilfe des Outputs aus Teil (b) das Hypothesenpaar

$$\mathrm{H}_0: \beta_2 = \beta_3 = \ldots = \beta_{11} = 0,$$
$$\mathrm{H}_1: \beta_j \neq 0 \text{ für mindestens ein } \beta_j,\ j = 2, \ldots, 11,$$

führen Sie also den (asymptotischen) Overall-F-Test für dieses Modell durch.

37. Aufgabe (9 Punkte) (*Die Spur einer Matrix - The trace operator*)

Die Spur einer $(n \times n)$-Matrix $\mathbf{A} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$ ist definiert als die Summe ihrer Diagonaleinträge, also

$$\mathrm{tr}(\mathbf{A}) = \sum_{i=1}^{n} a_{ii}.$$

Die Einträge des Produkts $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$ zweier Matrizen $\underbrace{\mathbf{A}}_{(n\times m)}$ und $\underbrace{\mathbf{B}}_{(m\times n)}$ sind

$$c_{ij} = \sum_{k=1}^{m} a_{ik}b_{kj},$$

also gilt insbesondere für die Diagonaleinträge $c_{ii} = \sum_{k=1}^{m} a_{ik}b_{ki}$.
Die Spur der $(n \times n)$-Produktmatrix $\mathbf{C}$ ist also

$$\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{C}) = \sum_{i=1}^{n} c_{ii} = \sum_{i=1}^{n}\sum_{k=1}^{m} a_{ik}b_{ki}.$$

Zeigen Sie die folgenden Aussagen.
Vorsicht mit den Dimensionen!

(a) (1 Punkt) Die Spur der $(n \times n)$-Einheitsmatrix ist gerade n, also $\mathrm{tr}(\mathbf{I}_n) = n$.

(b) (1 Punkt) Die Spur der transponierten Matrix $\mathbf{A}^T$ ist gleich der Spur der $(n \times n)$-Matrix $\mathbf{A}$, also $\mathrm{tr}(\mathbf{A}^T) = \mathrm{tr}(\mathbf{A})$.

(c) (1 Punkt) Die Spur der Summe zweier $(n \times n)$-Matrizen $\mathbf{A}$ und $\mathbf{B}$ ist gleich der Summe der Spuren der beiden Matrizen, also $\mathrm{tr}(\mathbf{A} + \mathbf{B}) = \mathrm{tr}(\mathbf{A}) + \mathrm{tr}(\mathbf{B})$.

(d) (1 Punkt) Die Spur einer mit einem Skalar α multiplizierten $(n \times n)$-Matrix $\mathbf{A}$ ist gleich der mit α multiplizierten Spur von $\mathbf{A}$, also $\mathrm{tr}(\alpha \cdot \mathbf{A}) = \alpha \cdot \mathrm{tr}(\mathbf{A})$.

(e) (3 Punkte) Der Spur-Operator ist invariant unter zyklischer Permutation, also z.B. für vier geeignet dimensionierte Matrizen $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$ und $\mathbf{D}$ ist $\mathrm{tr}(\mathbf{ABCD}) = \mathrm{tr}(\mathbf{BCDA}) = \mathrm{tr}(\mathbf{CDAB}) = \mathrm{tr}(\mathbf{DABC})$.
Zeigen Sie hierfür zunächst, dass dies für die beiden Matrizen $\underbrace{\mathbf{A}}_{(n\times m)}$ und $\underbrace{\mathbf{B}}_{(m\times n)}$ gilt, also $\mathrm{tr}(\mathbf{AB}) = \mathrm{tr}(\mathbf{BA})$.

(f) (2 Punkte) Im Modell $\mathbf{y} = \underbrace{\mathbf{X}}_{(n\times k)}\boldsymbol{\beta} + \mathbf{u}$ gilt $\mathrm{rk}(\mathbf{P_X}) = k$ und $\mathrm{rk}(\mathbf{M_X}) = n - k$. Zeigen Sie dies, indem Sie verwenden, dass der Rang idempotenter Matrizen gleich ihrer Spur ist.
Das Ergebnis $\mathrm{rk}(\mathbf{P_X}) = k$ und $\mathrm{rk}(\mathbf{M_X}) = n - k$ zeigt auch, dass $\delta(\mathbf{X})$ die Dimension k und $\delta^{\perp}(\mathbf{X})$ die Dimension $n - k$ hat.

Bemerkung: Dank den Eigenschaften c) und d) ist die Spur eine lineare Abbildung vom Raum der quadratischen Matrizen mit der Matrizenaddition $(\mathbb{R}^{n\times n}, +)$ nach $\mathbb{R}$.

38. Aufgabe (4 Punkte) (*Verzerrter und unverzerrter Schätzer der Fehlervarianz*)

(a) (3 Punkte) Machen Sie sich mit den Eigenschaften des Spur-Operators (zum Beispiel mit Aufgabe 37) vertraut und zeigen Sie damit, dass im Modell

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{u}, \quad \mathbf{u}|\mathbf{X} \sim (\mathbf{0}, \sigma^2\mathbf{I})$$

für den unkorrigierten Varianzschätzer $\hat{\sigma}^2 = \frac{1}{n}\sum_{t=1}^{n} \hat{u}_t^2$

$$E(\hat{\sigma}^2|\mathbf{X}) = \frac{n-k}{n}\sigma_0^2$$

gilt (vgl. Skript Abschnitt "Schätzen der Fehlervarianz").

(b) (1 Punkt) Welcher Schätzer für die Fehlervarianz ist unverzerrt?

39. Aufgabe (20 Punkte) (*Asymptotik und B3 am Beispiel eines Binomialmodells*)
Betrachten Sie für die Zufallsstichprobe $t = 1, \ldots, n$ das (korrekt spezifizierte) Modell:

$$wage_t = \mathbf{X}_t\boldsymbol{\beta} + u_t = \beta_1 + \beta_2\, male_t + u_t, \quad male_t \sim IID\, B(p),\, 0 < p < 1, \quad u_t|\mathbf{X} \sim NID(0, \sigma^2). \tag{10}$$

Dabei bezeichnet $male_t$ eine Dummyvariable für das Geschlecht eines beobachteten Individuums (1=männlich, 0=weiblich). Wenn in der Grundgesamtheit $p \cdot 100\%$ der Individuen männlich sind, dann folgt beim Ziehen einer Zufallsstichprobe für das Geschlecht eines beobachteten Individuums $male_t$ der Bernoulli-Verteilung mit "Erfolgswahrscheinlichkeit" p, kurz $male_t \sim B(p)$:

$$male_t = \begin{cases} 1 & \text{mit } \mathrm{P}(male_t = 1) = p \\ 0 & \text{mit } \mathrm{P}(male_t = 0) = 1 - p \end{cases}.$$

(a) (1 Punkt) Erläutern Sie, wann der KQ-Schätzer für $\boldsymbol{\beta}$ für eine gegebene Stichprobe nicht berechnet werden kann.

(b) (3 Punkte) Zeigen Sie, dass

$$p^n + (1-p)^n \tag{11}$$

die Wahrscheinlichkeit angibt, eine Stichprobe mit perfekter Kollinearität zu erhalten.
Geben Sie danach die Wahrscheinlichkeit, dass Annahme (B3) verletzt ist, für $p = 0.6$ und $n = 20$ bzw. $n = 50$ an.
Begründen Sie anhand Gleichung (11), warum die Wahrscheinlichkeit für perfekte Kollinearität mit steigender Stichprobengröße abnimmt.

(c) (2 Punkte) Bestimmen Sie (für eine feste Stichprobengröße n) den Wert des Parameters p, für den die Wahrscheinlichkeit eine Stichprobe mit perfekter Kollinearität zu erhalten am geringsten ist.

(d) (1 Punkt) Im Folgenden werde $\hat{\gamma} = \hat{\beta}_1 + \hat{\beta}_2$ untersucht:
Wie lässt sich der Parameter $\hat{\gamma}$ für Modell (10) interpretieren?

(e) (1 Punkt) Ermitteln Sie den Vektor $\mathbf{w}$, für den gilt: $\hat{\gamma} = \mathbf{w}^T\hat{\boldsymbol{\beta}}$.

(f) (1 Punkt) Bestimmen Sie $\mathrm{Var}(\hat{\gamma}|\mathbf{X})$ in Abhängigkeit von σ_0^2, $\mathbf{X}$ und $\mathbf{w}$.

(g) (2 Punkte) Zeigen Sie mit Hilfe Ihres Ergebnisses der vorherigen Teilaufgabe, dass $\mathrm{Var}(\hat{\gamma}|\mathbf{X}) = \frac{\sigma_0^2}{n\,\hat{p}_n}$ gilt und interpretieren Sie das Ergebnis.
Dabei bezeichne $\hat{p}_n$ den Anteil der Männer in der gezogenen Stichprobe, also $\hat{p}_n = \frac{1}{n}\sum_{t=1}^{n} male_t$.
Nehmen Sie an, dass $\hat{p}_n$ weder 0 noch 1 ist.

(h) (1 Punkt) Bestimmen Sie $\mathrm{E}\left[\frac{\mathbf{X}^T\mathbf{X}}{n}\right]$.

(i) (2 Punkte) Berechnen Sie zuerst $\boldsymbol{X}^T\boldsymbol{X}$ für beliebige Designmatrix $\boldsymbol{X} = \begin{pmatrix}\boldsymbol{x}_1 & \cdots & \boldsymbol{x}_k\end{pmatrix}$.
Was hängt in der Matrix $\frac{1}{n}\boldsymbol{X}^T\boldsymbol{X}$ alles von n ab?
Bestimmen Sie dann $\underset{n\to\infty}{\mathrm{plim}}\left(\frac{\mathbf{X}^T\mathbf{X}}{n}\right) =: \boldsymbol{S}_{\mathbf{X}^T\boldsymbol{X}}$.

(j) (1 Punkt) Erklären Sie nun mit eigenen Worten, was die Annahme (A1) bedeutet bzw. welche Theoreme hier implizit ihre Anwendung finden?

(k) (2 Punkte) Bestimmen Sie $\underset{n\to\infty}{\mathrm{plim}}\left(\frac{\mathbf{X}^T\mathbf{u}}{n}\right)$.

(l) (2 Punkte) Zeigen Sie

$$\underset{n\to\infty}{\mathrm{plim}}\ \hat{\boldsymbol{\beta}}_n = \boldsymbol{\beta}_0.$$

(m) (1 Punkt) Bestimmen Sie $\underset{n\to\infty}{\mathrm{plim}}\ \hat{\gamma}_n$.

40. Aufgabe (12 Punkte) (*Asymptotik im unterspezifizierten Modell*)

Betrachten Sie das Modell

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\boldsymbol{\gamma} + \mathbf{u}, \quad \mathbf{u}|\mathbf{X},\mathbf{Z} \sim IID(\mathbf{0}, \sigma^2\mathbf{I}),$$

das den DGP für $\boldsymbol{\beta} = \boldsymbol{\beta}_0$, $\boldsymbol{\gamma} = \boldsymbol{\gamma}_0 \neq \mathbf{0}$, $\sigma^2 = \sigma_0^2$ enthalte.
Allerdings wird nun das unterspezifizierte Modell

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

geschätzt. Der zugehörige KQ-Schätzer wird mit $\widehat{\ }$ gekennzeichnet, während der aus dem korrekt spezifizierten Modell mit $\widetilde{\ }$ gekennzeichnet wird.

(a) (2 Punkte) Zeigen Sie, dass für $MSE(\hat{\boldsymbol{\beta}}|\mathbf{X},\mathbf{Z}) = \mathrm{E}\left[(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}_0)^T|\mathbf{X},\mathbf{Z}\right]$ gilt

$$MSE(\hat{\boldsymbol{\beta}}|\mathbf{X},\mathbf{Z}) = \sigma_0^2(\mathbf{X}^T\mathbf{X})^{-1} + (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Z}\boldsymbol{\gamma}_0\boldsymbol{\gamma}_0^T\mathbf{Z}^T\mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}.$$

(b) (1 Punkt) Zeigen Sie

$$MSE(\tilde{\boldsymbol{\beta}}|\mathbf{X},\mathbf{Z}) = \sigma_0^2(\mathbf{X}^T\mathbf{M_Z}\mathbf{X})^{-1}.$$

(c) (2 Punkte) In den folgenden Teilaufgaben wird nun gezeigt, dass der MSE von $\tilde{\boldsymbol{\beta}}$ asymptotisch verschwindet (also wenn das korrekt spezifizierte Modell geschätzt wird), der MSE von $\hat{\boldsymbol{\beta}}$ aus dem unterspezifizierten Modell dagagen nicht.
Dazu wird hier lediglich der Fall betrachtet, dass die Matrix $\mathbf{X}$ eine $(n \times 2)$-Matrix sei, die in der ersten Spalte eine Konstante enthält und in der zweiten Spalte Ausprägungen der Zufallsvariablen x_t, und $\mathbf{Z}$ sei ein $(n \times 1)$-Vektor, der die Ausprägungen der Zufallsvariablen z_t enthält. Dabei seien x_t und z_t jeweils IID-verteilt mit endlichem Erwartungswert und

endlicher Varianz. Die Kovarianz zwischen x_t und z_t sei ebenfalls endlich.

Zeigen Sie

$$\operatorname*{plim}_{n\to\infty} \frac{\mathbf{X}^T\mathbf{X}}{n} = \begin{pmatrix} 1 & \mathrm{E}[x_t] \\ \mathrm{E}[x_t] & \mathrm{Var}(x_t) + \mathrm{E}[x_t]^2 \end{pmatrix} =: \mathbf{S}_{\mathbf{X}^T\mathbf{X}},$$

(d) (1 Punkt)

$$\operatorname*{plim}_{n\to\infty} \frac{\mathbf{Z}^T\mathbf{Z}}{n} = \mathrm{Var}(z_t) + \mathrm{E}[z_t]^2 =: \mathbf{S}_{\mathbf{Z}^T\mathbf{Z}},$$

(e) (1 Punkt)

$$\operatorname*{plim}_{n\to\infty} \frac{\mathbf{X}^T\mathbf{Z}}{n} = \begin{pmatrix} \mathrm{E}[z_t] \\ \mathrm{Cov}(x_t, z_t) + \mathrm{E}[x_t]\mathrm{E}[z_t] \end{pmatrix} =: \mathbf{S}_{\mathbf{X}^T\mathbf{Z}},$$

(f) (1 Punkt) Für beliebige Matrizen $\mathbf{X}$ und $\mathbf{Z}$ gilt (falls $\mathbf{S}_{\mathbf{Z}^T\mathbf{X}}$ existiert):

$$\operatorname*{plim}_{n\to\infty} \frac{\mathbf{Z}^T\mathbf{X}}{n} =: \mathbf{S}_{\mathbf{Z}^T\mathbf{X}} = \mathbf{S}_{\mathbf{X}^T\mathbf{Z}}^T.$$

(g) (2 Punkte) Nehmen Sie im Folgenden an, dass $\mathbf{S}_{\mathbf{X}^T\mathbf{X}}$ und $\mathbf{S}_{\mathbf{Z}^T\mathbf{Z}}$ invertierbar sind (dies ist in diesem Beispiel erfüllt, falls x_t und z_t positive Varianz haben).

Außerdem sei $\operatorname*{plim}_{n\to\infty} \frac{1}{n} \begin{pmatrix} \mathbf{X}^T \\ \mathbf{Z}^T \end{pmatrix} \begin{pmatrix} \mathbf{X} & \mathbf{Z} \end{pmatrix}$ invertierbar. (Mit der Formel zur Invertierung partitionierter Matrizen $\left(\begin{smallmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} \mathbf{W} & -\mathbf{W}\mathbf{B}\mathbf{D}^{-1} \\ -\mathbf{D}^{-1}\mathbf{C}\mathbf{W} & \mathbf{D}^{-1}+\mathbf{D}^{-1}\mathbf{C}\mathbf{W}\mathbf{B}\mathbf{D}^{-1} \end{smallmatrix}\right)$ mit $\mathbf{W} = (\mathbf{A} - \mathbf{B}\mathbf{D}^{-1}\mathbf{C})^{-1}$ ist dann der Wahrscheinlichkeitslimes von $\frac{1}{n}\mathbf{W} = \frac{1}{n}(\mathbf{X}^T\mathbf{X} - \mathbf{X}^T\mathbf{Z}(\mathbf{Z}^T\mathbf{Z})^{-1}\mathbf{Z}^T\mathbf{X})$ invertierbar.)

Zeigen Sie

$$\operatorname*{plim}_{n\to\infty} MSE(\tilde{\boldsymbol{\beta}}|\mathbf{X}, \mathbf{Z}) = \mathbf{0}.$$

(h) (2 Punkte) Bestimmen Sie

$$\operatorname*{plim}_{n\to\infty} MSE(\hat{\boldsymbol{\beta}}|\mathbf{X}, \mathbf{Z})$$

und begründen Sie, warum $MSE(\hat{\boldsymbol{\beta}}|\mathbf{X}, \mathbf{Z})$ asymptotisch nicht verschwindet.

41. Aufgabe (12 Punkte) (*Erstellung einer Simulation*)

Ziel dieser Aufgabe ist es, in `R` eine Simulation mit R Replikationen für das Modell

$$y_t = \beta_1 + \beta_2 x_{t1} + \beta_3 x_{t2} + \beta_4 x_{t3} + u_t, \qquad t = 1, \dots, n$$

durchzuführen, um zu untersuchen, wie sich die Parameterschätzungen mit zunehmender Beobachtungszahl und zunehmender Replikationsanzahl verhalten.

(a) Öffnen Sie ein neues R-Skript.

(b) (1 Punkt) Da die Daten simuliert sein sollen, müssen Sie zunächst die Variablen x_1, x_2 und x_3, sowie die Fehler u erzeugen. Dabei sei $x_1 \sim i.i.d.N(2, 8)$, $x_2 \sim i.i.d.N(-5, 5)$ und x_3 bestehe aus einem $i.i.d.N(0, 1)$-verteilten Teil plus einem Teil, der x_2 enthält. Die Fehler u seien $i.i.d.N(0, 4)$-verteilt. Übergeben Sie zunächst die Parameter für die Erstellung der Variablen. Dies geschieht beispielsweise für x_1 mit

```
parX1 <- c(2,8)
```

(c) (1 Punkt) Übergeben Sie nun die wahren Parameterwerte des DGPs in den Vektor `beta`. Diese seien $\beta_1 = 2$, $\beta_2 = -5$, $\beta_3 = 7$ und $\beta_4 = 10$.

(d) (1 Punkt) Legen Sie die Anzahl der Beobachtungen n als 100 fest.

(e) (1 Punkt) Verwenden Sie, um Ihre Ergebnisse wieder reproduzieren, bzw. mit denen anderer vergleichen zu können, als Randomseed die Zahl 1234.

(f) (1 Punkt) Erzeugen Sie nun die x-Variablen und die Fehler. Dabei sollen x_1 und x_2 nur die oben beschriebenen normalverteilten Komponenten enthalten und x_3 außerdem noch einen Teil, der mit x_2 zusammenhängt (s. Bsp.). Die Variablen x_2 und x_3 sind somit nicht unabhängig, x_1 dagegen ist unabhängig von x_2 und x_3. Beispielsweise für x_3 sieht der Befehl folgendermaßen aus

```
x3 <- rnorm(n, mean=parX3[1], sd=sqrt(parX3[2])) + 0.2*x2
```

Der Befehl `rnorm(n, mean=<mean>, sd=<standard deviation>)` erzeugt n Zufallsvariablen aus der $N(\text{mean}, \text{sd}^2)$-Verteilung in einem Vektor. Für standardnormalverteilte Zufallsvariablen können `mean=` und `sd=` weggelassen werden.

(g) (1 Punkt) Generieren Sie y.

(h) (1 Punkt) Plotten Sie y jeweils gegen x_1, x_2 und x_3 und die x-Variablen untereinander. Mit der Zeile `par(mfrow=c(2,3))` vor den `plot`-Befehlen können Sie alle 6 Plots in einem Fenster ausgeben. Wie können Sie bei den x-x-Plots die Zusammenhänge der x-Variablen erkennen?

(i) (1 Punkt) Schätzen Sie mit allen möglichen Kombinationen an Regressoren x_1, x_2 und x_3 (jeweils mit Konstante) die Parameter, d.h. das korrekt spezifizierte Modell und alle 6 unterspezifizierten. Die einzelnen Koeffizienten können Sie über

```
coef(<name>)
```

abfragen. Füllen Sie folgende Tabelle aus.

Parameter / Schätzung	β_1	β_2	β_3	β_4
Wert im DGP	2	-5	7	10
`y ~ x1 + x2 + x3`				
`y ~ x1 + x2`				—
`y ~ x1 + x3`			—	
`y ~ x2 + x3`		—		
`y ~ x1`			—	—
`y ~ x2`		—		—
`y ~ x3`		—	—	

(j) (1 Punkt) Schreiben Sie ausgehend von Ihrem bisherigen Programm eines, in dem Sie die Teilaufgaben (a) bis (i) (ohne Plots) 100 mal wiederholen und bilden Sie die Mittelwerte der Parameterschätzungen der einzelnen Modelle. Füllen Sie analog zu oben die Tabelle aus. Was stellen Sie fest, wenn Sie die Ergebnisse mit denen aus Teilaufgabe (i) vergleichen? Lassen Sie sich außerdem für das korrekt spezifizierte Modell und das ohne x_3 jeweils ein Histogramm zu $\hat{\beta}_3$ ausgeben.

Hinweise: Schreiben Sie eine Schleife, die die Erzeugung der Variablen 100 mal wiederholt. Achten Sie darauf, den Randomseed außerhalb der Schleife zu setzen. Warum ist dies wichtig? Speichern Sie die Koeffizientenschätzer der einzelnen Durchläufe in einer Matrix oder einer Matrix pro Schätzung, die in ihren Spalten die Parameterschätzer enthält / enthalten und in den 100 Zeilen die Durchläufe. Bilden Sie jeweils die Mittelwerte der Spalten. Dies geschieht mit dem Befehl `colMeans(<matrix name>)`.

(k) (1 Punkt) Wiederholen Sie die Teilaufgaben (a) bis (i) für einen Stichprobenumfang von 1000. Was stellen Sie fest? Wiederholen Sie auch Teil (j).

(l) (1 Punkt) Wie sehen die Histogramme bei $n = 1000$ und $R = 1000$ aus? Kommentieren Sie diese.

(m) (1 Punkt) Berechnen Sie die Verzerrung, die der Parameterschätzer $\hat{\beta}_3$ im Modell $y = \beta_1 + \beta_2 x_1 + \beta_3 x_2 + u$ aufweist.

Partitionieren Sie dazu am besten die Regressoren in x_3 und den Rest und bestimmen Sie die Verzerrung des gesamten $\boldsymbol{\beta}$-Vektors aus dem Modell ohne x_3.

Passt das zu den Ergebnissen der Simulation?

42. Aufgabe (10 Punkte) (*Monte-Carlo-Simulation zum Mittelwertschätzer*)

Betrachten Sie die Zufallsvariablen $y_t \sim NID(3, 1)$ und $x_t = y_t^2$, $t = 1, \dots, n$.

(a) (1 Punkt) Berechnen Sie $\mathrm{E}[x_t]$.

(b) (2 Punkte) Nun wird der Mittelwertschätzer $\hat{\mu}_n$ von x_t untersucht:

$$\hat{\mu}_n = \frac{1}{n} \sum_{t=1}^{n} x_t.$$

Geben Sie den Wahrscheinlichkeitslimes von $\hat{\mu}_n$ an, d.h. bestimmen Sie

$$\operatorname*{plim}_{n \to \infty} \hat{\mu}_n.$$

Begründen Sie alle Ihre Schritte.

(c) Im Folgenden soll die Konvergenz in Wahrscheinlichkeit mit Hilfe von Monte-Carlo-Simulationen graphisch veranschaulicht werden. Hierfür werden R Stichproben der Größe n gezogen und für jede dieser Stichproben wird der Mittelwertschätzer $\hat{\mu}_n$ berechnet. Nun hat man R Mittelwertschätzungen, deren (empirische) Dichte geplottet werden kann. Indem Sie solche Simulationen mit unterschiedlichen Stichprobengrößen durchführen, können Sie den Einfluss der Stichprobengröße auf die Schätzeigenschaften veranschaulichen.

Erstellen Sie dazu mit `R` ein Programm, in dem Sie die folgenden Schritte berücksichtigen:

i. (1 Punkt) Setzen Sie den Randomseed `42`. Definieren Sie die Variablen `n1`, `n2` und `R`, die die beiden Stichprobengrößen $n_1 = 100$ und $n_2 = 1000$ und die Anzahl der Wiederholungen $R = 1000$ festlegen.

ii. (4 Punkte) Ziehen Sie R mal eine Stichprobe der Größe n_1, so dass Sie für die r-te Stichprobe $y_t^{(r)}$ und damit $x_t^{(r)}$, $t = 1, 2, \ldots, n_1$, erhalten. Dabei bezeichnet der Index (r) die r-te Stichprobe. Berechnen Sie für jede Stichprobe $r = 1, \ldots, R$ den Mittelwertschätzer $\hat{\mu}_{n_1}^{(r)}$ und speichern diesen in der r-ten Zeile (oder Spalte) des Vektors `mu1`.
Hinweis: Sie können die Stichprobenwerte für x_t für alle Wiederholungen in einer $(n_1 \times R)$-Matrix zusammenzufassen und den Vektor `mu1` mittels des Befehls `colMeans` bilden.
Anschließend führen Sie diese Aufgabe für die Stichprobengröße n_2 durch und speichern die Ergebnisse in `mu2`.
(Hinweis: Alternativ kann die gesamte Teilaufgabe mit einer `for`-Schleife programmiert werden, die die Vektoren `mu1` und `mu2` füllt.)

iii. (2 Punkte) Plotten Sie die empirischen Dichten des Mittelwertschätzers $\hat{\mu}_n^{(r)}$ von $x_t^{(r)}$ für die beiden Stichprobengrößen $n_1 = 100$ und $n_2 = 1000$. Verwenden Sie dazu beispielsweise den folgenden Programmcode und erklären Sie, was dabei geschieht:

```
plot( density(mu1), ylim=c(0,max(density(mu1)$y, density(mu2)$y)),
            xlab=paste("Dichten von mu für n=", n1, " und n=", n2, sep=""),
            main="Dichteplots",
            col="red", lwd=2)
lines(density(mu2), col="blue", lwd=2)
legend("topright", inset=0.02,
      legend=c(paste("Dichte von mu für n=", n1, sep=""),
               paste("Dichte von mu für n=", n2, sep="")),
      lwd=c(2,2), col=c("red", "blue"))
```

Hinweise: `mu1` und `mu2` in diesem Programmcode bezeichnen die Vektoren, die $\hat{\mu}_{n_1}^{(r)}$ bzw. $\hat{\mu}_{n_2}^{(r)}$, $r = 1, \ldots, R$ enthalten, und `n1` und `n2` die Stichprobengrößen n_1 und n_2.
Mittels `density` wird eine Kernschätzung (nicht relevant in dieser Veranstaltung) der empirischen Dichte durchgeführt. Einfach ausgedrückt kann man sich das Ergebnis als geglättetes Histogramm vorstellen.
Mittels `ylim` wird angegeben, in welchem Bereich sich die Ordinate bewegt. Hierbei soll erreicht werden, dass beide Dichten vollständig im Plot enthalten sind.
`density(mu1)$y` gibt die Dichtewerte von `density` zurück (muss nicht erklärt werden).

iv. Probieren Sie Ihr Programm für andere Werte für die Stichprobengrößen und die Anzahl der Wiederholungen aus.

43. Aufgabe (22 Punkte) (*Bootstrap(simulation)*)

Diese Aufgabe ist unterteilt in mehrere kleine Abschnitte. Der erste Abschnitt befasst sich mit dem Verständnis von Bootstrap und der Einordnung im Vergleich zu Monte-Carlo-Simulationen. Im zweiten Abschnitt wird eine kleine Stichprobe genauer untersucht, wohingegen im dritten Abschnitt eine Simulation durchgeführt wird.

(a) Für den ersten Teil der Aufgabe lesen Sie neben den Vorlesungsmaterialien auch ökonometrische Begleitliteratur, um folgende Begriffe in eigenen Worten zu erklären.

i. (2 Punkte) Erklären Sie den Unterschied zwischen der Herangehensweise per Monte-Carlo/Bootstrap und asymptotischer Theorie.

ii. (1 Punkt) Erklären Sie, was man unter parametrischen und semi- bzw. nichtparametrischen Bootstraping versteht?

iii. (2 Punkte) Was ist gleich und was ist unterschiedlich bei einer Monte-Carlo-Simulation und einem Bootstrap? Wann funktioniert das eine, das andere aber nicht?

iv. (1 Punkt) Für diese Aufgabe lohnt sich eine kurze Internetrecherche. Erklären Sie anhand einer beliebigen Stichprobe $\boldsymbol{x} = (x_1, ..., x_n)$, was man unter den Resampling-Methoden (Wiederholungsprobenentnahme) *Jackknife*, *nicht-parametrischen Bootstrap* und *Subsampling* versteht.

(b) Sei nun X eine Zufallsvariable mit unbekannter Verteilungsfunktion F_X. Nach fünf unabhängigen Ziehungen aus F_X liege die Stichprobe

$$\boldsymbol{x} = (1.879049, 2.539645, 6.117417, 3.141017, 3.258575)$$

vor. Von Interesse sei der Erwartungswert μ bzgl. der Verteilung F_X.

i. (1 Punkt) Nennen Sie einen erwartungstreuen Schätzer $\hat{\mu}$ des Erwartungswerts μ.

ii. (2 Punkte) Erklären Sie im Hinblick auf den Mittelwert den Unterschied zwischen den Begriffen Standardabweichung und Standardfehler.

iii. (2 Punkte) Da der Stichprobenumfang gering ist, möchten Sie eine Resampling-Methode, genauer gesagt einen nicht-parametrischen Bootstrap, auf diese Stichprobe anwenden. Um die Verteilung zu approximieren, kreieren Sie nun aus obiger Stichprobe der Größe 5 durch „Ziehen mit Zurücklegen" neue Stichproben. Wie viele verschiedene solcher Stichproben gibt es? Berechnen Sie auch die Anzahl, falls die Stichprobe 10 verschiedene Werte enthält.
Hinweis: Die Stichprobe enthält 5 *verschiedene* Werte.

iv. (1 Punkt) Erzeugen Sie mit dieser Stichprobe in `R` 4 Bootstrap-Stichproben $\boldsymbol{x}^{\star 1}, ..., \boldsymbol{x}^{\star 4}$, berechnen Sie zuerst jeweils den Mittelwert jeder Stichprobe, danach den Gesamtmittelwert und den Standardfehler des Mittelwerts. Hinweise:

- Benutzen Sie der Reproduzierbarkeit halber einen `set.seed(123)`.
- Einlesen: `x <- c(1.879049, 2.539645, 6.117417, 3.141017, 3.258575)`.
- Bootstrap-Stichproben, -Mittelwert, - Standardfehler: `?sample`, `?mean`, `?sd`.

v. (2 Punkte) Erhöhen Sie nun mit einer Schleife die Anzahl der Bootstrap-Replikationen. Welche Anzahl B von Replikationen erscheint für Sie maximal sinnvoll? Wieso? Wiederholen Sie die Berechnungen der vorangegangen Aufgaben, speichern Sie den Standardfehler .

vi. (2 Punkte) Nutzen Sie dazu anschließend die Funktion `bootstrap()` aus dem gleichnamigen `R`-Paket und vergleichen Sie die Ergebnisse.

(c) Für diesen Abschnitt sollen Sie einen Bootstrap mit dem Häuserdatensatz bss-data.zip einer vorangegangenen Aufgabe durchzuführen. Der Programmcode hierfür finden Sie im Anhang.

Das folgende Modell unterscheidet sich vom alten durch die Interaktion $age_t \cdot repair_t$:

$$\begin{aligned}\log(price) = \beta_1 + \beta_2 age + \beta_3 repair + \beta_4 rooms + \beta_5 lot + \\ + \beta_6 county + \beta_7 vacant + \beta_8 pop + \beta_9 avinc + \\ + \beta_{10} tmead + \beta_{11} nitro + + \beta_{12} age \cdot repair + u\end{aligned}$$

i. (2 Punkte) Erläutern Sie das Vorgehen bei Bootstrap-Tests an Hand des angegebenen `R`-Programmcodes.

Hinweis: Erklärungen zu einzelnen Zeilen des Programmcodes reichen nicht aus, vielmehr soll ein Überblick der Idee und Herangehensweise vermittelt werden.

ii. (1 Punkt) Formulieren Sie den angegebenen Programmcode so um, dass die Varianz der Residuen u_t^* der geschätzen Varianz der Fehler entspricht.

iii. (1 Punkt) Nehmen Sie im Folgenden an, dass die Fehler normalverteilt sind, und ändern Sie den Programmcode so ab, dass ein parametrischer Bootstrap durchgeführt wird.

iv. (1 Punkt) Welche Werte können sich für den Bootstrap-p-value theoretisch ergeben, falls Sie $B = 99$ wählen?

v. (1 Punkt) Sollten Sie an Stelle des Bootstrap-Tests lieber einen entsprechenden asymptotischen Test durchführen? Begründen Sie Ihre Antwort.

Anhang:

R-Programmcode zum semiparametrischen Bootstrap:

```
daten <- read.table("bss-data.txt", header=FALSE,
               col.names=c("price_log", "age", "repair", "rooms", "lot",
               "county", "vacant", "pop", "avinc", "tmead", "nitro"))
attach(daten)
repage     <- repair*age # Interaktion

# Schätzen des Modells mit KQ
ols        <- lm(price_log ~ 1 + age + repair + rooms + lot + county
                        + vacant + pop + avinc + tmead + nitro + repage)
( ols.s    <- summary(ols) )

( tau_hat  <- coef(ols.s)["repage","t value"] ) # Teststatistik

# Modell unter H_0
ols.H0     <- lm(price_log ~ 1 + age + repair + rooms + lot + county
                        + vacant + pop + avinc + tmead + nitro)

# X-Matrix, beta_tilde, u_tilde und n
X          <- model.matrix(ols.H0)
beta_tilde <- coef(ols.H0)
u_tilde    <- resid(ols.H0)
n          <- length(price_log)

set.seed(42) # Randomseed

B     <- 999 # Anzahl der Bootstrap-Replikationen

tau   <- rep(NA, B) # Definieren des Vektors tau, der die Bootstrap-Teststatistiken enthalten soll

for (j in 1:B) # Schleife, die die j-te Stichprobe generiert und die Teststatistik bestimmt
{
# u_star und y_star erzeugen (-> j-te Stichprobe y_star, X)
      u_star      <- sample(u_tilde, n, replace=TRUE)
      y_star      <- X %*% beta_tilde + u_star

      # j-te Teststatistik
      tau[j]      <- coef(summary(
                          lm(y_star ~ 1 + age + repair + rooms + lot + county
+ vacant + pop  + avinc + tmead + nitro + repage))
                     )["repage","t value"]
}

( p_hat    <- 1 - (sum (abs(tau) <= abs(tau_hat)) )/B ) # Bootstrap-p-value
```

44. Aufgabe (5 Punkte) (*Simulation der Schätzerverteilung*)

Verwenden Sie für diese Aufgabe den Datensatz bss-data.zip.

Betrachten Sie wieder das Modell

$$\begin{aligned}\log(price) = \beta_1 + \beta_2 age + \beta_3 repair + \beta_4 rooms + \beta_5 lot + \\ + \beta_6 county + \beta_7 vacant + \beta_8 pop + \beta_9 avinc + \\ + \beta_{10} tmead + \beta_{11} nitro + u.\end{aligned}$$

(a) (1 Punkt) Überprüfen Sie für das obige, mit OLS geschätzte Modell mit Hilfe eines asymptotischen F-Tests die gemeinsame Signifikanz der Variablen, die die Region (nicht die Umwelt) charakterisieren und interpretieren Sie anschließend das Ergebnis.
Hinweis: Sie können für den F-Test die Funktion `anova` verwenden.

(b) Ziel dieser Teilaufgabe ist es, die Verteilung des geschätzen Parametervektors zu simulieren. Gehen Sie dazu in folgenden Schritten vor.

i. (1 Punkt) Ziehen Sie $R = 100$ mal eine Stichprobe mit jeweils $N = 80$ Elementen. Verwenden Sie als Randomseed beispielsweise `1234`.
Hinweis: Mit `smpl <- daten[trunc(runif(N,min=1,max=nrow(daten)+1)),]` erhalten Sie die eingeschränkte Datenmatrix `smpl`, wobei `daten` die zuvor eingelesene Gesamtdatenmatrix bezeichne. Für die Schätzung mit `lm` wird lediglich diese Stichprobe verwendet, wenn Sie nach der Schätzgleichung noch `,data=smpl` einfügen.

ii. (1 Punkt) Schätzen Sie mit jeder Stichprobe obiges Modell und speichern Sie die Parameterschätzer.

iii. (1 Punkt) Geben Sie die Histogramme von $\hat{\beta}_1, \ldots, \hat{\beta}_{11}$ aus und interpretieren Sie diese.

iv. (1 Punkt) Führen Sie die Simulation noch mit drei weiteren Kombinationen von R und N durch.

45. Aufgabe (10 Punkte) (*Monte-Carlo zu Konfidenzintervallen*)
Betrachten Sie den folgenden (unkommentierten) R-Programmcode:

```
n     <- 50
B     <- 20
alpha <- 0.10
sigma2      <- 3
beta        <- c(5, 1.4)

beta_2_hat  <- rep(NA, B)
intervals   <- matrix(NA, ncol=2, nrow=B)

set.seed(42)
x     <- rnorm(n, mean=2, sd=1)

for (j in 1:B)
{
      u     <- rnorm(n, mean=0, sd=sqrt(sigma2))
      y     <- beta[1] + beta[2]*x + u

      ols   <- lm(y ~ 1 + x)

      beta_2_hat[j]     <- coef(ols)["x"]
      intervals[j,]     <- confint(ols, parm="x", level=1-alpha)
}
intervals

plot(c(min(intervals[,1]),max(intervals[,2])), c(1,B), type="n",
      font=2, font.lab=2, xlab="beta", ylab="Stichprobennummer",
      main=paste((1-alpha)*100, "%-Konfidenzintervalle", sep=""))
abline(v=beta[2], lwd=2)
axis(1, at = beta[2], labels = expression(beta[2]))

for (j in 1:B)
{
      lines(c(intervals[j,1],intervals[j,2]), c(j,j), lwd=2)
      points(beta_2_hat[j], j, pch=19, col="blue")
}
```

(a) (2 Punkte) Führen Sie den Programmcode in R aus und erklären Sie was in den einzelnen Zeilen/Abschnitten geschieht.

(b) (2 Punkte) Interpretieren und kommentieren Sie die Graphik, die mittels des Programmcodes erstellt wurde.

(c) (2 Punkte) Was passiert, wenn Sie das Signifikanzniveau α verändern? Erklären Sie dies theoretisch und anschließend, indem Sie den Programmcode entsprechend modifizieren.

(d) (2 Punkte) Was passiert, wenn Sie die Beobachtungszahl erhöhen? Erklären Sie auch dies theoretisch und anschließend durch entsprechende Änderung des Programmcodes.

(e) (2 Punkte) Welche Änderungen erwarten Sie für die Graphik, falls Sie die Varianz der Fehler verdoppeln? Überprüfen Sie Ihre Vermutung mit Hilfe des Programmcodes.

46. Aufgabe (6 Punkte) (*Konfidenzinzervalle und -ellipsoide*)

Verwenden Sie für diese Aufgabe den Datensatz (bss-data.zip). Es werde wieder das Modell

$$\begin{aligned}\log(price) = \beta_1 + \beta_2\, age + \beta_3\, repair + \beta_4\, rooms + \beta_5\, lot + \\ + \beta_6\, county + \beta_7\, vacant + \beta_8\, pop + \beta_9\, avinc + \\ + \beta_{10}\, tmead + \beta_{11}\, nitro + u\end{aligned}$$

betrachtet.

Hinweise:

- Konfidenzintervalle können in `R` mit Hilfe der Funktion `confint` berechnet werden.
- Falls Sie Ihre `lm`-Schätzung aus Teil (a) als `ols` bezeichnen, wird mittels `cov2cor(vcov(ols))` die Schätzung der bedingten Korrelationsmatrix des KQ-Schätzers bestimmt.
- Für die Konfidenzellipsen kann der Befehl `confidenceEllipse` aus dem Zusatzpaket car verwendet werden.

(a) (2 Punkte) Schätzen Sie das Modell mit KQ und geben Sie unter der Annahme asymptotisch normalverteilter Fehler ein 80%-Konfidenzintervall für β_4 an. Berechnen Sie dieses zunächst selbst und anschließend mit Hilfe der Funktion `confint` in `R`.

(b) (1 Punkt) Wovon hängen Form und Ausrichtung von Konfidenzellipsen ab?

(c) (3 Punkte) Schätzen Sie die bedingte Varianz-Kovarianz-Matrix der Parameterschätzer und bestimmen Sie daraus die entsprechende Korrelationsmatrix.
Welche Form erwarten Sie für die einzelnen Konfidenzellipsen? Wählen Sie die beiden extremsten Fälle (in Absolutbeträgen) bezüglich der Form aus und zeichnen Sie die beiden zugehörigen 80%-Konfidenzellipsen mit `R`.
Zeichnen Sie außerdem jeweils die Grenzen der entsprechenden Einzel-Konfidenzintervalle ein.

47. Aufgabe (8 Punkte) (*Konfidenzintervall für den Varianzschätzer*)

Gegeben sei folgender DGP

$$u_i \sim NID(0, \sigma^2).$$

Ihre Aufgabe ist es, auf Basis genau einer von Ihnen simulierten Stichprobe $\{u_i, i = 1, \ldots, n\}$ die Varianz zu schätzen, wobei der Mittelwert $\mu = 0$ als bekannt vorausgesetzt wird.

Verwenden Sie folgenden 'Mittelwert'-Schätzer

$$\tilde{\sigma}_n^2 = \frac{1}{n}\sum_{i=1}^{n} u_i^2 = \frac{1}{n} S_n$$

mit

$$S_n \equiv \sum_{i=1}^{n} u_i^2.$$

(a) (1 Punkt) Bestimmen Sie Erwartungswert und Varianz von $\tilde{\sigma}_n^2$.

(b) (2 Punkte) Zeigen Sie, dass gilt: $n\frac{\tilde{\sigma}_n^2}{\sigma^2} \sim \chi^2(n)$.

(c) (1 Punkt) Leiten Sie das $(1-\alpha)$-Konfidenzintervall für σ^2 ab.

(d) (2 Punkte) Bestimmen Sie die asymptotische Verteilung von $\tilde{\sigma}^2_n$.

(e) (1 Punkt) Leiten Sie das asymptotische $(1-\alpha)$-Konfidenzintervall für σ^2 ab.

(f) (1 Punkt) Vergleichen Sie die Grenzen des exakten und asymptotischen Konfidenzintervalls für $n = 10, 25, 50, 100, 500$ und $\alpha = 0.05$. Generieren Sie hierzu jeweils eine Stichprobe mit n Beobachtungen, $\sigma^2 = 1$ und `set.seed(1234)`.

48. Aufgabe (10 Punkte) (*Interpretation von Konfidenzintervallen*)

Ziel dieser Aufgabe ist es, zu veranschaulichen, dass Konfidenzintervalle zufällig sind, also von der gezogenen Stichprobe abhängen.

Nehmen Sie für diese Aufgabe den folgenden DGP an

$$y = \beta_1 + \beta_2 x + u$$

mit

$$\boldsymbol{\beta} = \begin{pmatrix} 5 \\ 1.4 \end{pmatrix}, \quad x \sim N(2,1), \quad u \sim N(0,3).$$

Schreiben Sie ein Programm, das die folgenden Punkte beinhaltet.

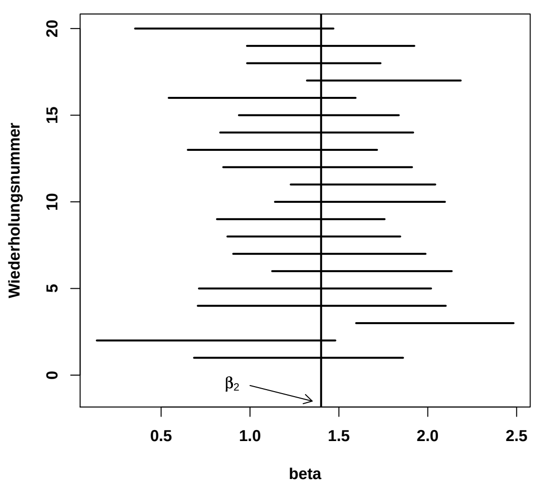
(a) (4 Punkte) Bilden Sie mit diesem DGP `R = 20` Stichproben mit je `n = 50` Beobachtungen und schätzen Sie jeweils das Modell

$$y_i = \beta_1 + \beta_2 x_i + u_i, \quad i = 1, \ldots, n.$$

Bestimmen Sie für `alpha = 0.05` jeweils $(1-\alpha)$-Konfidenzintervalle für β_2 und speichern Sie die Grenzen in einer Matrix ab.

Für Konfidenzintervalle können Sie den Befehl `confint` verwenden.

Tragen Sie diese Konfidenzintervalle gemeinsam in einem Plot ab, der auch den wahren Wert von β_2 einhält. Ihr Ergebnis solle bei einem Randomseed von `1234` (abgesehen von Äußerlichkeiten) so aussehen::



Hinweise:

- Erzeugen Sie zunächst einen Plot, der nur die Achseninformationen enthält (`plot`-Option: `type=''n''`). Dazu müssen Sie wissen, welche Ausmaße die Achsen haben sollen.
- Zeichnen Sie das wahre β_2 ein (`abline`). Sie können auch Text in Ihrer Graphik platzieren. Mit `expression` können Sie beispielsweise auch Formeln oder griechische Buchstaben erzeugen. Um β_2 einzutragen, können Sie die Zeile `text(<x-Koord.>, <y-Koord.>, expression(beta[2]))` verwenden.
- Rufen Sie die Einträge der Matrix, die die Grenzen der Konfidenzintervalle enthält, auf und zeichnen Sie damit für jedes Intervall eine Linie (`lines`) von der linken zur rechten Grenze.

(b) (2 Punkte) Interpretieren Sie die Graphik.

(c) (1 Punkt) Um die Zufälligkeit der Konfidenzintervalle noch besser sehen zu können, entfernen Sie den Randomseed und lassen Sie Ihr Programm mehrmals hintereinander durchlaufen.

(d) (1 Punkt) Was passiert, wenn Sie α verändern?

(e) (2 Punkte) Was passiert, wenn Sie die Beobachtungszahl erhöhen?

49. Aufgabe (22 Punkte) (*Momentenschätzer und (F)GLS*)
Gegeben sei das korrekt spezifizierte Modell

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

sowie eine Matrix $\mathbf{W}$ mit $\mathbf{u}|\mathbf{X},\mathbf{W} \sim (\mathbf{0},\boldsymbol{\Omega})$ und $\boldsymbol{\Omega} \neq \sigma^2\mathbf{I}$.
Der allgemeine Momentenschätzer für die Matrix $\mathbf{W}$ wird bestimmt aus

$$\mathbf{W}^T(\mathbf{y} - \mathbf{X}\boldsymbol{\beta}) \stackrel{!}{=} \mathbf{0}.$$

Er ist der Prototyp für viele Schätzer in der Ökonometrie, wie zum Beispiel der Instrumentvariablenschätzer. In den kommenden zwei Teilbereichen werden theoretische und angewandte Eigenschaften des GLS-Schätzers erläutert.

(a) Theorie:

i. (1 Punkt) Zeigen Sie, dass sich aus obiger Bedingung der Momentenschätzer

$$\hat{\boldsymbol{\beta}}_{\mathbf{W}} = (\mathbf{W}^T\mathbf{X})^{-1}\mathbf{W}^T\mathbf{y}$$

ergibt. Wann ist $\hat{\boldsymbol{\beta}}_{\boldsymbol{W}} = \hat{\boldsymbol{\beta}}_{OLS}$?

ii. (2 Punkte) Leiten Sie die bedingte Kovarianz-Matrix des Momentenschätzers $\hat{\boldsymbol{\beta}}_{\mathbf{W}}$,

$$\mathrm{Var}(\hat{\boldsymbol{\beta}}_{\mathbf{W}}|\mathbf{X},\mathbf{W}) = (\mathbf{W}^T\mathbf{X})^{-1}\mathbf{W}^T\boldsymbol{\Omega}\mathbf{W}(\mathbf{X}^T\mathbf{W})^{-1},$$

her.

iii. (2 Punkte) Zeigen Sie, dass sich für $\mathbf{W} = \boldsymbol{\Omega}^{-1}\mathbf{X}$ der GLS-Schätzer

$$\hat{\boldsymbol{\beta}}_{GLS} = \left(\mathbf{X}^T\boldsymbol{\Omega}^{-1}\mathbf{X}\right)^{-1}\mathbf{X}^T\boldsymbol{\Omega}^{-1}\mathbf{y}$$

ergibt und bestimmen Sie mit Hilfe von Teil (ii) die bedingte Kovarianz-Matrix $\mathrm{Var}(\hat{\boldsymbol{\beta}}_{GLS}|\mathbf{X})$ des GLS-Schätzers.

iv. (2 Punkte) Zeigen Sie, dass $\hat{\boldsymbol{\beta}}_{\mathbf{W}}$ ein unverzerrter Schätzer für $\boldsymbol{\beta}$ ist. Warum ist damit auch der GLS-Schätzer mit $\mathbf{W} = \boldsymbol{\Omega}^{-1}\mathbf{X}$ unverzerrt?

v. (2 Punkte) Vergleichen Sie die statistischen Eigenschaften von $\hat{\boldsymbol{\beta}}_{GLS}$ mit denen des KQ-Schätzers für obiges Modell. Nehmen Sie dabei an, dass $\boldsymbol{\Omega}$ bekannt ist.

(b) Praxis: Laden Sie den Datensatz `Hartnagel` aus dem `R` Package `car`:

```
library(car)
data(Hartnagel)
```

i. (3 Punkte) Erklären Sie alle vorkommenden Variablen, schätzen Sie das Modell

$$\text{fconvict}_t = \beta_1 + \beta_2\text{tfr}_t + \beta_3\text{partic}_t + \beta_4\text{degrees}_t + \beta_5\text{mconvict}_t + u_t, t = 1, \ldots, T, \quad (12)$$

mit KQ und erklären Sie die zugrundeliegende Frage. Wieso wurde die Variable mconvict_t aufgenommen? Interpretieren Sie die Schätzung kurz.

ii. (2 Punkte) Visualisieren Sie die Residuen obiger Schätzung, was fällt Ihnen auf? Betrachten Sie außerdem die (partielle) Autokorrelationsfunktion (Befehl `acf`). Zu welchem Ergebnis kommen Sie?

iii. (2 Punkte) Installieren sie die Packages `Matrix`, `lme4` und `nlme`. Schätzen Sie obiges Modell mit der `gls` Funktion mit Hilfe von Maximum Likelihood (`method = "ML"`) und verschiedenen Korrelationsstrukturen. Die ML-Schätzung ist ein sehr gängiges, alternatives Schätzverfahren zu KQ, welches in Applied Financial Econometrics oder in Fortgeschrittener Ökonometrie besprochen wird. Für welches Modell entscheiden Sie sich?

iv. (2 Punkte) Vergleichen Sie nun die Schätzung von (12) mit dem FGLS geschätzten Modell (mit der AR(2) Spezifikation). Was fällt auf? Für welche Schätzung entscheiden Sie sich?

(c) Algorithmik: Entwickeln Sie nun selbst einen Schätzer für eine solche FGLS-AR(1) Schätzung.

i. (2 Punkte) Betrachten Sie das Modell

$$y_t = \boldsymbol{X}_t\beta + u_t, \quad t = 1, \ldots T, \quad (13)$$

wobei $\mathrm{E}[u_t \mid \boldsymbol{X}] = 0, \mathrm{E}[u_t^2 \mid \boldsymbol{X}] = \sigma_u^2$ und $\mathrm{E}[u_t u_s] \neq 0$ für $s \neq t$, insbesondere also $\boldsymbol{u} \mid \boldsymbol{X} \sim (\boldsymbol{0}, \boldsymbol{\Omega})$ mit $\boldsymbol{\Omega}$ keine Diagonalmatrix, gilt. Nehmen Sie an, dass der Fehler u_t einem kovarianzstationären AR(1)-Prozess folgt, d. h.

$$u_t = \rho u_{t-1} + \epsilon_t \quad \text{mit } \epsilon_t \mid u_t \sim IID(0, \sigma_\epsilon^2), |\rho| < 1.$$

Berechnen Sie $\boldsymbol{\Omega} := \mathrm{E}[\boldsymbol{u}\boldsymbol{u}^T \mid \boldsymbol{X}]$.
Hinweis: Berechnen Sie $\mathrm{E}[u_t^2 \mid \boldsymbol{X}]$ und $\mathrm{E}[u_t u_s \mid \boldsymbol{X}]$ einzeln.

ii. (2 Punkte) Führen Sie nun eine FGLS Schätzung von Hand durch, d. h. schätzen Sie (13) mit KQ um die Residuen $\hat{u}_t^{KQ}$ zu erhalten. Regressieren Sie $\hat{u}_t^{KQ}$ gegen $\hat{u}_{t-1}^{KQ}$ um ρ zu schätzen und konstruieren Sie im Anschluss den FGLS Schätzer

$$\hat{\boldsymbol{\beta}}^{FGLS} := (\boldsymbol{X}^T \hat{\boldsymbol{\Omega}}^{-1} \boldsymbol{X})^{-1} \boldsymbol{X}^T \hat{\boldsymbol{\Omega}}^{-1} \boldsymbol{y},$$

wobei ρ durch die Schätzung $\hat{\rho}$ ersetzt wird. Berechnen Sie außerdem die Standardfehler. Erklären Sie im Anschluss noch kurz, wieso $\hat{\boldsymbol{\beta}}^{FGLS}$ durch dieses Verfahren konsistent geschätzt werden kann.
Hinweis: Nehmen Sie an, dass (B1), (A1) und (A2) gelten.

50. Aufgabe (8 Punkte) (*Stochastische Prozesse*)

Betrachten Sie den stochastischen Prozess $y_t = x_t - 1$, $t = 1, 2, ...$, wobei die Dichte von x_t gegeben ist als

$$f_{x_t}(x) = \begin{cases} e^{-x} & \text{für } t \text{ ungerade, } x \geq 0, \\ 0, & \text{für } t \text{ ungerade, } x < 0 \\ \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-1)^2}, & \text{für } t \text{ gerade.} \end{cases}$$

Für die gemeinsame Dichte von x_t und x_s, $s \neq t$ gilt $f_{x_t,x_s}(y,z) = f_{x_t}(y) f_{x_s}(z)$.

Gegeben sei auch der Prozess z_t, der sich ergibt aus

$$z_t = \cos(tw) \quad \text{mit} \quad f(w) = \begin{cases} \frac{1}{2\pi}, & \text{für } x \in [0; 2\pi] \\ 0, & \text{sonst,} \end{cases} \qquad t = 1, 2, 3, ...$$

Prüfen Sie jeweils, ob y_t bzw. z_t

(a) (2 Punkte) *i.i.d.*,

(b) (2 Punkte) streng stationär,

(c) (2 Punkte) schwach stationär,

(d) (2 Punkte) Weißes Rauschen sind.

Hinweis: Für die explizite Berechnung der Integrale (sofern nötig) benutzen Sie geeignete Software, etwa Maple. Sie können diese auch numerisch bestimmen, indem Sie die den Integrationsbereich $[a; b]$ in n gleich lange Teilintervalle $[x_{i-1}; x_i]$ der Länge h zerlegen und für kleine h die Approximation $\int_a^b f(x)dx \approx \sum_{i=1}^n f(x_i)h$ verwenden. In R geht das etwa so:

```
a <- 0; b <- 4
f <- function(x) x^2
n <- 100000
x_i <- seq(a,b,length.out=n)
int <- sum(f(x_i)*(b-a)/n)
```

51. Aufgabe (18 Punkte) (*Markov-Ketten*)

In dynamischen makroökonomischen Modellen spielen diskrete Markovprozesse eine wichtige Rolle. Hierbei springt die Ausprägung eines Zufallsprozesses über die Zeit hinweg zwischen endlich vielen Zuständen (z.B. hohe, mittle und geringe Produktivität), wobei die Wahrscheinlichkeitsverteilung jeweils nur vom Zustand in der Vorperiode abhängt; vgl. etwa Kapitel 2 in Ljungqvist und Sargent (2004). Man stellt nun einen solchen dar als n-Vektor-Prozess $\{\boldsymbol{x}_t\}_{t=0,1,2,...}$, der als Werte $\boldsymbol{e}_j$, $j = 1, 2, ..., n$, n-Vektoren mit 1 an der j-ten Position und 0 sonst, annehmen kann. Man bezeichne mit π_0 den Vektor der unbedingten Wahrscheinlichkeiten für den Startzustand zum Zeitpunkt 0. Allgemein ist das i-te Element von π_t definiert als

$$\pi_{ti} = Prob(\boldsymbol{x}_t = \boldsymbol{e}_i), \quad \text{wobei} \sum_{i=1}^{n} \pi_{ti} = 1.$$

Die $(n \times n)$ Transitionsmatrix $\boldsymbol{P}$ enthält alle Übergangswahrscheinlichkeiten vom i-ten Zustand in den j-ten, $i, j = 1, 2, ..., n$. Also gilt

$$P_{ij} = Prob(\boldsymbol{x}_{t+1} = \boldsymbol{e}_j | \boldsymbol{x}_t = \boldsymbol{e}_i) \quad \text{mit} \quad \sum_{j=1}^{n} P_{ij} = 1.$$

Betrachten Sie nun

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \boldsymbol{P} = \begin{pmatrix} 0.4 & 0.6 \\ 0.9 & 0.1 \end{pmatrix} \quad \text{und} \quad \pi_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

(a) (3 Punkte) Berechnen Sie die unbedingten Verteilungen π_t für $t = 1, 2$. Finden Sie eine allgemeine Lösung für π_t in Abhängigkeit von π_0 und $\boldsymbol{P}$. Ist $\boldsymbol{x}_0$, $\boldsymbol{x}_1$, ... ein streng stationärer Prozess?

(b) (4 Punkte) Finden Sie einen alternativen Vektor der Anfangsverteilung π_0^*, der für die Folge der unbedingten Verteilungen π_t^* impliziert, dass $\pi_0^* = \pi_t^*$ für alle $t = 1, 2, ...$ Zeigen Sie, dass hiermit die gemeinsame Verteilung von $(\boldsymbol{x}_t, \boldsymbol{x}_{t+i})$ nur von i, nicht aber von t abhängt.

(c) (3 Punkte) Nun soll der univariate Prozess $y_t = \boldsymbol{a}'\boldsymbol{x}_t$ untersucht werden. Hier ist $\boldsymbol{a} = (-8\ 2)'$ ein fester Vektor. Unter Verwendung von π^* aus der letzten Aufgabe (nehmen Sie $\pi^* = (0.4\ 0.6)'$): Berechnen Sie $\mathrm{E}[y_t]$, $\mathrm{E}[y_t | y_{t-1} = 2]$ und $\mathrm{E}[\mathrm{E}[()y_t | y_{t-1})]$.

(d) (4 Punkte) Ermitteln Sie $\mathrm{Cov}\,(()\, y_t, y_{t+1})$.

(e) (4 Punkte) Ein weiterer Zustand soll jetzt möglich sein:

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \boldsymbol{P} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0.4 & 0.6 \\ 0 & 0.9 & 0.1 \end{pmatrix} \text{ und } \pi_0 = \begin{pmatrix} 0.4 \\ 0.6 \\ 0 \end{pmatrix}.$$

Gibt es hier eine eindeutige stationäre Verteilung $\pi^* = \pi_t^*$? Ist dieser Prozess ergodisch?

52. Aufgabe (4 Punkte) (*Markov-Ungleichung*)

Wählen Sie 2 diskrete oder stetige Wahrscheinlichkeitsverteilungen (etwa Poisson, Exponential, Normal, t, F oder Gleichverteilung; Wahrscheinlichkeits- oder Dichtefunktionen finden Sie z.B. in Wikipedia) und prüfen Sie, ob die Markov-Ungleichung

$$P(|X| \geq \varepsilon) \leq \frac{\mathrm{E}[|X|^p]}{\varepsilon^p}$$

für selbstgewählte Parameterwerte, $\varepsilon = 2, 4$ und $p = 2, 4, 6$ gilt.

(a) (4 Punkte) Kann man diese jeweils als gute Approximation für die Wahrscheinlichkeiten betrachten?

Hinweis: Die Berechnung kann in Maple durchgeführt werden. Als Alternative können Sie auch den R-Code für numerische Integration (n groß genug wählen) und die in R enthaltenen Tabellierungen der Dichte- und Verteilungsfunktionen verwenden. Etwa `pnorm`, `pf`, `ppois`, `pchisq`, `pbinom` (in der Hilfe nachschauen, auch für weitere Verteilungen) liefern die Verteilungsfunktion, während `dnorm` etc. die Dichte berechnet.

53. Aufgabe (9 Punkte) (*Definitionen in der Zeitreihentheorie*)

Sei $\epsilon_t \sim NID(0,1), t = 0, 1, 2, ..., y_0 = 0$ und seien $a, b \in \mathbb{R}_0^+$. Sind die folgenden stochastischen Prozesse $\{y_t\}_{t=0}^{\infty}$

- Martingal-Differenzenfolgen bezüglich der Informationsmenge, die durch ϵ_0 und den letzten Beobachtungen von y_t erzeugt wurde?
- Weißes Rauschen?
- streng stationär?
- kovarianzstationär?

(a) (2 Punkte) $y_t = a + b\epsilon_0$.

(b) (2 Punkte) $y_t = t\epsilon_t$.

(c) (2 Punkte) $y_t = \epsilon_t + \epsilon_{t-1}^2 - 1$.

(d) (2 Punkte) $y_t = \epsilon_t \epsilon_{t-1}$.

Diese Teilaufgabe soll unabhängig oben genannter Prozesse behandelt werden:

(e) (1 Punkt) Erklären Sie mit eigenen Worten, was das Ergodentheorem besagt.

54. Aufgabe (10 Punkte) (*Yule-Walker und Anwendung auf einen AR(2)-Prozess*)

Im Folgenden wollen Sie das Theorem von Yule-Walker beweisen:

Die Autokovarianzfunktion $\gamma(k)$ eines kovarianzstationären AR(p)-Prozesses $\{y_t\}_{t=0}^{\infty}$ mit Erwartungswert 0

$$y_t = \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \epsilon_t = \sum_{i=1}^{p} \phi_i y_{t-i} + \epsilon_t \quad \text{where } \epsilon_t \sim IID(0, \sigma^2) \tag{14}$$

lässt sich schreiben als

$$\gamma(k) = \begin{cases} \phi_1\gamma(1) + \cdots + \phi_p\gamma(p) + \sigma^2 & \text{if } k = 0 \\ \phi_1\gamma(k-1) + \cdots + \phi_p\gamma(k-p) & \text{if } k > 0 \end{cases}$$

Gehen Sie dabei folgendermaßen vor:

(a) (1 Punkt) Benutzen Sie den Verschiebungssatz für Cov (y_t, y_{t+k}) und vereinfachen Sie den Ausdruck.

(b) (2 Punkte) Multiplizieren Sie nun (14) mit y_{t+k} für $k < 0$ auf beiden Seiten und nehmen Sie den Erwartungswert.

(c) (2 Punkte) Wiederholen Sie die Prozedur von (b) mit $k = 0$.

Betrachten Sie folgenden AR(2)-Prozess:

$$y_t = \phi_1 y_{t-1} - \phi_2 y_{t-2} + \epsilon_t = 1.2 y_{t-1} - 0.5 y_{t-2} + \epsilon_t \quad \text{mit} \quad \epsilon_t \sim IID(0, \sigma^2) \tag{15}$$

(d) (2 Punkte) Ist der Prozess (15) stabil? Begründung.

(e) (1 Punkt) Nehmen Sie nun an, dass (15) bereits unendlich lange läuft. Sind die Voraussetzungen für die Yule-Walker-Gleichungen erfüllt? Begründung.

(f) (2 Punkte) Nutzen Sie die Yule-Walker-Gleichungen um alle Autokovarianzen bzw. die Autokovarianzfunktion von (15) zu berechnen.

55. Aufgabe (10 Punkte) (*ARMA-Prozesse*)

Gegeben sei der ARMA(1,4) Prozess

$$x_t = \alpha_1 x_{t-1} + \phi_1 u_{t-1} + \phi_4 u_{t-4} + u_t, \quad u_t \sim WN(0, \sigma^2), \quad t \in \mathbb{Z}.$$

(a) (2 Punkte) Unter welchen Voraussetzungen an α_1, ϕ_1 und ϕ_4 ist der Prozess stabil? Hat er in jedem Fall eine AR(∞)-Darstellung? Gehen Sie im Folgenden davon aus, dass die jeweiligen Bedingungen erfüllt sind.

(b) (2 Punkte) Berechnen Sie die ersten 5 Koeffizienten ψ_j aus der unendlichen MA darstellung

$$x_t = \rho_0 u_t + \rho_1 u_{t-1} + \rho_2 u_{t-2} + \ldots = \sum_{j=0}^{\infty} \rho_j u_{t-j} = \rho(L) u_t$$

Leiten Sie eine Formel ab, mit der Sie die weiteren Koeffizienten rekursiv berechnen können.

Hinweis: Die Parameter lassen sich durch Koeffizientenvergleich berechnen.

(c) (2 Punkte) Benutzen Sie die berechnete MA-Darstellung, um die Autokovarianzfunktion von x_t für die ersten Lags zu berechnen.
Hinweis: Benutzen Sie `R` mit den Werten $\sigma = 1$, $\alpha_1 = 0.9$, $\phi_1 = -0.6$, $\phi_4 = 0.7$.

(d) (2 Punkte) Generieren Sie mit `R` einige Realisierungen des obigen ARMA(1,4) Prozesses. Verwenden Sie die vorher angegebenen Parameterwerte. Dazu können Sie zum Beispiel die folgenden Befehle verwenden und für den hier betrachteten Prozess anpassen:

```
model <- list(ar=c(1.5,-.6),ma=c(-0.5))
n <- 240
x <- arima.sim(model,n,n.start=200)
ts.plot(x)
```

Alternativ kann man Schleifen benutzen:

```
n <- 240
u <- rnorm(n)
x <- rep(0,n)
for (t in 3:n) x[t] <- 1.5 * x[t-1] - 0.6 * x[t-2] - 0.5 * u[t-1] + u[t]
ts.plot(x)
```

Beschreiben Sie die Trajektorien für verschiedene Längen. Sind die so generierten Prozesse jeweils stationär? Welche Unterschiede bestehen?

(e) (2 Punkte) Schätzen Sie jeweils die Autokovarianzfunktion für verschiedene Stichprobengrößen. Verwenden Sie dazu die `acf` Funktion mit geeigneten Optionen. Vergleichen Sie die Werte mit Ihren Ergebnissen von vorher.

56. Aufgabe (7 Punkte) (*Zeitreihenschätzung in R*)
Verwenden Sie den Datensatz zum ifo-Geschäftsklimaindex auf der Homepage.
Stellen Sie im den im Anhang folgenden Code das Arbeitsverzeichnis auf den Ordner, in dem der Datensatz liegt und gehen Sie Zeile für Zeile durch den Code, um die folgenden Fragen zu beantworten.

(a) (1 Punkt) Wieso steht beim Einlesen der Daten `startRow=20`? Beschreiben Sie zudem den Datensatz.

(b) (1 Punkt) Was ändert sich durch die `ts()`-Funktion? Was bedeuten die Parameter

```
start = c(1991, 1), end = c(2012,10), frequency = 12
```

Hinweis: `?ts()` bzw. `?class()`.

(c) (1 Punkt) Welches Modell wird in `klima_ar1` abgespeichert?

(d) (1 Punkt) Beschreiben Sie, was die Funktion `stepAIC` macht.

(e) (2 Punkte) Was wird im Abschnitt `(P)ACF` berechnet? Beschreiben Sie beide Konzepte kurz und geben Sie die Definitionen an.

(f) (1 Punkt) Was macht der Befehl `sim.arima`?

```
# setwd("<Arbeitsverzeichnis>")
# install.packages("xlsx")
library(xlsx) # Einlesen von Dateien im Format xls oder xlsx

# Einlesen der Daten
ifo_index <- read.xlsx("ifo-geschaeftsklima_d_10_2012.xla",
                       sheetIndex = 1, colIndex = c(2,3,4),startRow=20,
                       colClasses = c("numeric","numeric","numeric"))

# In R gibt es das Time Series Object (ts), zu dem man diesen Datensatz konvertieren muss
ifo_ts <- ts((ifo_index [-1,1:3]),
             start = c(1991, 1), end = c(2012,10), frequency = 12,
             names = c("Geschaftsklima", "Geschaftsbeurteilung","Geschaftserwartungen"))
plot(ifo_ts) # Zeitreihen plotten

# AR-Modell schätzen; in R sehr viele Möglichkeiten, analog zu EViews: dynlm Package
# install.packages("dynlm")
library(dynlm)
klima_ar1 <- dynlm(Geschaftsklima ~ L(Geschaftsklima,1), data=ifo_ts)
summary(klima_ar1) # Üblicher R-Output

# Beliebiger AR-Prozess, z. B. 2ter Ordnung
klima_ar2 <- dynlm(Geschaftsklima ~ L(Geschaftsklima,1)+L(Geschaftsklima,2), data=ifo_ts)
# alternativ: dynlm(Geschaftsklima ~ L(Geschaftsklima,1:2), data=ifo_ts)

# Einzeichnen der gefitteten Werte
plot(ifo_ts[,2], xlim=c(1990, 2015)) # Graph mit den Rohdaten
lines(klima_ar2\$fit, col="red", lwd=2)

# Lag-Wahl mit Informationskriterien:
# Idee: Schätze Modell mit vielen Lags und verkürze mittels AIC (stepAIC)
klima_ar4 <-dynlm(Geschaftsklima ~ L(Geschaftsklima,1)+L(Geschaftsklima,2)+L(Geschaftsklima,3)
                  +L(Geschaftsklima,4), data=ifo_ts)
klima_aic <- stepAIC(klima_ar4)
summary(klima_aic)

# (P)ACF
acf(ifo_ts[,2],main="ACF of Geschäftsklima") # 1. ACF
pacf(ifo_ts[,2],main="PACF of Geschäftsklima") # 2. PACF

# Andere Untersuchungsmöglichkeiten: Histogramm mit Kerndichteschätzung
hist(ifo_ts[,2], prob=TRUE, breaks=50)
lines(density(ifo_ts[,2]), col="blue")

# Simulation von 100 Realisationen eines AR(1) Prozesses mit ar=0.5
plot(arima.sim(n=100, list(order=c(1,0,0), ar=0.9,sd=1)))
```

57. Aufgabe (8 Punkte) (*Lineare Prozesse*)

Betrachten Sie (für vorerst festes n und eine positive Konstante $C \in \mathbb{N}$) den stochastischen Prozess

$$x_t(n) = \sum_{j=0}^{n} \phi_j u_{t-j}, \quad u_t \sim WN(0, \sigma^2), \quad \phi_j = \begin{cases} 1, \; j < C, \\ \psi^{(j-C)}, \; j \geq C. \end{cases}$$

(a) (4 Punkte) Berechnen Sie den Erwartungswert und die Varianz von $x_t(n)$. Bestimmen Sie auch
Cov $(x_t(n), x_{t-1}(n))$. Hängen diese jeweils von n und/oder t ab?

(b) (4 Punkte) Für welche Werte von ψ ist $x_t(1)$, $x_t(2)$, $x_t(3)$, ... eine Cauchy-Folge bezüglich der L^2-Norm? Hierzu muss

$$\lim_{n,m\to\infty} ||x_t(n) - x_t(m)||^2 = \lim_{n,m\to\infty} \mathrm{E}[(x_t(n) - x_t(m))^2] = 0$$

gelten. Illustrieren Sie Ihr Ergebnis graphisch.

58. Aufgabe (28 Punkte) (*Schwache Exogenität*)

In dieser Aufgabe wird die Definition der schwachen Exogenität zuerst mithilfe einer bivariaten Normalverteilung wiederholt, um sie dann in einem zweiten Schritt auf ein einfaches Modell anzuwenden.

(a) Betrachten Sie die übliche Produktschreibweise einer bivariaten Dichte der Normalverteilung in eine bedingte und eine marginale Dichte:

$$f_{X,Y}(x, y; \Psi) = f_{Y|X}(y \mid x; \Psi) \cdot f_X(x; \Psi), \qquad \begin{pmatrix} y \\ x \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_y \\ \mu_x \end{pmatrix}, \begin{pmatrix} \sigma_y^2 & \sigma_{yx}^2 \\ \sigma_{xy}^2 & \sigma_y^2 \end{pmatrix}\right)$$

Wir nennen nun y die abhängige Variable und x den Regressor, wobei die bedingte Dichte $f_{Y|X}(y \mid x)$ die typische Idee der Regressionsgleichung verkörpert. Falls nun x eine (schwach) exogene Variable ist, so ist man in der Lage die marginale Dichte $f_X(x)$ zu ignorieren, um inferenzielle Rückschlüsse über die Parameter der bedingten Dichte zu treffen (vgl. Definition der schwachen Exogenität).

i. (5 Punkte) Berechnen Sie β_1 und β_2 der Parametrisierung von

$$\mu_{Y|X} := \mathrm{E}[Y \mid X = x] = \beta_1 + \beta_2 x$$

in Abhängigkeit der gegebenen Parameter $\Psi = \{\mu_y, \mu_x, \sigma_y^2, \sigma_{yx}^2, \sigma_y^2\}$.
Hinweis:

- Berechnen Sie $f_{Y|X}(y \mid x) := \frac{f_{X,Y}(x,y)}{f_X(x)}$ und bilden davon den Erwartungswert.
- Obige Formel wird auch als *regression lemma* bezeichnet.

Man definiert nun allgemein die Störterme der bedingten und marginalen Dichten als

$$\epsilon_t := y_t - \mu_{Y|X} \qquad \epsilon_t \sim N(0, \sigma^2) \tag{16}$$

$$\nu_t := x_t - \mu_x \qquad \nu_t \sim N(0, \sigma_x^2) \tag{17}$$

ii. (4 Punkte) Zeigen Sie, dass $\operatorname{Cov}(\epsilon_t, \nu_t) = 0$ und $\sigma^2 = \sigma_y^2(1-\rho^2) = \sigma_{xy} - \frac{\sigma_{xy}^2}{\sigma_x}$ gilt.

Die Faktorisierung der bivariaten Verteilung partitionierte den Parametervektor Ψ in zwei Teilvektoren $\Psi_1 = \{\alpha, \beta, \sigma^2\}$ und $\Psi_2 = \{\mu_X, \sigma_X^2\}$, womit obige Faktorisierung folgendermaßen geschrieben werden kann

$$f_{X,Y}(x, y; \Psi) = f_{Y|X}(y \mid x; \Psi_1) \cdot f_X(x; \Psi_2)$$

und Ψ_1 als $g(\Psi_1) \subset \Psi$ nicht als Teilmenge, sondern als injektive Abbildung aufgefasst werden soll.

iii. (2 Punkte) Erklären Sie die Definition der schwachen Exogenität an obigem Beispiel.

(b) Betrachten Sie als Anwendung nun ein sogenanntes *Cobweb*-Modell in einem landwirtschaftlichen Umfeld, indem der Preis p_t vom kontemporären Angebot q_t abhängt und das Angebot vom vorangegangen Preis. Die daraus resultierenden, strukturellen Gleichungen sind:

$$p_t = \beta_1 q_t + \epsilon_t \qquad \epsilon_t \sim N(0, \sigma^2) \tag{18}$$

$$q_t = \pi_q p_{t-1} + \nu_{qt} \qquad \nu_{qt} \sim N(0, \sigma_q^2) \tag{19}$$

Die Variablen p_t und q_t wurden hierbei logarithmiert und danach mittelwertbereinigt, um die Konstanten aus dem Modell entfernen zu können.
Der Wert $1/\beta$ kann als die Preiselastizität der Nachfrage interpretiert werden und π_q als die Elastizität des Angebots.

i. (4 Punkte) Nehmen Sie an, dass $\operatorname{Cov}(\epsilon_t, \nu_{qt}) = 0$ gilt und folgern Sie daraus, dass $\operatorname{Cov}(q_t, \nu_{qt}) = 0$. Zeigen Sie, dass diese Annahme impliziert, dass q_t exogen bzgl. (18) ist. Interpretieren Sie kurz, inwieweit obige Annahme sinnvoll ist.

ii. (4 Punkte) Die Variablen p_t und q_t haben auch eine gemeinsame Verteilung mit konstanten Mittelwerten, welche die folgende, genannt reduzierte, Form von (18) und (19) motiviert:

$$p_t = \pi_p p_{t-1} + \nu_{pt} \qquad \nu_{pt} \sim N(0, \sigma_p^2) \tag{20}$$

$$q_t = \pi_q p_{t-1} + \nu_{qt} \qquad \nu_{qt} \sim N(0, \sigma_q^2) \tag{21}$$

Hierbei wird angenommen, dass $\operatorname{Cov}(\nu_{pt}, \nu_{qt}) = \sigma_{pq} \neq 0$. Zeigen Sie, indem Sie $\mathrm{E}[p_t \mid q_t]$ berechnen, dass

$$p_t = (\pi_p - \beta\pi_q)p_{t-1} + \beta q_t + \epsilon_t \tag{22}$$

iii. (3 Punkte) Vergleichen Sie (18) mit (22) und gehen Sie auf den Fall ein, wenn Sie nur an $\beta \in \Psi_1$ interessiert sind.

iv. (3 Punkte) Es ist nun sinnvoll anzunehmen, dass die Störterme dämpfenden Einfluss auf die Zyklen in Preise und Mengen haben, was eine Restriktion auf (20) notwendig macht: $|\pi_p| < 1$. Setzen Sie nun (19) in (22) ein und zeigen Sie wieder die Bedingung von π_p und π_q.

In den beiden vorangegangenen Aufgaben haben Sie gesehen, dass $\pi_p = \beta\pi_q$ und dass die Parameter $\beta \in \Psi_1$ und $\pi_q \in \Psi_2$ eine funktionale Abhängigkeit bilden. Weiß man nun den Wert von π_q, so schränkt man die möglichen Werte von β deutlich ein, weswegen in Systemen, in denen die dynamische Stabilität eine notwendige Annahme ist, p_{t-1} nicht mehr shwach exogen für β ist.

v. (3 Punkte) Definieren sie für den Fall q_t und den Fehlern ϵ_t die Begriffe „vorherbestimmt“ und „strenge Exogenität“. Was wäre notwendig, um ϵ_t streng exogen bzgl. q_t zu machen?

59. Aufgabe (14 Punkte) (*Exogenität im bivariaten Modell*)

Zu lesen: Davidson (2000, Kapitel 4 und 5)

Betrachten Sie das folgende bivariate ökonometrische Modell (vgl. Davidson, 2000, S. 81)

$$y_t + \alpha_{12} z_t = \gamma_1 + \beta_{11} y_{t-1} + \beta_{12} z_{t-1} + \sigma_1 u_{1t} \tag{23}$$

$$z_t = \gamma_2 + \beta_{21} y_{t-1} + \sigma_2 u_{2t}, \qquad \boldsymbol{u}_t \overset{iid}{\sim} N(\boldsymbol{0}, \boldsymbol{I}) \tag{24}$$

(a) (3 Punkte) Stellen Sie das System in Matrixschreibweise als

$$\boldsymbol{B}\boldsymbol{x}_t = \boldsymbol{c} + \boldsymbol{C}\boldsymbol{x}_{t-1} + \boldsymbol{u}_t^*$$

dar. Geben Sie die jeweiligen Matrizen explizit an. Bringen Sie es dann in die Form

$$\boldsymbol{x}_t = \boldsymbol{a} + \boldsymbol{A}\boldsymbol{x}_{t-1} + \boldsymbol{\varepsilon}_t.$$

Wie sind hier $\boldsymbol{a}$, $\boldsymbol{A}$ und $\boldsymbol{\varepsilon}_t$ definiert?

(b) (4 Punkte) Bestimmen Sie die gemeinsame Dichtefunktion

$$f_{\boldsymbol{x}|\boldsymbol{x}_{-1}}(\boldsymbol{x}_t|\boldsymbol{x}_{t-1}; \boldsymbol{\psi}) = f_{y,z|y_{-1},z_{-1}}(y_t, z_t|y_{t-1}, z_{t-1}; \boldsymbol{\psi})$$

$\boldsymbol{\psi}$ ist hier der Vektor der Parameter α_{12}, γ_1, β_{11}, β_{12}, σ_1, γ_2, β_{21} und σ_2.

(c) (4 Punkte) Sie seien zunächst am Parameter α_{12} interessiert. Bestimmen Sie eine Partition der Modellparameter $\boldsymbol{\psi}_1, \boldsymbol{\psi}_2$, um zu einem konditionalen Modell für $y_t|z_t, z_{t-1}, y_{t-1}$ in Abhängigkeit des Parametervektors $\boldsymbol{\psi}_1$ zu gelangen. Leiten Sie dazu die (bedingte) Dichtefunktion

$$f_{y|z,z_{-1},y_{-1}}(y_t|z_t, y_{t-1}, z_{t-1}; \boldsymbol{\psi}_1)$$

her. (vgl. Davidson, 2000, S.75f)

Was passiert mit Ihrem Ergebnis, falls $\gamma_1 = \gamma_2$ gilt?

(d) (3 Punkte) Sie wollen nun z_t prognostizieren und verwenden hierfür den bedingten Erwartungswert $\mathrm{E}[z_t|y_t, y_{t-1}, z_{t-1}] = \delta_0 + \delta_1 y_t + \delta_2 y_{t-1} + \delta_3 z_{t-1}$. Bestimmen Sie δ_j, $j = 0, 1, 2, 3$ explizit in Abhängigkeit der Modellparameter. Ist hier eine Zerlegung der bedingten Dichte möglich?

60. Aufgabe (21 Punkte) (*Dynamisches Lohn-Philipps Modell*)

Ein dynamisches Modell der Lohnsetzung lässt sich darstellen als (vgl. Bardsen et al. 2005, S.48)

$$\Delta w_t = \beta_{w0} - \beta_{w1} u_t + \beta_{w2} \Delta a_t + \beta_{w3} \Delta q_t + \beta_{w4} \Delta p_t + \varepsilon_{w,t} \tag{25}$$

wobei (alle Variablen in natürlichen Logarithmen) w_t den Nominallohn, u_t die Arbeitslosenquote, a_t Arbeitsproduktivität, q_t Produzentenpreise und p_t den Konsumentenpreisindex bezeichnen. Es gelte

$$\varepsilon_{w,t} \sim NID(0, \sigma_w^2).$$

Es sei nun ferner eine Gleichung für die Bestimmung der Arbeitslosigkeit gegeben als

$$\Delta u_t = \beta_{u0} - \beta_{u1} u_{t-1} + \beta_{u2}(w - q - a)_{t-1} - \boldsymbol{\delta}' \boldsymbol{z}_t + \varepsilon_{u,t}, \quad \varepsilon_{u,t} \sim NID(0, \sigma_w^2) \tag{26}$$

Die CPI-Inflationsrate erfüllt die Gleichung

$$\Delta p_t = \beta_{p1}(\Delta w_t - \Delta a_t) + \beta_{p2} \Delta q_t + \varepsilon_{p,t}, \quad \varepsilon_{p,t} \sim NID(0, \sigma_p^2)$$

(a) (2 Punkte) Erläutern Sie die ökonomische Intuition hinter den Gleichungen. Welche Annahmen müssen Sie treffen, damit (25) als valides Modell für

$$\mathrm{E}[w_t | w_{t-1}, u_t, u_{t-1}, ..., a_t, a_{t-1}, ..., q_t, q_{t-1}, ..., p_t, p_{t-1}, ..., \boldsymbol{z}_t; \beta_{wj}]$$

angesehen werden kann? Nennen Sie gegebenenfalls Einwände gegen die getroffenen Annahmen.

(b) (3 Punkte) In den jährlichen Daten für Norwegen (siehe Passwortgeschützter Bereich) sind für $\boldsymbol{x}_t$ noch weitere Determinanten der Lohnsetzung enthalten, etwa die Log-Lohnersatzrate rpr_t sowie die Lohnsteuer im Industriesektor in Levels, $t1_t$, und die Veränderungen der täglichen Arbeitsstunden Δh_t. Lesen Sie die norwegischen Daten (Passwortgeschützter Bereich) unter dem Namen `data` in `R` ein. Wandeln Sie sie dann mit

```
dat.ts <- ts(data,start=1954)
```

in eine Zeitreihe um und stellen sie graphisch dar. Was verbirgt sich hinter den $i1967$ und IP Variablen?

Erstellen Sie dann einen `data.frame`, in den Sie auch verzögerte Variablen miteinbeziehen. Dies kann etwa wie folgt geschehen:

```
PCM.dat <- ts.intersect(
d.w = diff(dat.ts[,"w"]),
d.p =  diff(dat.ts[,"p"]),
d.p.l1 =  lag(diff(dat.ts[,"p"]),-1),
d.p.l2 =  lag(diff(dat.ts[,"p"]),-2),
d.a = diff(dat.ts[,"a"]),
d.a.l1 =  lag(diff(dat.ts[,"a"]),-1),
d.a.l2 =  lag(diff(dat.ts[,"a"]),-2),
```

```
d.q = diff(dat.ts[,"q"]),
d.q.l1 =  lag(diff(dat.ts[,"q"]),-1),
d.q.l2 =  lag(diff(dat.ts[,"q"]),-2),
u = dat.ts[,"u"],
u.l1 = lag(dat.ts[,"u"],-1),
d.t1 = diff(dat.ts[,"t1"]),
rpr.l1 = lag(dat.ts[,"rpr"],-1),
d.h = diff(dat.ts[,"h"]),
i1967 = dat.ts[,"i1967"],
IP = dat.ts[,"IP"],dframe=TRUE)
```

Um alle Beobachtungen mit fehlenden Werten zu entfernen, kann der Befehl

```
PCM.dat <- na.omit(PCM.dat)
```

verwendet werden.

(c) (4 Punkte) Schätzen Sie nun die Parameter der Phillipskurve (25). Plotten Sie die Residuen als Zeitreihe. Schätzen Sie die Autokorrelationsfunktion. Ist eine der Autokorrelationen signifikant von 0 verschieden? Sind wichtige Variablen in der Regression weggelassen worden? Welches Licht werfen die Ergebnisse auf die in a) genannten Annahmen?

(d) (4 Punkte) Betrachten Sie nun als mögliches Modell

$$\Delta w_t - \Delta p_{t-1} = \beta_0 + \beta_2 \Delta p_{t-1} + \beta_7 \Delta q_t + \beta_8 \Delta q_{t-1} + \beta_{10} u_t + \gamma_5 IP_t + \varepsilon_t$$

Interpretieren Sie die geschätzten Koeffizienten. Prüfen Sie, ob (auch) hier eine Fehlspezifikation vorliegen könnte.

(e) (4 Punkte) Plotten Sie die rekursiven Schätzer: Schätzen Sie das Modell mit den Daten bis zum Zeitpunkt $t_0 = 1976$. Wiederholen Sie die Schätzung mit je einer weiteren Beobachtung und stellen Sie die Koeffizientenschätzer sowie Konfidenzintervalle in Abhängigkeit des Endzeitpunkts dar. Gibt es Hinweise auf Parameterinstabilität?

(f) (4 Punkte) Eine andere Interpretation des Zusammenhangs zwischen Lohn-/Preissetzung und der realen ökonomischen Aktivität ist die Lucas-Angebotskurve (vgl. etwa Romer, 2006, Kapitel 6 A). Diese stellt eine Inversion des in d) geschätzten Modells dar mit u_t als abhängiger Variable. Schätzen Sie auch diesen Zusammenhang. Ist dieses bedingte Modell stabil? Erläutern Sie die Ergebnisse. Gehen Sie dabei auch auf verschiedene Exogenitätskonzepte ein. (5 Punkte)

61. Aufgabe (9 Punkte) (*Dynamische Regression*)

Sie kommen zum Schluss, dass bei der Modellierung der Skiverkäufe in der Tat schwache Exogenität von x_t für die relevanten Parameter vorliegt, Sie aber wichtige dynamische Wirkungsmechanismen übersehen haben. Also betrachten Sie nun

$$y_t = \delta + \alpha_1 y_{t-1} + \beta_0 x_t + \beta_1 x_{t-1} + u_t, \quad u_t \sim WN(0, \sigma) \tag{27}$$

(a) (2 Punkte) Stellen Sie das Modell mit expliziter Angabe von $\alpha(L)$ und $\beta(L)$ als $\alpha(L)y_t = \delta + \beta(L)x_t + u_t$ dar.

(b) (3 Punkte) Unter welchen Voraussetzungen an $\alpha(L)$ und $\beta(L)$ gilt das "common factors"-Modell, das zu

$$y_t = \delta^+ + \beta^+ x_t + u_t^+$$

führt? Welchem Prozess folgt hier der Fehler u_t^+?

(c) (4 Punkte) Führen Sie die Regressionsgleichung (27) in die Fehlerkorrekturdarstellung

$$\alpha^{**}(L)\Delta y_t = \delta + \beta^{**}(L)\Delta x_t - \alpha(y_{t-1} - \theta x_{t-1}) + u_t$$

über. Berechnen Sie explizit $\alpha^{**}(L)$, $\beta^{**}(L)$, α und θ.

62. Aufgabe (8 Punkte) (*Homo- und Heteroskedastizität*)

Gegeben sei folgender Code, der ein simples lineares Regressionsmodell (SLR) im Fall von homo- und heteroskedastischen Daten schätzt.

```
set.seed(1)

N  = 100              # Number of Obervations
beta_0 = 2            # True Beta_0
beta_1 = 0.5          # True Beta_1

sigma2_homo = 5       # Homoskedastic Variance
p1 = 1.5              # Heteroskedastic Variance Process Parameter 1
p2 = 0.015            # Heteroskedastic Variance Process Parameter 2

x        = runif(N, min=0, max=100)
y_homo   = beta_0 + beta_1*x + rnorm(N, mean=0, sd=sqrt(sigma2_homo   ))
y_hete   = beta_0 + beta_1*x + rnorm(N, mean=0, sd=sqrt(exp(p1 + p2*x)))

model.homo = lm(y_homo~x)
model.hete = lm(y_hete~x)

summary(model.homo)
summary(model.hete)
```

(a) (2 Punkte) Erklären Sie in maximal zwei wohlüberlegten (!) Sätzen, was für eine Simulation hier durchgeführt wird. Spezifizieren Sie hierfür beide Modelle vollständig.

(b) (1 Punkt) Zeichnen Sie die beiden Scatterplots der homo- und heteroskedastischen `y_homo / y_hete` gegen `x` und verdeutlichen Sie hierbei die unterschiedlichen Varianzeigenschaften.

(c) (2 Punkte) Zeichnen Sie die beiden Scatterplots der homo- und heteroskedastischen Residuen gegen die gefitteten Werte. Erklären Sie die Unterschiede.

(d) (1 Punkt) Zeichnen Sie die beiden Scatterplots der homo- und heteroskedastischen quadrierten Residuen gegen die gefitteten Werte und vergleichen Sie diesen mit dem integrierten *Scale – Location*-Plot in `R`.

Hinweis: `plot(model.homo, which=3)` erzeugt diesen Plot im homoskedastischen Fall.

(e) (2 Punkte) Visualisieren Sie den *Residual – Leverage*-Plot in beiden Fällen. Machen Sie sich dabei klar, was *leverage* im ökonometrischen Sinn (insbesondere bei der KQ-Schätzung) und *Cook′s distance* bedeuten.

Hinweis: `plot(model.homo, which=5)` erzeugt diesen Plot im homoskedastischen Fall.

63. Aufgabe (16 Punkte) (*HC und HAC - Einführung*)

Diese Aufgabe befasst sich mit dem Artikel von Achim Zeileis: Econometric Computing with HC and HAC Covariance Matrix Estimators, Journal of Statistical Software, Volume 11, Issue 10, November 2004.

(a) (3 Punkte) Beschreiben Sie mit eigenen Worten, inwiefern HC- bzw. HAC-Standardfehlerschätzung eine empirische Analyse beeinträchtigen. Gehen Sie dabei folgendermaßen vor:

- Interpretieren Sie die Annahme aus der Vorlesung, die die Benutzung von HC- und HAC-Schätzern unnötig macht.
- Was ist das Problem, das entsteht, wenn diese Annahme nicht gilt?
- Welche Teile einer empirischen Analyse sind davon betroffen?

(b) (1 Punkt) Fassen Sie die Vorgehensweise in Zeileis' Artikel zusammen.

(c) (1 Punkt) Wie wird $\hat{\sigma}^2$ normalerweise geschätzt? Wann ist diese Schätzung konsistent?

(d) (2 Punkte) Welche Schätzer werden im Fall von Heteroskedastizität vorgeschlagen? Interpretieren Sie jeden Einzelnen.

(e) (2 Punkte) Formulieren Sie die Idee der HAC-Schätzung in eigenen Worten.

(f) (1 Punkt) Erklären Sie die Intuition hinter dem Gewichtsvektor $\omega = (\omega_0, ..., \omega_{n-1})^T$ bei einer HAC-Schätzung.

(g) (2 Punkte) Welche Gewichtsschätzer von ω werden vorgeschlagen? Interpretieren Sie alle bis auf den HAC-Schätzer von Andrews (1991), dessen Herangehensweise über die Spektraldichte einer Zeitreihe motiviert ist und über die vermittelte Technik dieses Kurses hinausgeht.

Mit den folgenden Fragen sollen HC- und HAC-Schätzern anhand von oben genanntem Artikel, Kapitel 4, implementiert werden. Abschnitt 4.3 ist zwar eine schöne Anwendung, führt aber über die Intention dieses Übungsblatts hinaus und wird deshalb nicht nachvollzogen.

(h) (2 Punkte) HC: Schätzen Sie das Modell

$$Expenditure_t = \beta_1 + \beta_2 Income_t + \beta_3 Income_t^2 + u_t \tag{28}$$

mit dem Datensatz `PublicSchools` des `R`-Packages `sandwich`. Nutzen Sie die Funktion `coeftest` mit Hilfe aller angegebenen HC-Schätzern und vergleichen Sie die Standardfehler und die Teststatistiken miteinander.

(i) (2 Punkte) HAC: Schätzen Sie das Modell

$$RealInv_t = \beta_1 + \beta_1 RealGNP_t + \beta_2 RealInt_t + \epsilon_t \tag{29}$$

mit dem Datensatz `Investment` des `R`-Packages `sandwich`.

64. Aufgabe (5 Punkte) (*Frisch-Waugh-Lovell bei Paneldaten*)

In einem Panel mit m Querschnitts- und T Längsschnittbeobachtungen enthalte der Fehler u_{it} für Individuum i, $i = 1, \dots, m$, und Zeitpunkt t, $t = 1, \dots, T$, einen individuenspezifischen Teil v_i und eine Restkomponente ε_{it}. Dabei sei v_i stochastisch unabhängig von allen ε_{jt}, $j = 1, \dots, m$, aber möglicherweise mit den erklärenden Variablen $\mathbf{X}$ korreliert.

(a) (5 Punkte) Zeigen Sie nun, dass die Anwendung des Frisch-Waugh-Lovell-Theorems auf

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{D}\boldsymbol{\eta} + \boldsymbol{\varepsilon}$$

und das Modell

$$y_{it} - \bar{y}_{i\cdot} = \beta_1(x_{it,1} - \bar{x}_{i\cdot,1}) + \cdots + \beta_k(x_{it,k} - \bar{x}_{i\cdot,k}) + \varepsilon_{it} - \varepsilon_{i\cdot}$$

zu gleichen Schätzergebnissen für $\boldsymbol{\beta}$ führen.

65. Aufgabe (4 Punkte) (*Verteilung des Fixed-Effects-Schätzers*)

(a) (3 Punkte) Zeigen Sie, dass für den Schätzer für $\boldsymbol{\beta}$ im Fixed Effects Modell

$$\hat{\boldsymbol{\beta}}_{FE}|\mathbf{X} \sim N(\boldsymbol{\beta}, \sigma^2(\mathbf{X}^T\mathbf{M}_{\mathbf{D}}\mathbf{X})^{-1})$$

gilt, wenn $\boldsymbol{\varepsilon}|\mathbf{X} \sim N(\mathbf{0}, \sigma^2\mathbf{I})$ und die sonstigen üblichen Annahmen gelten.

(b) (1 Punkt) Warum wird σ^2 mit $m(T-1) - k$ statt $mT - k$ Freiheitsgraden geschätzt?

66. Aufgabe (8 Punkte) (*Fehlerkorrelationen im Panel*)

Für die Fehler des Modells

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

gelte

$$u_{it} = v_i + \varepsilon_{it}, \quad i = 1, \dots, m, \ t = 1, \dots, T.$$

Dabei seien die v_i und die ε_{jt} für alle i, j, t unabhängig und $v_i \sim IID(0, \sigma_v^2)$, $\varepsilon_{it} \sim IID(0, \sigma_\varepsilon^2)$. Bestimmen Sie

(a) (2 Punkte) $\mathrm{Var}(u_{it}|\mathbf{X})$,

(b) (2 Punkte) $\mathrm{Cov}(u_{it}, u_{is}|\mathbf{X})$ für $s \neq t$,

(c) (2 Punkte) $\mathrm{Cov}(u_{it}, u_{js}|\mathbf{X})$ für $i \neq j$,

(d) (2 Punkte) und daraus $\mathrm{Var}(\mathbf{u}|\mathbf{X})$, wobei $\mathbf{u} = \begin{pmatrix} u_{11} & u_{12} & \cdots & u_{1T} & u_{21} & \cdots & u_{mT} \end{pmatrix}^T$.

67. Aufgabe (7 Punkte) (*Kausalität*)

Wiederholen Sie den entsprechenden Teil in der Vorlesung, um im Anschluss die folgenden Fragen beantworten zu können. Für Philosophie-Interessierte lohnt sich der Artikel des Ökonometrikers Arnold Zellner "Causality and Econometrics" (über die Uni abrufbar), dieser ist zur Beantwortung der Fragen nicht nowendig, mindestens jedoch hilfreich.

(a) (2 Punkte) Erklären Sie die Definition von „Kausalität“ aus der Vorlesung.

(b) (1 Punkt) Welche Art ökonometrischer Modellierung betrachten wir in diesem Kurs?

(c) (2 Punkte) Erklären Sie den Begriff „kontrolliertes Zufallsexperiment“ und nennen Sie eine typische, ökonometrische Fragestellung, die mit diesem untersucht werden könnte und gehen Sie dabei genauer auf den vorher erklärten Begriff ein.

(d) (2 Punkte) Schreiben Sie das simultane Gleichungsmodell (in struktureller Form) aus der Vorlesung in die reduzierte Form um.

68. Aufgabe (24 Punkte) (*Modellspezifikation*)

Mit den folgenden Zeilen laden Sie das EViews-Workfile wage2.wf1 in `R`.

```
setwd("<Arbeitsverzeichnis>")  # Ihr Arbeitsverzeichnis
install.packages("hexView")    #
library(hexView)               # Einlesen von Dateien im Format ".wf1"
data <- readEViews("wage2.wf1", as.data.frame = TRUE)
str(data)                      # Variablennamen sind groß geschrieben
```

Der Datensatz enthält 935 Beobachtungen von verschiedenen Variablen, die für die Schätzung einer monatlichen Lohnfunktion (`wage` in US-Dollar) nützlich sind. Dazu gehören die Ausbildungsdauer in Jahren (`educ`), der Intelligenzquotient (`iq`) in Punkten und die Anzahl an Berufsjahren beim aktuellen Arbeitgeber (`tenure`). In einem ersten Schritt sehen Sie das folgende lineare Regressionsmodell als geeignet an:

$$\ln wage_t = \beta_0 + \beta_1 tenure_t + \beta_2 iq_t + \beta_3 educ_t + u_t, \quad u_t | \mathbf{tenure}, \mathbf{iq}, \mathbf{educ} \sim IID(0, \sigma^2). \tag{30}$$

(a) (2 Punkte) Führen Sie eine KQ-Schätzung dieses Modells in `R` durch und interpretieren Sie kurz die Variable `educ`.

(b) (2 Punkte) Berechnen Sie das AIC Kriterium in `R` mit Hilfe einer geeigneten Funktion. Berechnen Sie es danach mit der Formel aus der Vorlesung und erklären Sie den Unterschied.

Ein Kollege schlägt nun das alternative Modell

$$\ln wage_t = \gamma_0 + \gamma_1 tenure_t + \gamma_2 iq_t + \epsilon_t \tag{31}$$

vor.

(c) (2 Punkte) Nehmen Sie nun an, dass Modellspezifikation (30) den richtigen Zusammenhang zur Erklärung der Variable `wage` angibt, es aber stattdessen Modell (31) geschätzt wurde. Erklären Sie <u>kurz</u>, welches Problem hier vorliegt, und wie es sich auf die KQ-Schätzung $\widehat{\gamma}_1$ von γ_1 im Modell (31) auswirkt.

(d) (5 Punkte) Leiten Sie nun die „omitted variables bias"-Formel ab, d. h. zeigen Sie, dass mit der Annahme von (c) gilt:

$$\mathrm{E}[\widehat{\gamma}_1 \mid \mathbf{tenure}, \mathbf{iq}, \mathbf{educ}] = \beta_{0,11} + \beta_{0,2} \frac{\widehat{\mathrm{Cov}}(\mathbf{tenure}, \mathbf{educ})}{\widehat{\mathrm{Var}}(\mathbf{tenure})}$$

Hinweis:

- Definieren Sie $\boldsymbol{X}_1 = (\mathbf{1}, \mathbf{tenure}, \mathbf{iq})$ und $\boldsymbol{X}_2 = (\mathbf{educ})$ und wiederholen Sie die Eigenschaften des KQ-Schätzers in fehlspezifizierten Modellen aus der Vorlesung.
- Im simplen linearen Regressionsmodell $\boldsymbol{y} = \boldsymbol{x}\tau + \boldsymbol{u}$ gilt: $\hat{\tau} = \frac{\widehat{\mathrm{Cov}}(\boldsymbol{y}, \boldsymbol{x})}{\widehat{\mathrm{Var}}(\boldsymbol{x})}$

(e) (2 Punkte) Schätzen Sie nun in `R` Modell (31), berechnen Sie den Ausdruck $\hat{\beta}_1 + \hat{\beta}_3 \frac{\widehat{\mathrm{Cov}}(\mathbf{tenure}, \mathbf{educ})}{\widehat{\mathrm{Var}}(\mathbf{tenure})}$ und vergleichen Sie das Ergebnis mit $\widehat{\gamma}_1$. Interpretieren Sie die Verzerrungsformel kurz.

(f) (2 Punkte) Geben Sie anhand obigen Beispiels an, wie Paneldaten benutzt werden könnten um bestimmte Arten des „omitted variable bias“ auszuschließen.

Ein weiterer Kollege schlägt nun ein größeres Modell

$$\ln wage_t = \delta_0 + \delta_1 tenure_t + \delta_2 iq_t + \delta_3 educ_t + \delta_4 age_t + v_t \tag{39}$$

vor.

(g) (2 Punkte) Nehmen Sie nun wieder an, dass Modellspezifikation (30) den richtigen Zusammenhang zur Erklärung der Variable `wage` angibt. Erklären Sie kurz, welches Problem hier vorliegt, und wie es numerisch sich auf die KQ-Schätzung $\widehat{\delta_1}$ von δ_1 im Modell (39) auswirkt.

(h) (2 Punkte) Wählen Sie nun in `R` aus allen drei Modellen mit Hilfe des angepassten R^2 und der Informationskriterien AIC, HQ, SC das jeweils beste Modell aus.
Hinweis: Sie können die `R`-Funktion `SelectCritEViews` aus der Vorlesung benutzen.

(i) (2 Punkte) Vergleichen Sie nun die drei Informationskritieren hinsichtlich ihrer Tendenzen über- bzw. fehlspezifizierte Modelle zu wählen. Wie würden Sie sich entscheiden, würden AIC und HQ verschiedene Modelle bevorzugen?

(j) (3 Punkte) Gegeben der Annahme, dass Modellspezifikation (30) den richtigen Zusammenhang zur Erklärung der Variable `wage` angibt, geben Sie einen kurzen Überblick über die asymptotischen Schätzeigenschaften des KQ-Schätzers in Modell (31) und (39).