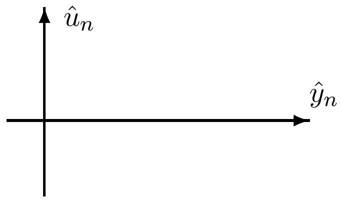
(a) Erläutern Sie, wieso sich das sog. *Tukey-Anscombe Diagramm*



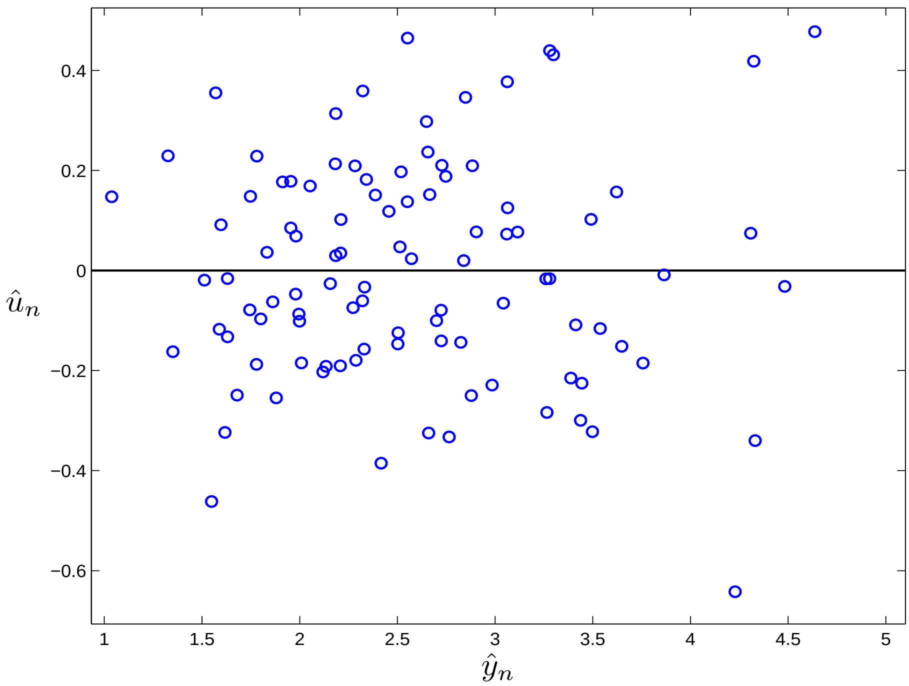
zur graphischen Modelldiagnose eignet! Auf welcher Eigenschaft von $\hat{y}_n$ und $\hat{u}_n$ basiert die Analysefähigkeit des Tukey-Anscombe Diagramms?

Lösungsvorschlag: Wie wir bereits in Aufgabenblatt 1 und 5 gesehen haben, spielen die statistischen Eigenschaften der Störgröße $\boldsymbol{u}$ eine bedeutsame Rolle bei der Entscheidung, ob und wie gut die gewählte Spezifikation die Annahmen des klassischen, linearen Regressionsmodells erfüllt. Obwohl die Modellspezifikation $\boldsymbol{X\beta}$ durch den Ökonometriker vorgegeben und damit bekannt ist, bleiben die statistischen Eigenschaften der Störgröße $\boldsymbol{u}$ unbeobachtbar. Die Analyse der statistischen Eigenschaften von $\boldsymbol{u}$ ist allerdings unabdingbar, da sich nur hieraus die Güte der Spezifikation $\boldsymbol{X\beta}$ und die Wahl der Schätzmethodik beurteilen lassen. Als eine praktikable und gangbare Approximation gilt es, die Störgröße u_n durch das Residuum $\hat{u}_n$ zu ersetzen. Viele statistische Testverfahren machen sich diese Approximation zunutze. Um die statistischen Eigenschaften von $\hat{\boldsymbol{u}}$ graphisch zu visualisieren, benötigt man eine Bezugsgröße. Im Tukey-Anscombe Diagramm werden die aus der KQ-Schätzung resultierenden Residuen $\hat{u}_n$ auf der Ordinate und die aus der KQ-Schätzung resultierenden Projektionen $\hat{y}_n$ auf der Abszisse abgetragen. Alternativ könnte man auch $\hat{u}_n$ gegen einen beliebigen Regressor plotten. Dies läßt sich, wie folgt, begründen: Sind die Annahme der Linearität und die Annahme eines vollen Spaltenrangs von $\boldsymbol{X}$ richtig, dann gelten die Orthogonalitätsbedingungen der KQ-Schätzung $\boldsymbol{X}'\hat{\boldsymbol{u}} = \boldsymbol{0}_{k\times 1}$. Hieraus folgt direkt

$$\hat{\boldsymbol{u}}'\hat{\boldsymbol{y}} = (\hat{\boldsymbol{u}}'\hat{\boldsymbol{y}})' = \hat{\boldsymbol{y}}'\hat{\boldsymbol{u}} = (\boldsymbol{X}\hat{\boldsymbol{\beta}}_{KQ})'\hat{\boldsymbol{u}} = \hat{\boldsymbol{\beta}}'_{KQ}\boldsymbol{X}'\hat{\boldsymbol{u}} = 0 \,,$$

d.h. es liegt kein lineare Zusammenhang zwischen Residuen und Projektion vor. Trotz $\mathrm{Cov}[\,\hat{u}_n, \hat{y}_n\,] = 0$ ist jedoch Heteroskedastizität der Form $\mathrm{Var}[\,\hat{u}_n\,] = f(\hat{y}_n) \neq \sigma^2 = \text{const}$ möglich. Gelten stattdessen jedoch die Annahmen des klassischen, linearen Regressionsmodells, $\mathrm{E}[\,\boldsymbol{u}\,] = \boldsymbol{0}_{N\times 1}$ und $\mathrm{E}[\,\boldsymbol{u}\boldsymbol{u}'\,] = \sigma^2\boldsymbol{I}_N$, dann hat das Tukey-Anscombe Diagramm das idealtypisches Aussehen aus Abbildung 1, bei dem weder das erste noch das zweite Moment von $\hat{u}_n$ von $\hat{y}_n$ abhängen.

Abbildung 1: *Tukey-Anscombe Diagramm für eine KQ-Schätzung mit Störgrößenvarianz* $\mathrm{Var}[\,u_n\,] = \sigma_n^2 = \sigma^2$.



(b) Für eine Stichprobe aus Querschnittsdaten wird die Regressionsgleichung

$$y_n = \beta_1 + \beta_2 x_{n2} + \beta_3 x_{n3} + u_n$$

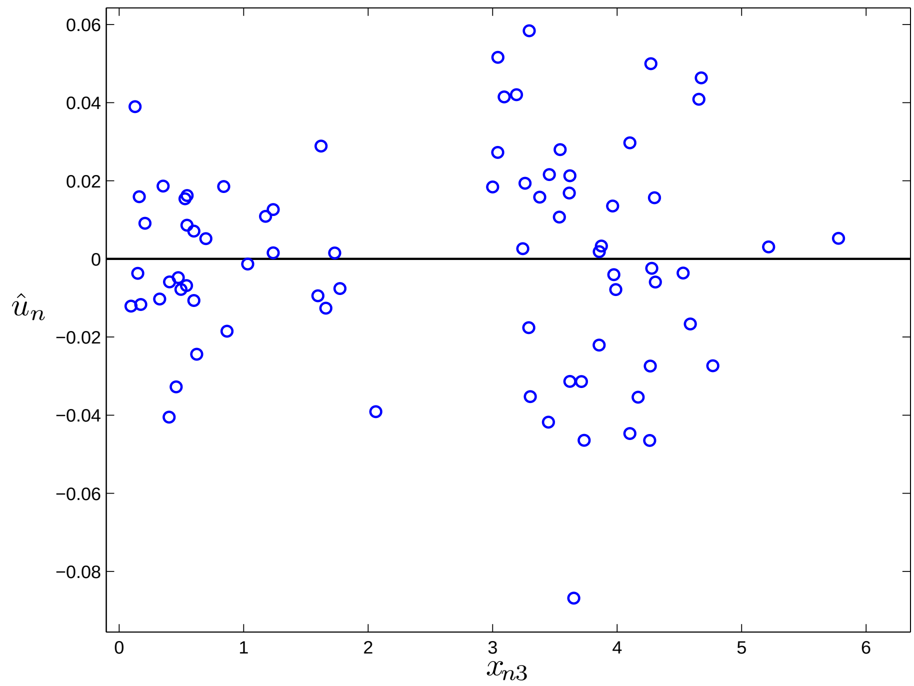
mit $n = 1, \ldots, 76$ geschätzt. Werden die aus der KQ-Schätzung resultierenden Residuen $\hat{u}_n$ gegen die entsprechenden Werte des dritten Regressors x_{n3} abgetragen, erhält man Abbildung 2.

Welche Eigenschaft der vorliegenden Stichprobe vermuten Sie? Beschreiben Sie ein geeignetes Verfahren, um Ihre Vermutung zu testen!

Lösungsvorschlag: Aufgrund des in Teilaufgabe (a) beschriebenen Zusammenhangs zwischen $\hat{\boldsymbol{u}}$ und $\boldsymbol{x}_3$ läßt sich vermuten, dass die Annahme homoskedastischer Störgrößen $\boldsymbol{u}$ verletzt ist. Offensichtlich lassen sich die Beobachtungen der Querschnittsdaten in zwei Gruppen mit unterschiedlicher Störgrößenvarianz unterteilen. Die Variable $\boldsymbol{x}_3$ scheint dabei als die sog. *Threshold-Variable* zu fungieren. Der Test von Goldfeld und Quandt (1965) ist für solche Situationen konzipiert. Er lässt sich für das vorliegende Problem wie folgt durchführen:

1) Aufgrund des Diagramms lassen sich alle Beobachtungen des Datensatzes mittels der Ausprägung der Variablen $\boldsymbol{x}_3$ in eine der beiden Gruppen mit σ_1^2 oder σ_2^2 einordnen, wobei $\sigma_1^2 < \sigma_2^2$ gelten soll. Jede Beobachtung des Querschnittsdatensatzes besteht aus dem Tripel (y_n, x_{n2}, x_{n3}), so dass bei der Sortierung nach x_{n3} automatisch auch y_n und x_{n2} umsortiert

Abbildung 2: *Tukey-Anscombe Diagramm.*



werden. Nach der Sortierung erhält man zwei Teilstichproben. Die erste Teilstichprobe ist charakterisiert durch σ_1^2 und subsumiert N_1 Beobachtungspunkte in $\boldsymbol{y}_1$ und $\boldsymbol{X}_1$. Die zweite Teilstichprobe ist charakterisiert durch σ_2^2 und subsumiert N_2 Beobachtungspunkte (mit $N = N_1 + N_2$) in $\boldsymbol{y}_2$ und $\boldsymbol{X}_2$.

2) Man schätzt die Modelle $\boldsymbol{y}_1 = \boldsymbol{X}_1\boldsymbol{\beta}_1 + \boldsymbol{u}_1$ für die erste Teilstichprobe und $\boldsymbol{y}_2 = \boldsymbol{X}_2\boldsymbol{\beta}_2 + \boldsymbol{u}_2$ für die zweite Teilstichprobe. Danach berechnet man die entsprechenden Summen der quadrierten Residuen $\hat{\boldsymbol{u}}_1'\hat{\boldsymbol{u}}_1$ und $\hat{\boldsymbol{u}}_2'\hat{\boldsymbol{u}}_2$.

3) Der Test lautet dann

$$\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2 \qquad \text{vs.} \qquad \mathrm{H}_1 : \sigma_1^2 < \sigma_2^2 \,.$$

Ein geeigneter Ausgangspunkt für eine diesbezügliche Teststatistik ist der Quotient σ_2^2/σ_1^2, welcher bei Homoskedastizität gleich 1 sein sollte. Ersetzt man σ_1^2 und σ_2^2 durch erwartungstreue Schätzer erhält man die Teststatitik

$$F := \frac{\hat{\sigma}_2^2}{\hat{\sigma}_1^2} = \frac{\hat{\boldsymbol{u}}_2'\hat{\boldsymbol{u}}_2/(N_2 - k)}{\hat{\boldsymbol{u}}_1'\hat{\boldsymbol{u}}_1/(N_1 - k)} = \frac{\hat{\boldsymbol{u}}_2'\hat{\boldsymbol{u}}_2}{\hat{\boldsymbol{u}}_1'\hat{\boldsymbol{u}}_1}\,\frac{N_1 - k}{N_2 - k} \,.$$

Wie bereits in Aufgabenblatt 2 gesehen, sind Teststatistiken, die auf einen Vergleich zwischen zwei Summen quadrierter Residuen abzielen, F-verteilt. Demnach ist F unter H_0 F-verteilt mit $N_2 - k$ Zählerfreiheitsgraden und $N_1 - k$ Nennerfreiheitsgraden.

4) H_0 wird zum Signifikanzniveau α abgelehnt, wenn $F > F_{krit}$, wobei der kritische Wert F_{krit} definiert ist als $F^{-1}_{N_2-k;N_1-k}(1 - \alpha)$.

(c) Führen Sie den unter Teilaufgabe (b) beschriebenen Test zum Signifikanzniveau 5% mit $\hat{\boldsymbol{u}}_1'\hat{\boldsymbol{u}}_1 = 0.0102$ für $n = 1, \ldots, 33$ und $\hat{\boldsymbol{u}}_2'\hat{\boldsymbol{u}}_2 = 0.0428$ für $n = 34, \ldots, 76$ durch! Kann diese Testprozedur verbessert werden?

Lösungsvorschlag:

Hypothese:

$$\mathrm{H}_0 : \sigma_1^2 = \sigma_2^2 \qquad \text{vs.} \qquad \mathrm{H}_1 : \sigma_1^2 < \sigma_2^2 \,.$$

Teststatistik:

$$F = \frac{\hat{\boldsymbol{u}}_2'\hat{\boldsymbol{u}}_2}{\hat{\boldsymbol{u}}_1'\hat{\boldsymbol{u}}_1}\,\frac{N_1 - k}{N_2 - k} = \frac{0.0428}{0.0102}\,\frac{33-3}{43-3} = 3.1471 \,.$$

Kritischer Wert:

$$F_{krit} = F^{-1}_{40;30}(0.95) = 1.792 \,.$$

Testentscheidung: H_0 wird abgelehnt, da $F_{krit} = 1.792 < 3.1471 = F$, d.h. es liegen keine homoskedastischen Störgrößen vor.

Die Trennschärfe des Goldfeld-Quandt Tests kann erhöht werden, wenn Beobachtungen eliminiert werden, die das Vorliegen unterschiedlicher Varianzen kaschieren. Eine Faustregel schlägt vor, das mittlere Drittel der sortierten Beobachtungen zu eliminieren. Jedoch trägt dieses Drittel im vorliegenden Datensatz sowohl zur Bestimmung von $\hat{\sigma}_1^2$ als auch zur Bestimmung von $\hat{\sigma}_2^2$ und somit auch zur Kontrastierung der beiden Teilstichprobenvarianzen bei. Damit wäre dieser Schritt wohl mit einem Verlust an Information und Schätzpräzision verbunden. Diese Faustregel sollte demnach flexibel und der Situation entsprechend angewendet werden.

(d) Zeigen Sie, dass bei Vorliegen heteroskedastischer Störgrößen

$$(\boldsymbol{y} - \boldsymbol{X\beta})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X\beta}) = \sum_{n=1}^{N}\left(\frac{u_n}{\sigma_n}\right)^2 \qquad \text{mit} \quad \boldsymbol{\Sigma} \equiv \mathrm{E}\left[\boldsymbol{uu}'\right] = \sigma^2\boldsymbol{\Omega}$$

gilt! Welche Interpretation lässt diese Beziehung zu?

Lösungsvorschlag: Zunächst führt man die Störgrößen auf der linken Seite der Gleichung ein,

$$(\boldsymbol{y} - \boldsymbol{X\beta})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X\beta}) = \boldsymbol{u}'\boldsymbol{\Sigma}^{-1}\boldsymbol{u} \,,$$

um u_n überhaupt ins Spiel zu bringen. Letztlich muss man nur noch die Elemente dieses Vektor-Matrix-Produkts betrachten und berechnen:

$$\boldsymbol{u}'\boldsymbol{\Sigma}^{-1}\boldsymbol{u} = \begin{pmatrix} u_1 & u_2 & \cdots & u_N \end{pmatrix} \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_N^2 \end{bmatrix}^{-1} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{pmatrix}$$

$$= \begin{pmatrix} u_1 & u_2 & \cdots & u_N \end{pmatrix} \begin{bmatrix} \frac{1}{\sigma_1^2} & & & 0 \\ & \frac{1}{\sigma_2^2} & & \\ & & \ddots & \\ 0 & & & \frac{1}{\sigma_N^2} \end{bmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{pmatrix}$$

$$= \begin{pmatrix} \frac{u_1}{\sigma_1^2} & \frac{u_2}{\sigma_2^2} & \cdots & \frac{u_N}{\sigma_N^2} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_N \end{pmatrix}$$

$$= \frac{u_1^2}{\sigma_1^2} + \frac{u_2^2}{\sigma_2^2} + \cdots + \frac{u_N^2}{\sigma_N^2} = \sum_{n=1}^{N} \left(\frac{u_n}{\sigma_n} \right)^2 .$$

Da

$$(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) = \boldsymbol{u}'\boldsymbol{u} = \sum_{n=1}^{N} u_n^2$$

der Summe der quadrierten Störgrößen entspricht und die Zielfunktion des KQ-Schätzers in der klassischen linearen Regression ist, kann somit die Zielfunktion des VKQ-Schätzers im Fall von heteroskedastischen Störgrößen

$$\begin{aligned} \boldsymbol{u}^{*'}\boldsymbol{u}^* &= (\boldsymbol{P}\boldsymbol{u})'(\boldsymbol{P}\boldsymbol{u}) = \boldsymbol{u}'\boldsymbol{P}'\boldsymbol{P}\boldsymbol{u} = \boldsymbol{u}'\boldsymbol{\Sigma}^{-1}\boldsymbol{u} \\ &= (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) = \sum_{n=1}^{N} \left(\frac{u_n}{\sigma_n} \right)^2 \end{aligned}$$

als gewichtete Summe der quadrierten Störgrößen aufgefasst werden.

(e) Zeigen Sie, dass bei Vorliegen heteroskedastischer Störgrößen

$$(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} = \left(\sum_{n=1}^{N} \sigma_n^{-2} \boldsymbol{x}_n \boldsymbol{x}_n' \right)^{-1} \sum_{n=1}^{N} \sigma_n^{-2} \boldsymbol{x}_n y_n \qquad \text{mit} \quad \boldsymbol{\Sigma} \equiv \mathrm{E}\left[\, \boldsymbol{u}\boldsymbol{u}' \,\right]$$

gilt! Welche Interpretation lässt diese Beziehung zu?

Lösungsvorschlag: Um diese Beziehung zu beweisen, geht man am besten in zwei Schritten vor. Als Erstes zeigt man

$$(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1} = \left(\sum_{n=1}^{N} \sigma_n^{-2} \boldsymbol{x}_n \boldsymbol{x}_n' \right)^{-1} .$$

Offensichtlich benötigt man eine geeignete Zerlegung der Regressormatrix $\boldsymbol{X}$ durch den Vektor $\boldsymbol{x}_n$. Dabei hilft es, sich über die Dimension der folgenden Objekte Klarheit zu verschaffen:

$$\underbrace{\boldsymbol{x}_n}_{k\times 1} = \begin{pmatrix} 1 \\ x_{n2} \\ \vdots \\ x_{nk} \end{pmatrix} \qquad \underbrace{\boldsymbol{X}}_{N\times k} = \begin{bmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \vdots \\ \boldsymbol{x}_N' \end{bmatrix} \qquad \underbrace{\boldsymbol{X}'}_{k\times N} = \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_N \end{bmatrix} .$$

Dann gilt

$$\begin{aligned}
(\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1} &= \left[\begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_N \end{bmatrix} \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_N^2 \end{bmatrix}^{-1} \begin{bmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \vdots \\ \boldsymbol{x}_N' \end{bmatrix} \right]^{-1} \\
&= \left[\begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_N \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_1^2} & & & 0 \\ & \frac{1}{\sigma_2^2} & & \\ & & \ddots & \\ 0 & & & \frac{1}{\sigma_N^2} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \vdots \\ \boldsymbol{x}_N' \end{bmatrix} \right]^{-1} \\
&= \left[\begin{bmatrix} \frac{\boldsymbol{x}_1}{\sigma_1^2} & \frac{\boldsymbol{x}_2}{\sigma_2^2} & \cdots & \frac{\boldsymbol{x}_N}{\sigma_N^2} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_1' \\ \boldsymbol{x}_2' \\ \vdots \\ \boldsymbol{x}_N' \end{bmatrix} \right]^{-1} \\
&= \left[\frac{\boldsymbol{x}_1\boldsymbol{x}_1'}{\sigma_1^2} + \frac{\boldsymbol{x}_2\boldsymbol{x}_2'}{\sigma_2^2} + \cdots + \frac{\boldsymbol{x}_N\boldsymbol{x}_N'}{\sigma_N^2} \right]^{-1} = \left[\sum_{n=1}^{N} \sigma_n^{-2} \boldsymbol{x}_n \boldsymbol{x}_n' \right]^{-1} .
\end{aligned}$$

Zuletzt beweist man, dass

$$\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} = \sum_{n=1}^{N} \sigma_n^{-2} \boldsymbol{x}_n y_n .$$

Es gilt

$$\begin{aligned}
\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} &= \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_N \end{bmatrix} \begin{bmatrix} \sigma_1^2 & & & 0 \\ & \sigma_2^2 & & \\ & & \ddots & \\ 0 & & & \sigma_N^2 \end{bmatrix}^{-1} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix} \\
&= \begin{bmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \cdots & \boldsymbol{x}_N \end{bmatrix} \begin{bmatrix} \frac{1}{\sigma_1^2} & & & 0 \\ & \frac{1}{\sigma_2^2} & & \\ & & \ddots & \\ 0 & & & \frac{1}{\sigma_N^2} \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}
\end{aligned}$$

$$= \begin{bmatrix} \frac{\boldsymbol{x}_1}{\sigma_1^2} & \frac{\boldsymbol{x}_2}{\sigma_2^2} & \cdots & \frac{\boldsymbol{x}_N}{\sigma_N^2} \end{bmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{pmatrix}$$

$$= \frac{\boldsymbol{x}_1 y_1}{\sigma_1^2} + \frac{\boldsymbol{x}_2 y_2}{\sigma_2^2} + \cdots + \frac{\boldsymbol{x}_N y_N}{\sigma_N^2} = \sum_{n=1}^{N} \sigma_n^{-2} \boldsymbol{x}_n y_n \, .$$

Zusammengenommen lässt sich somit

$$\hat{\boldsymbol{\beta}}_{VKQ} = (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} = \left(\sum_{n=1}^{N} \sigma_n^{-2} \boldsymbol{x}_n \boldsymbol{x}_n' \right)^{-1} \sum_{n=1}^{N} \sigma_n^{-2} \boldsymbol{x}_n y_n$$
$$= \left(\sum_{n=1}^{N} \frac{\boldsymbol{x}_n}{\sigma_n} \frac{\boldsymbol{x}_n'}{\sigma_n} \right)^{-1} \sum_{n=1}^{N} \frac{\boldsymbol{x}_n}{\sigma_n} \frac{y_n}{\sigma_n}$$

als ein gewichteter KQ-Schätzer interpretieren.

(f) Für eine Stichprobe von Querschnittsdaten wird die Regressionsgleichung

$$y_n = \beta_1 + \beta_2 x_{n2} + \beta_3 x_{n3} + u_n \qquad \text{mit} \quad n = 1, \ldots, 300$$

geschätzt. Werden die aus der KQ-Schätzung resultierenden Residuen $\hat{u}_n$ gegen die entsprechenden Werte der Projektion $\hat{y}_n$ abgetragen, erhält man Abbildung 3. Welche Eigenschaft der vorliegenden Stichprobe vermuten Sie?

Beschreiben Sie ein möglichst allgemeines Verfahren, um Ihre Vermutung zu testen!

Lösungsvorschlag: Der Test von Breusch and Pagan (1979) ist ein sehr flexibles Verfahren, um allgemeine Formen von Heteroskedastizität zu testen. Sei z.B. das Regressionsmodell

$$y_n = \boldsymbol{x}_n'\boldsymbol{\beta} + u_n \qquad \text{für} \quad n = 1, \ldots, N$$

und $\boldsymbol{x}_n = (1 \;\; x_{n2} \;\; \cdots \;\; x_{nk})'$ gegeben. Die Störgröße u_n habe folgende Eigenschaften: $\mathrm{E}[\, u_n \,] = 0$ und eine flexible Varianzfunktion

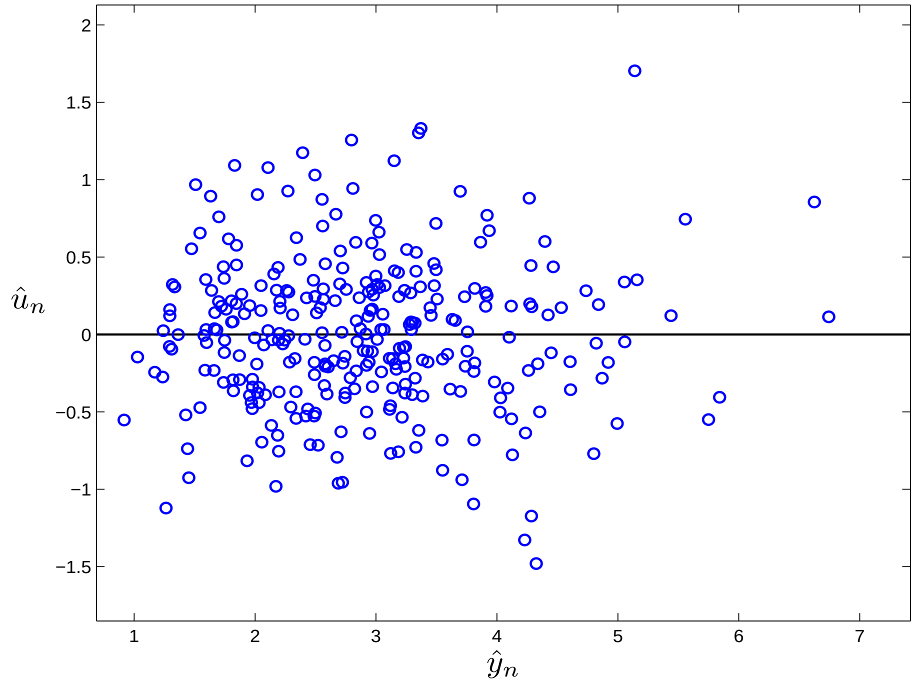
$$\sigma_n^2 := \mathrm{E}\left[\, u_n^2 \,\right] = h(\boldsymbol{z}_n'\boldsymbol{\alpha}) \, ,$$

mit $\boldsymbol{z}_n = (1 \;\; z_{n2} \;\; \cdots \;\; z_{np})'$ und $\boldsymbol{\alpha} = (\alpha_1 \;\; \alpha_2 \;\; \cdots \;\; \alpha_p)'$. Die invariante Varianzfunktion $h(\cdot)$ kann viele Formen von Heteroskedastizität abbilden, aber auch Homoskedastizität $\sigma_n^2 = h(\alpha_1)$, für $\alpha_2 = \cdots = \alpha_p = 0$. Letzteres kann als Nullhypothese für das Fehlen von Heteroskedastizität verwendet werden:

$$\mathrm{H}_0 : \alpha_2 = \cdots = \alpha_p = 0 \qquad \text{vs.} \qquad \mathrm{H}_1 : \exists\, j = 2, \ldots, p : \; \alpha_j \neq 0 \, .$$

Als Testverfahren haben Breusch und Pagan folgende Schritte vorgeschlagen:

Abbildung 3: *Tukey-Anscombe Diagramm.*



(1) Die KQ-Schätzung des linearen Modells $y_n = \boldsymbol{x}_n'\boldsymbol{\beta} + u_n$ und die Berechnung von $\hat{u}_n$ und $\hat{\sigma}^2 = \sum_{n=1}^{N} \hat{u}_n^2/N$.

(2) Die KQ-Schätzung der Hilfsregression

$$v_n = \boldsymbol{z}_n'\boldsymbol{\alpha} + \epsilon_n \qquad \text{mit} \quad v_n = \frac{\hat{u}_n^2}{\hat{\sigma}^2}$$

und Berechnung der erklärten Streuung

$$SQE = \sum_{n=1}^{N} (\hat{v}_n - \bar{v})^2 \ .$$

(3) Unter H_0 und der Annahme, dass u_n normalverteilt ist, gilt

$$\frac{1}{2} SQE \xrightarrow{d} \chi^2(p-1) \ .$$

(4) H_0 wird zum Signifikanzniveau α abgelehnt, wenn $\frac{1}{2}SQE > \chi^2_{krit}$, wobei der kritische Wert χ^2_{krit} definiert ist als das Quantil $F^{-1}_{p-1}(1-\alpha)$.

Eine asymptotisch äquivalente Teststatistik ist $N \cdot R^2$ ($\frac{1}{2}SQE \overset{asy}{\sim} NR^2$), wobei R^2 das multiple Bestimmtheitsmaß der Hilfsregression $v_n = \boldsymbol{z}_n'\boldsymbol{\alpha} + \epsilon_n$ ist. Diese Teststatistik hat den Vorteil, dass sie auch dann gilt, wenn u_n nicht normalverteilt ist.

(g) Erläutern Sie, wie Sie den in Teilaufgabe (f) beschriebenen Test unter Berücksichtigung der Information des Tukey-Anscombe Diagramms formulieren!

Führen Sie diesen Test zum Signifikanzniveau von $\alpha = 5\%$ und für ein R^2 der entsprechenden Hilfsregression von 0.0687 durch!

Lösungsvorschlag: Die Heteroskedastizität scheint im Tukey-Anscombe Diagramm eine Funktion der Projektion $\hat{y}_n = \boldsymbol{x}_n'\hat{\boldsymbol{\beta}}_{KQ}$ zu sein. Daher sollte die Hilfsregression als

$$v_n = \boldsymbol{x}_n'\boldsymbol{\alpha} + \epsilon_n \ ,$$

mit $\boldsymbol{x}_n = (1\ x_{n2}\ x_{n3})'$ und $\boldsymbol{\alpha} = (\alpha_1\ \alpha_2\ \alpha_3)'$, formuliert werden. Bei genaueren Informationen könnte man die Anzahl der Regressoren in der Hilfsregression einschränken.
Hypothese:

$$\mathrm{H}_0 : \alpha_2 = \alpha_3 = 0 \qquad \text{vs.} \qquad \mathrm{H}_1 : \exists\, j = 2,3 : \ \alpha_j \neq 0 \ .$$

Teststatistik:

$$N \cdot R^2 = 300 \cdot 0.0687 = 20.61 \ .$$

Kritischer Wert:

$$\chi^2_{krit} = F^{-1}_{p-1}(1-\alpha) = F^{-1}_2(0,95) = 5.99 \ .$$

Testentscheidung: H_0 wird abgelehnt, da $\chi^2_{krit} = 5.99 < 20.61 = N \cdot R^2$. d.h. es liegen keine homoskedastischen Störgrößen vor.

(h) Die Schätzung des linearen Regressionsmodells aus Teilaufgabe (f) ergab

$$\hat{\boldsymbol{\beta}}_{KQ} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{y} = \begin{pmatrix} 0.8891 \\ 1.2430 \\ 1.2545 \end{pmatrix} .$$

Benutzen Sie die analytischen Ergebnisse aus Aufgabenblatt 5, um zu beurteilen, welche Auswirkungen die Verwendung der Kovarianzmatrizen

$$\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}_{KQ}} = \hat{\sigma}^2_{KQ}(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{bmatrix} 0.0040 & -0.0021 & -0.0018 \\ -0.0021 & 0.0025 & 0.0001 \\ -0.0018 & 0.0001 & 0.0022 \end{bmatrix}$$

bzw.

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{KQ}} = (\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}\boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1} = \begin{bmatrix} 0.0038 & -0.0025 & -0.0019 \\ -0.0025 & 0.0037 & 0.0001 \\ -0.0019 & 0.0001 & 0.0023 \end{bmatrix}$$

haben werden, wobei angenommen wird, dass $\boldsymbol{\Sigma} := \mathrm{E}[\,\boldsymbol{u}\boldsymbol{u}'\,]$ in $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{KQ}}$ bekannt sei!

Lösungsvorschlag: $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}_{KQ}}$ ist die Kovarianzmatrix, die daraus resultiert, dass man fälschlicherweise homoskedastische Störgrößen annimmt. Wie wir in Aufgabenblatt 5 gesehen haben, ist $\hat{\sigma}^2_{KQ}$ ein verzerrter Schätzer für σ^2 und der Matrixteil von $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}_{KQ}}$, $(\boldsymbol{X}'\boldsymbol{X})^{-1}$, ist nur für homoskedastische Störgrößen korrekt. Damit ist $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}_{KQ}}$ die falsche Kovarianzmatrix für $\hat{\boldsymbol{\beta}}_{KQ}$. Bei Heteroskedastie ist $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{KQ}}$ die richtige Kovarianzmatrix für $\hat{\boldsymbol{\beta}}_{KQ}$, falls $\boldsymbol{\Sigma}$ bekannt ist. Falls $\boldsymbol{\Sigma}$ unbekannt ist, muss $\boldsymbol{X}'\boldsymbol{\Sigma}\boldsymbol{X}$ durch den Kovarianzschätzer von White (1980) bestimmt werden. Die Quadratwurzel der Hauptdiagonalenelemente gibt die Schätzfehler des KQ-Schätzers $\hat{\boldsymbol{\beta}}_{KQ}$ an. Für $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}_{KQ}}$ folgt hieraus

$$\begin{aligned}
\hat{\sigma}_{\hat{\boldsymbol{\beta}}_{KQ,1}} &= \sqrt{\left(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}_{KQ}}\right)_{11}} = \sqrt{0.0040} = 0.0629 \\
\hat{\sigma}_{\hat{\boldsymbol{\beta}}_{KQ,2}} &= \sqrt{\left(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}_{KQ}}\right)_{22}} = \sqrt{0.0025} = 0.0502 \\
\hat{\sigma}_{\hat{\boldsymbol{\beta}}_{KQ,3}} &= \sqrt{\left(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}_{KQ}}\right)_{33}} = \sqrt{0.0022} = 0.0467\,.
\end{aligned}$$

Für $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{KQ}}$ folgt

$$\begin{aligned}
\sigma_{\hat{\boldsymbol{\beta}}_{KQ,1}} &= \sqrt{\left(\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{KQ}}\right)_{11}} = \sqrt{0.0038} = 0.0616 \\
\sigma_{\hat{\boldsymbol{\beta}}_{KQ,2}} &= \sqrt{\left(\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{KQ}}\right)_{22}} = \sqrt{0.0037} = 0.0608 \\
\sigma_{\hat{\boldsymbol{\beta}}_{KQ,3}} &= \sqrt{\left(\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{KQ}}\right)_{33}} = \sqrt{0.0023} = 0.0480\,.
\end{aligned}$$

Da die Schätzfehler der KQ-Schätzer, die sich aus $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}_{KQ}}$ ergeben, bis auf den für $\hat{\beta}_{KQ,1}$, durchweg kleiner sind als die Schätzfehler aus $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{KQ}}$, liegt der Schluss nahe, dass wir (für die vorliegende Stichprobe) das Risiko der Schätzung von β_1 bzw. β_i für $i = 2, 3$ überschätzen bzw. unterschätzen. Für Signifikanztests einzelner Regressoren bedeutet dies, dass bei fälschlicher Verwendung der Schätzfehler aus $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}_{KQ}}$ die Tendenz besteht, die Nullhypothese $\beta_i = 0$ für $i = 1$ bzw. $i = 2, 3$ eher nicht abzulehnen bzw. abzulehnen.

(i) Ist die Struktur der Kovarianzmatrix der Störgrößen $\boldsymbol{\Sigma}$ bekannt, lässt sich für den gegebenen Datensatz der VKQ-Schätzer

$$\hat{\boldsymbol{\beta}}_{VKQ} = (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{y} = \begin{pmatrix} 0.8852 \\ 1.2301 \\ 1.2706 \end{pmatrix}$$

und seine Kovarianzmatrix

$$\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{VKQ}} = (\boldsymbol{X}'\boldsymbol{\Sigma}^{-1}\boldsymbol{X})^{-1} = \begin{bmatrix} 0.0033 & -0.0020 & -0.0017 \\ -0.0020 & 0.0030 & 0.0001 \\ -0.0017 & 0.0001 & 0.0020 \end{bmatrix}$$

berechnen.

Vergleichen Sie die Präzision von $\hat{\boldsymbol{\beta}}_{KQ}$ und $\hat{\boldsymbol{\beta}}_{VKQ}$, die sich aus $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{KQ}}$ und $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{VKQ}}$ ergeben! Diskutieren Sie, inwiefern diese Vergleiche mit den theoretischen Überlegungen von Aufgabenblatt 5 übereinstimmen!

<u>**Lösungsvorschlag:**</u> Die Schätzfehler aus der gegebenen Kovarianzmatrix $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{VKQ}}$ lauten

$$
\begin{aligned}
\sigma_{\hat{\beta}_{VKQ,1}} &= \sqrt{\left(\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{VKQ}}\right)_{11}} = \sqrt{0.0033} = 0.0574 \\
\sigma_{\hat{\beta}_{VKQ,2}} &= \sqrt{\left(\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{VKQ}}\right)_{22}} = \sqrt{0.0030} = 0.0548 \\
\sigma_{\hat{\beta}_{VKQ,3}} &= \sqrt{\left(\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{VKQ}}\right)_{33}} = \sqrt{0.0020} = 0.0447 \ .
\end{aligned}
$$

Gemessen an der richtigen Kovarianzmatrix $\boldsymbol{\Sigma}_{\hat{\boldsymbol{\beta}}_{KQ}}$ sind die Schätzfehler von $\hat{\boldsymbol{\beta}}_{VKQ}$ für alle Parameter in $\boldsymbol{\beta}$ kleiner als die Schätzfehler von $\hat{\boldsymbol{\beta}}_{KQ}$. Dieses Ergebnis ist in Übereinstimmung mit den theoretischen Überlegungen aus Aufgabenblatt 5, wo gezeigt wurde, dass $\hat{\boldsymbol{\beta}}_{VKQ}$ gegenüber $\hat{\boldsymbol{\beta}}_{KQ}$ relativ effizient ist.

(j) Nun sei $\boldsymbol{\Sigma}$ unbekannt, so dass sie zunächst geschätzt werden muss, um $\hat{\boldsymbol{\beta}}_{EVKQ}$ operationalisieren zu können.

Erklären Sie, wie—unter der Annahme von multiplikativer Heteroskedastizität $\sigma_n^2 = \exp(\alpha_1 + \alpha_2 x_{n2} + \alpha_3 x_{n3})$—die Elemente von $\boldsymbol{\Sigma}$ geschätzt werden können! Das Ergebnis dieser Schätzung sei:

	α_1	α_2	α_3
Schätzer	-3.3546	0.8257	-0.0840
Schätzfehler	0.2753	0.2196	0.2042

Wie lautet unter der Information in obiger Tabelle Ihre Spezifikation von $\hat{\sigma}_n^2$?

Erklären Sie, wieso $\hat{\boldsymbol{\beta}}_{EVKQ}$ konsistent geschätzt werden kann, obwohl $\hat{\alpha}_1$ nicht konsistent geschätzt werden kann!

<u>**Lösungsvorschlag:**</u> Zunächst logarithmiert man $\sigma_n^2 = \exp\big(\alpha_1 + \alpha_2 x_{n2} + \alpha_3 x_{n3}\big)$:

$$\ln\big(\sigma_n^2\big) = \alpha_1 + \alpha_2 x_{n2} + \alpha_3 x_{n3} \ .$$

Eine gangbare Approximation für $\ln\big(\sigma_n^2\big)$ ist $\ln\big(\hat{u}_n^2\big)$, welche man in einem ersten Schritt aus der KQ-Schätzung von $y_n = \beta_1 + \beta_2 x_{n2} + \beta_3 x_{n3} + u_n$ gewinnt. Addiert man $\ln\big(\hat{u}_n^2\big)$ auf beiden Seiten

$$\ln\big(\hat{u}_n^2\big) + \ln\big(\sigma_n^2\big) = \alpha_1 + \alpha_2 x_{n2} + \alpha_3 x_{n3} + \ln\big(\hat{u}_n^2\big) \ ,$$

folgt eine Regressionsgleichung,

$$\begin{aligned}\ln\left(\hat{u}_n^2\right) &= \alpha_1 + \alpha_2 x_{n2} + \alpha_3 x_{n3} + \ln\left(\hat{u}_n^2\right) - \ln\left(\sigma_n^2\right) \\ &= \alpha_1 + \alpha_2 x_{n2} + \alpha_3 x_{n3} + \ln\left(\frac{\hat{u}_n^2}{\sigma_n^2}\right) \\ &= \alpha_1 + \alpha_2 x_{n2} + \alpha_3 x_{n3} + \epsilon_n \ ,\end{aligned}$$

aus der α_2 und α_3 konsistent geschätzt werden können.[1]

Nach der KQ-Schätzung können mit Hilfe der Schätzfehler t-Tests zum Signifikanzniveau 5% durchgeführt werden, um die Spezifikation von $\sigma_n^2 = \exp(\alpha_1 + \alpha_2 x_{n2} + \alpha_3 x_{n3})$ festzulegen.

$$\begin{aligned}|t_{\hat{\alpha}_1}| &= \left|\frac{\hat{\alpha}_1 - 0}{\hat{\sigma}_{\hat{\alpha}_1}}\right| = \left|\frac{-3.3546}{0.2753}\right| = 12.19 > 1,96 = t_{krit} \\ |t_{\hat{\alpha}_2}| &= \left|\frac{\hat{\alpha}_2 - 0}{\hat{\sigma}_{\hat{\alpha}_2}}\right| = \left|\frac{0.8257}{0.2196}\right| = 3.76 > 1,96 = t_{krit} \\ |t_{\hat{\alpha}_3}| &= \left|\frac{\hat{\alpha}_3 - 0}{\hat{\sigma}_{\hat{\alpha}_3}}\right| = \left|\frac{-0.0840}{0.2042}\right| = 0.41 < 1,96 = t_{krit}\end{aligned}$$

Damit sind lediglich $\hat{\alpha}_1$ und $\hat{\alpha}_2$ signifikant zum Niveau 5% von 0 verschieden, und die bevorzugte parametrische Spezifikation der Varianzgleichung lautet

$$\sigma_n^2 = \exp(\alpha_1 + \alpha_2 x_{n2}) \ .$$

Allgemein gilt, dass man die volle Spezifikation von $\mathbf{\Sigma}$ nicht vernünftig schätzen kann, ohne Restriktionen auf $\mathbf{\Sigma}$ aufzuerlegen, da $\mathbf{\Sigma}$ bei Heteroskedastizität genau soviele zu schätzende Parameter besitzt wie die Stichprobe Beobachtungen hat. Eine solche Restriktion ist z.B. die Annahme multiplikativer Heteroskedastizität.

$$\begin{aligned}\hat{\sigma}_n^2 &= \exp\left(\hat{\alpha}_1 + \hat{\alpha}_2 x_{n2}\right) = \exp\left(\hat{\alpha}_1\right)\exp\left(\hat{\alpha}_2 x_{n2}\right) = \hat{\sigma}^2 \exp\left(\hat{\alpha}_2 x_{n2}\right) \\ \Longrightarrow \hat{\mathbf{\Sigma}} &= \begin{bmatrix} \hat{\sigma}_1^2 & & & 0 \\ & \hat{\sigma}_2^2 & & \\ & & \ddots & \\ 0 & & & \hat{\sigma}_N^2 \end{bmatrix} \\ &= \begin{bmatrix} \hat{\sigma}^2 \exp\left(\hat{\alpha}_2 x_{12}\right) & & & 0 \\ & \hat{\sigma}^2 \exp\left(\hat{\alpha}_2 x_{22}\right) & & \\ & & \ddots & \\ 0 & & & \hat{\sigma}^2 \exp\left(\hat{\alpha}_2 x_{N2}\right) \end{bmatrix}\end{aligned}$$

[1]Wie Judge et al. (1988) anmerken, lässt sich zeigen, dass (auch asymptotisch) $\mathrm{E}[\epsilon_n] \neq 0$ gilt, so dass α_1 weder unverzerrt noch konsistent geschätzt werden kann.

$$= \hat{\sigma}^2 \begin{bmatrix} \exp\big(\hat{\alpha}_2 x_{12}\big) & & & 0 \\ & \exp\big(\hat{\alpha}_2 x_{22}\big) & & \\ & & \ddots & \\ 0 & & & \exp\big(\hat{\alpha}_2 x_{N2}\big) \end{bmatrix} = \hat{\sigma}^2 \hat{\boldsymbol{\Omega}} \, .$$

Setzt man $\hat{\boldsymbol{\Sigma}}$ in

$$\hat{\boldsymbol{\beta}}_{EVKQ} = \big(\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X}\big)^{-1}\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{y}$$

ein, folgt

$$\begin{aligned} \hat{\boldsymbol{\beta}}_{EVKQ} &= \big(\boldsymbol{X}'\hat{\sigma}^{-2}\hat{\boldsymbol{\Omega}}^{-1}\boldsymbol{X}\big)^{-1}\boldsymbol{X}'\hat{\sigma}^{-2}\hat{\boldsymbol{\Omega}}^{-1}\boldsymbol{y} \\ &= \hat{\sigma}^2\big(\boldsymbol{X}'\hat{\boldsymbol{\Omega}}^{-1}\boldsymbol{X}\big)^{-1}\boldsymbol{X}'\hat{\boldsymbol{\Omega}}^{-1}\boldsymbol{y}\hat{\sigma}^{-2} \\ &= \big(\boldsymbol{X}'\hat{\boldsymbol{\Omega}}^{-1}\boldsymbol{X}\big)^{-1}\boldsymbol{X}'\hat{\boldsymbol{\Omega}}^{-1}\boldsymbol{y} \, . \end{aligned}$$

Damit ist bewiesen, dass $\hat{\boldsymbol{\beta}}_{EVKQ}$ zwar von der Strukturmatrix $\hat{\boldsymbol{\Omega}}$ abhängt, aber nicht von $\hat{\sigma}^2 = \exp(\hat{\alpha}_1)$. Deshalb spielt die Tatsache, dass $\hat{\alpha}_1$ nicht konsistent geschätzt werden kann keine Rolle für $\hat{\boldsymbol{\beta}}_{EVKQ}$.

(k) Nach Schätzung der unbekannten Kovarianzmatrix $\boldsymbol{\Sigma}$ unter Berücksichtigung Ihrer parametrischen Spezifikation aus Teilaufgabe (i) lässt sich nun für die gegebene Stichprobe der empirische verallgemeinerte KQ-Schätzer

$$\hat{\boldsymbol{\beta}}_{EVKQ} = (\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X})^{-1}\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{y} = \begin{pmatrix} 0.8812 \\ 1.2263 \\ 1.2784 \end{pmatrix}$$

und seine Kovarianzmatrix

$$\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}_{EVKQ}} = (\boldsymbol{X}'\hat{\boldsymbol{\Sigma}}^{-1}\boldsymbol{X})^{-1} = \begin{bmatrix} 0.0066 & -0.0041 & -0.0034 \\ -0.0041 & 0.0069 & 0.0001 \\ -0.0034 & 0.0001 & 0.0041 \end{bmatrix}$$

berechnen.

Berechnen Sie die Präzision von $\hat{\boldsymbol{\beta}}_{EVKQ}$ und vergleichen Sie diese mit den Ergebnissen aus Teilaufgabe (h)!

Diskutieren Sie, inwiefern diese Vergleiche mit den theoretischen Überlegungen aus Aufgabenblatt 5 übereinstimmen!

Lösungsvorschlag: Die Schätzfehler aus der Kovarianzmatrix $\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}_{EVKQ}}$ lauten

$$\begin{aligned} \hat{\sigma}_{\hat{\beta}_{EVKQ,1}} &= \sqrt{\big(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}_{EVKQ}}\big)_{11}} = \sqrt{0.0066} = 0.0812 \\ \hat{\sigma}_{\hat{\beta}_{EVKQ,2}} &= \sqrt{\big(\hat{\boldsymbol{\Sigma}}_{\hat{\boldsymbol{\beta}}_{EVKQ}}\big)_{22}} = \sqrt{0.0069} = 0.0831 \end{aligned}$$

$$\hat{\sigma}_{\hat{\beta}_{EVKQ,3}} = \sqrt{(\hat{\mathbf{\Sigma}}_{\hat{\boldsymbol{\beta}}_{EVKQ}})_{33}} = \sqrt{0.0041} = 0.0640 \; .$$

Damit sind sämtliche Schätzfehler für $\hat{\boldsymbol{\beta}}_{EVKQ}$ durchweg größer als jene für $\hat{\boldsymbol{\beta}}_{VKQ}$ oder $\hat{\boldsymbol{\beta}}_{KQ}$. Während dieses Ergebnis für den Vergleich zwischen $\hat{\boldsymbol{\beta}}_{VKQ}$ und $\hat{\boldsymbol{\beta}}_{EVKQ}$ logisch ist, erscheint die relative Effizienz von $\hat{\boldsymbol{\beta}}_{KQ}$ gegenüber $\hat{\boldsymbol{\beta}}_{EVKQ}$ kontraintuitiv zu sein, da man—wie Aufgabenblatt 5 veranschaulicht—durch die zusätzliche Information über die Kovarianzmatrix $\mathbf{\Sigma}$ eine (relativ) effiziente Schätzung durch $\hat{\boldsymbol{\beta}}_{EVKQ}$ erwarten würde. Doch leider gilt das Gauss-Markov Theorem aus Aufgabenblatt 5 nicht für $\hat{\boldsymbol{\beta}}_{EVKQ}$. Aufgrund der Nichtlinearität von $\hat{\boldsymbol{\beta}}_{EVKQ}$ sind lediglich asymptotische Effizienzvergleiche möglich, d.h. über die relative Effizienz von $\hat{\boldsymbol{\beta}}_{EVKQ}$ gegenüber $\hat{\boldsymbol{\beta}}_{KQ}$ in endlichen Stichproben ist nichts Analytisches bekannt und keine Aussage möglich. Eine mögliche Erklärung, wieso $\hat{\boldsymbol{\beta}}_{KQ}$ kleinere Schätzfehler als $\hat{\boldsymbol{\beta}}_{EVKQ}$ aufweist, ist der gleiche Grund, der dazu geführt hat, dass $\hat{\boldsymbol{\beta}}_{VKQ}$ geringere Schätzfehler besitzt als $\hat{\boldsymbol{\beta}}_{EVKQ}$. Da für $\hat{\boldsymbol{\beta}}_{EVKQ}$ in der ersten Stufe $\mathbf{\Sigma}$ geschätzt werden muss, weist der Schätzer $\hat{\boldsymbol{\beta}}_{EVKQ}$ eine höhere Unsicherheit auf als Schätzer, die wie $\hat{\boldsymbol{\beta}}_{EVKQ}$ und—in diesem unrealistischen Fall—$\hat{\boldsymbol{\beta}}_{KQ}$ ohne direkte Schätzung von $\mathbf{\Sigma}$ auskommen. Insofern ist der Effizienzvergleich zwischen $\hat{\boldsymbol{\beta}}_{EVKQ}$ und $\hat{\boldsymbol{\beta}}_{KQ}$ nicht ganz fair.[2]

[2] Effizienzvergleiche für komplizierte Schätzer, deren Eigenschaften in endlichen Stichproben unbekannt sind, können mittels Monte Carlo Simulationen durchgeführt werden.