

Der Satz besagt, unter welchen Bedingungen γ_τ durch $\hat{\gamma}_N(\tau) = \frac{1}{N} \sum_{t=1}^{N-\tau} (Y_t - \bar{Y}_N)(Y_{t+\tau} - \bar{Y}_N)$ konsistent geschätzt werden kann.

Ergodizität von ARMA-Prozessen

3.55

- Stationäre ARMA $[p, q]$ -Prozesse sind mittelwertergodisch.
- Stationäre ARMA $[p, q]$ -Prozesse mit $\mathbb{E}[\varepsilon_t^4] < \infty$ sind kovarianzergodisch.

Hinweis: Diese Aussagen gelten auch für MA-Prozesse und für stationäre AR-Prozesse.

Literaturhinweise

3.56

Schlittgen und Streitberg (1999), Zeitreihenanalyse: Kapitel 2.3.2, 2.3.3, 2.3.4, 2.3.5, 4.1.2, 6.2.3

Schlittgen (2015), Angewandte Zeitreihenanalyse mit R: Kapitel 3

Literaturhinweise ■ Wiederholung von mathematischen Grundlagen

3.57

Schlittgen und Streitberg (1999), Zeitreihenanalyse: Anhänge B.1, B.2

3.3 Anpassung linearer Modelle

Vorgehensweise in der Anpassung von linearen Zeitreihenmodellen

3.58

1. **Datentransformation** im Rahmen des klassischen Komponentenmodells: Erzeugen einer trend- und saisonbereinigten Reihe, die mit Hilfe eines stationären und ergodischen Prozesses modelliert werden kann.
2. **Modellspezifikation:** Bestimmung der Ordnung (p, q) des Modells. Zwei Ansätze: auf der Grundlage der ACF und PACF oder auf der Grundlage von Informationskriterien.
3. **Modellschätzung:** Schätzung der Modellparameter bei gegebener Ordnung
4. **Modelldiagnose:** Beurteilung der Anpassungsgüte
5. **Modellanwendung:** Interpretation, Prognose

3.3.1 Anpassung autoregressiver Modelle

Datentransformation

3.59

In Kapitel 1 haben wir den Trend einer Zeitreihe mit Hilfe von

- polynomialen Trendfunktionen

oder

- der glatten Komponente

bestimmt und bereinigt. Alternativ kann die Trendbereinigung mit Hilfe des *Differenzenfilters* durchgeführt werden.

Datentransformation ■ Der Differenzenfilter

3.60

Definition 30 (Differenzenfilter Δ). Wir nennen den linearen Filter Δ ,

$$\Delta Y_t = (1 - B) Y_t = Y_t - Y_{t-1}, \quad t = 2, 3, \dots, N.$$

Differenzenfilter 1. Ordnung.

Hintereinanderausführungen von Δ schreiben wir als Potenzen

$$\Delta^2 Y_t = (1 - B)^2 Y_t = (1 - B) ((1 - B) Y_t) = \Delta (\Delta Y_t)$$

und allgemein für $d \in \mathbb{N}$

$$\Delta^d Y_t = \Delta^{d-1} (\Delta Y_t).$$

Datentransformation ■ Der Differenzenfilter

3.61

Definition 31 (Integrierter Prozess). Ein Prozess (Y_t) heißt **integriert** vom Grade d , kurz vom Typ $I(d)$, wenn $\Delta^d Y_t$ stationär ist.

Definition 32 (ARIMA $[p, d, q]$ -Prozess). Ein stochastischer Prozess (Y_t) heißt **Autoregressiver-Integrierter-Moving-Average-Prozess der Ordnung** $[p, d, q]$, kurz ARIMA $[p, d, q]$ -Prozess, wenn der nach d -maligem Differenzenbilden entstandene Prozess $\Delta^d Y_t$ einem stationären ARMA $[p, q]$ -Prozess genügt.

Schätzung von AR-Modellen

3.62

Das zentrierte AR-Modell:

$$Y_t - \bar{y} = \alpha_1 (Y_{t-1} - \bar{y}) + \dots + \alpha_p (Y_{t-p} - \bar{y}) + \varepsilon_t \quad (3.4)$$

Das geschätzte Modell:

$$\widehat{Y_t - \bar{y}} = \hat{\alpha}_1 (Y_{t-1} - \bar{y}) + \dots + \hat{\alpha}_p (Y_{t-p} - \bar{y})$$

Die Residuen:

$$\hat{\varepsilon}_t = Y_t - \bar{y} - \widehat{Y_t - \bar{y}}$$

Wir minimieren die Residuenquadratsumme:

$$\sum_{t=p+1}^N \hat{\varepsilon}_t^2 \stackrel{!}{=} \min_{\hat{\alpha}_1, \dots, \hat{\alpha}_p}$$

Schätzung von AR-Modellen

3.63

Die Normalgleichungen:

$$\begin{aligned} \hat{\alpha}_1 \sum_{t=p+1}^N (y_{t-1} - \bar{y})^2 + \dots + \hat{\alpha}_p \sum_{t=p+1}^N (y_{t-p} - \bar{y}) (y_{t-1} - \bar{y}) &= \sum_{t=p+1}^N (y_t - \bar{y}) (y_{t-1} - \bar{y}) \\ \hat{\alpha}_1 \sum_{t=p+1}^N (y_{t-1} - \bar{y}) (y_{t-2} - \bar{y}) + \dots + \hat{\alpha}_p \sum_{t=p+1}^N (y_{t-p} - \bar{y}) (y_{t-2} - \bar{y}) &= \sum_{t=p+1}^N (y_t - \bar{y}) (y_{t-2} - \bar{y}) \\ \vdots & \\ \hat{\alpha}_1 \sum_{t=p+1}^N (y_{t-1} - \bar{y}) (y_{t-p} - \bar{y}) + \dots + \hat{\alpha}_p \sum_{t=p+1}^N (y_{t-p} - \bar{y})^2 &= \sum_{t=p+1}^N (y_t - \bar{y}) (y_{t-p} - \bar{y}) \end{aligned}$$

Schätzung von AR-Modellen

3.64

Drei Schätzverfahren durch Lösung der Normalgleichungen:

1. CLS (Conditional Least Squares)

Die Summationen in den Normalgleichungen fangen bei $t = 1$ anstatt $t = p + 1$ an. Für die Terme $y_0, y_{-1}, \dots, y_{-(p-1)}$ wird der Erwartungswert Null eingesetzt.

2. ULS (Unconditional Least Squares)

Das AR-Modell wird zuerst mit einem alternativen Verfahren geschätzt und es werden die Prognosen in die Vergangenheit $\hat{y}_0, \hat{y}_{-1}, \dots, \hat{y}_{-(p-1)}$ der ersten p Reihenwerte berechnet. Dann werden die Normalgleichungen gelöst, wobei die Summationen bei $t = 1$ anstatt $t = p + 1$ anfangen. Für die Terme $y_0, y_{-1}, \dots, y_{-(p-1)}$ werden ihre Prognosen eingesetzt.

Schätzung von AR-Modellen

3.65

3. Yule-Walker

Die Normalgleichungen werden mit Hilfe der empirischen Autokorrelationen r_τ geschätzt. Wir erhalten das empirische Analogon der Yule-Walker-Gleichungen:

$$\begin{pmatrix} 1 & r_1 & \dots & r_{p-1} \\ r_1 & 1 & \dots & r_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ r_{p-1} & r_{p-2} & \dots & 1 \end{pmatrix} \begin{pmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \\ \vdots \\ \hat{\alpha}_p \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \\ \vdots \\ r_p \end{pmatrix}$$

Durch Lösen des Gleichungssystems erhalten wir die Parameterschätzungen $\hat{\alpha}_i$, $1 \leq i \leq p$.

Schätzung von AR-Modellen ■ Modellannahmen im Rahmen des linearen Regressionsmodells (3.4)

3.66

- (i) $\mathbb{E}[\varepsilon_t] = 0$, $t = 1, 2, \dots, N$.
- (ii) $\mathbb{E}[\varepsilon_s \varepsilon_t] = \begin{cases} \sigma_\varepsilon^2, & \text{falls } s = t \\ 0, & \text{falls } s \neq t \end{cases}$
- (iii) Die Regressoren $Y_{t-1} - \bar{y}, \dots, Y_{t-p} - \bar{y}$ sind nichtstochastisch und die Matrix $(Y_{t-1} - \bar{y}, \dots, Y_{t-p} - \bar{y})$ hat den Rang p .
- (iv) Es gibt über $\hat{\alpha}_1, \hat{\alpha}_2, \dots, \hat{\alpha}_p$ und σ^2 keine Vorinformation.

Schätzung von AR-Modellen ■ Modellannahmen im Rahmen des linearen Regressionsmodells (3.4)

3.67

Annahmeverletzung für Zeitreihendaten:

- die Regressoren sind *stochastisch*. Daher besondere (asymptotische) Betrachtung der Schätzungsgüte.

Schätzung von AR-Modellen ■ Die ML-Methode

3.68

4. Maximum Likelihood (ML)

Die beobachtete Zeitreihe $(y_t) = (y_1, y_2, \dots, y_N)$ ist eine Realisation des zugrundeliegenden AR-Prozesses (Y_t) .

Die multivariate Wahrscheinlichkeitsdichtefunktion $f((y_t))$ fasst die Wahrscheinlichkeiten zusammen, mit welchen sich beliebige Reihen (y_t) oder Reihenabschnitte $(y_{i_1}, y_{i_2}, \dots, y_{i_k})$, $1 \leq i_1, i_2, \dots, i_k \leq N$ realisieren können. Sei θ der unbekannte Parametervektor des AR-Prozesses (Y_t) , $\theta = (\sigma^2, \alpha_1, \dots, \alpha_p)$.

Für jede Ausprägung θ_i , $i = 1, 2, \dots$ von θ betrachten wir die multivariate Wahrscheinlichkeitsdichtefunktion $f((y_t) | \theta_i)$ bedingt auf den Parameter θ_i .

Schätzung von AR-Modellen ■ Die ML-Methode

3.69

- Wir können $f((y_t) | \theta_i)$ als Funktion des AR-Prozesses (Y_t) ausdrücken. Dann sprechen wir, wie oben, von *einer bedingten Wahrscheinlichkeitsdichtefunktion*.
- Oder als Funktion des Parametervektors θ . Diese ist dann die so genannte *Likelihood-Funktion*

$$L(\theta | (y_t)) := f((y_t) | \theta).$$

Der ML-Schätzer $\hat{\theta}$ ist der Parametervektor, für den die Likelihood-Funktion ihr Maximum annimmt:

$$L(\theta | (y_t)) \stackrel{!}{=} \max_{\theta}$$

Alternativ kann $\hat{\theta}$ durch Maximierung der so genannten *Loglikelihood-Funktion* berechnet werden

$$l(\theta | (y_t)) := \ln(L(\theta | (y_t))) \stackrel{!}{=} \max_{\theta}$$

Dabei gehen wir von einem *Gauß-Prozess* (Y_t) aus.

Schätzung von AR-Modellen ■ Schätzungsgüte

3.70

Theorem 33. Sei

$$Y_t - \mu = \alpha_1 (Y_{t-1} - \mu) + \dots + \alpha_p (Y_{t-p} - \mu) + \varepsilon_t$$

ein stationärer $AR[p]$ -Prozess mit $\mathbb{E}[\varepsilon_t] = 0$ und $\text{Var}[\varepsilon_t] = \sigma^2$. Dann gilt für die nach einer der vier Methoden berechneten Schätzer $\hat{\alpha}_i$, $i = 1, \dots, p$:

- $\hat{\alpha}_i$ konvergiert fast sicher (mit Wahrscheinlichkeit 1) gegen den wahren Parameterwert α_i wenn $N \rightarrow \infty$

$$\mathbb{P}\left(\lim_{N \rightarrow \infty} \hat{\alpha}_i(N) = \alpha_i\right) = 1$$

Schätzung von AR-Modellen ■ Schätzungsgüte

3.71

Theorem 34. ■ die Zufallsgrößen

$$\sqrt{N}(\hat{\alpha}_1 - \alpha_1), \dots, \sqrt{N}(\hat{\alpha}_p - \alpha_p)$$

sind asymptotisch multivariat normalverteilt mit Erwartungswert **0** und Kovarianzmatrix $\sigma^2 \sum_p^{-1}$ mit $\sum_p$ die Kovarianzmatrix von p aufeinander folgenden Variablen des Prozesses (Y_t)

$$\sum_p = \begin{pmatrix} \gamma_0 & \gamma_1 & \cdots & \gamma_{p-1} \\ \gamma_1 & \gamma_0 & \cdots & \gamma_{p-2} \\ & & \ddots & \\ \gamma_{p-1} & \gamma_{p-2} & \cdots & \gamma_0 \end{pmatrix}.$$

Modellspezifikation

3.72

Unter *Spezifikation* oder *Identifikation* eines AR-Modells verstehen wir die Bestimmung der Ordnung p des Modells. 2 Ansätze:

1. Der klassische Box-Jenkins-Ansatz besteht in der Interpretation der empirischen ACF und PACF.
2. Wir passen AR-Modelle verschiedener Ordnungen an die Daten an und wählen das geeignete Modell mit Hilfe eines Informationskriteriums. Wir wählen das Modell mit dem *kleinsten* Wert des betrachteten Informationskriteriums aus.

Modellspezifikation ■ Interpretation der empirischen ACF und PACF

3.73

Wir unterscheiden zwischen drei Verhaltensmustern der empirischen Kennfunktionen ACF und PACF:

- die Kennfunktion klingt exponentiell ab (siehe die ACF eines AR-, die PACF eines MA- sowie beide die ACF und die PACF eines gemischten ARMA-Modells). Dieses Verhalten ist dadurch gekennzeichnet, dass die Werte schnell abfallen, jedoch sehr langsam gegen Null konvergieren.
- die Kennfunktion klingt langsam von $+1$ herkommend ab (Hinweis auf eine Zeitreihe mit Trend);
- die Kennfunktion bricht abrupt ab (siehe die ACF eines White-Noise- oder MA-Prozesses und die PACF eines AR-Prozesses).

Modellspezifikation ■ Informationskriterien

3.74

- AIC (Akaike's Information Criterion):

$$AIC := \ln\left(\frac{RSS}{N}\right) + \frac{2p}{N}$$

- AICC (AIC with correction):

$$AICC := \ln\left(\frac{RSS}{N}\right) + \frac{2p}{N-p-2}$$

- BIC (Bayesian Information Criterion):

$$BIC := \ln\left(\frac{RSS}{N}\right) + \frac{p \ln N}{N}$$

Modellspezifikation ■ Überprüfung der Parametersignifikanz

3.75

- Asymptotische Verteilung der Parameterschätzer $\hat{\alpha}_i$, $1 \leq i \leq p$ in einem $AR[p]$ -Modell

$$\sqrt{N}(\hat{\alpha}_i - \alpha_i) \xrightarrow{d} N(0, \sigma_i^2).$$

Die Varianz σ_i^2 kann mit Hilfe des Satzes über die Schätzungsgüte bestimmt werden.

- Approximative Verteilung für N hinreichend groß

$$\hat{\alpha}_i \dot{\approx} N(\alpha_i, \sigma_{\hat{\alpha}_i}^2).$$

Modellspezifikation ■ Überprüfung der Parametersignifikanz

3.76

Testhypothesen:

$$H_0 : \alpha_i = 0 \quad \text{gegen} \quad H_1 : \alpha_i \neq 0$$

Teststatistik:

$$z = \frac{\hat{\alpha}_i}{\hat{\sigma}_{\hat{\alpha}_i}} \dot{\approx} N(0, 1) | H_0$$

Testentscheidung:

- Bestimme den kritischen Wert κ_α zu gegebenem $\alpha \in (0, 1)$:

$$\kappa_\alpha = z_{1-\frac{\alpha}{2}}.$$

- Lehne H_0 mit Irrtumswahrscheinlichkeit α ab, falls

$$|z| > \kappa_\alpha.$$

Modelldiagnose - Untersuchung der Residuen

3.77

- **Untersuchung der Abhängigkeitsstruktur** anhand der empirischen ACF und PACF der Residuen oder mit Hilfe von statistischen Tests (bspw. der Ljung-Box-Pierce-Test)
- **Überprüfung der Normalverteilungsannahme im Falle von ML-Schätzungen** mit dem QQ-Diagramm oder mit Hilfe von statistischen Tests

Modelldiagnose - Untersuchung der Residuen ■ Der Ljung-Box-Pierce-Test

3.78

Sei $(\tilde{r}_\tau)$ die empirische ACF der Residuen $(\hat{\varepsilon}_t)$.

Testhypothesen:

$$H_0 : (\varepsilon_t) \text{ White-Noise} \quad \text{gegen} \quad H_1 : (\varepsilon_t) \text{ kein White-Noise}$$

Teststatistik:

$$Q = N(N+2) \sum_{\tau=1}^k \frac{\tilde{r}_\tau^2}{N-\tau} \approx \chi_{k-p}^2 | H_0 .$$

Erfahrungsgemäß sollte Q für verschiedene Werte von $k = 6, 12, 18, 24, \dots$ geprüft werden, wobei $k > p$ gelten soll.

Testentscheidung:

- Bestimme den kritischen Wert κ_α zu gegebenem $\alpha \in (0, 1)$:

$$\kappa_\alpha = \chi_{1-\alpha; k-p}^2 .$$

- Lehne H_0 mit Irrtumswahrscheinlichkeit α ab, falls

$$Q > \kappa_\alpha .$$

3.3.2 Anpassung von Moving-Average-Modellen

Schätzung von MA-Modellen

3.79

Sei (Y_t) ein MA[q]-Prozess

$$Y_t = \varepsilon_t - \beta_1 \varepsilon_{t-1} - \dots - \beta_q \varepsilon_{t-q}$$

mit (ε_t) White-Noise. Y_t ist eine *nichtlineare* Funktion der unbekannten Parameter $\beta_1, \dots, \beta_q$.
Bspw. für das MA[1]-Modell

$$Y_t = - \sum_{u=1}^{t-1} \beta^u Y_{t-u} + \varepsilon_t - \beta^t \varepsilon_0.$$

Schätzung von MA-Modellen ■ Schätzung eines MA[1]-Modells

3.80

Das MA[1]-Modell:

$$Y_t = - \sum_{u=1}^{t-1} \beta^u Y_{t-u} + \varepsilon_t - \beta^t \varepsilon_0.$$

Das geschätzte MA[1]-Modell:

$$\hat{Y}_t = - \sum_{u=1}^{t-1} \hat{\beta}^u Y_{t-u}$$

Die Residuen:

$$\hat{\varepsilon}_t = Y_t - \hat{Y}_t = Y_t + \sum_{u=1}^{t-1} \hat{\beta}^u Y_{t-u}$$

Wir minimieren die Residuenquadratsumme:

$$\sum_t \hat{\varepsilon}_t^2 \stackrel{!}{=} \min_{\hat{\beta}}$$

Schätzung von MA-Modellen

3.81

Zwei Schätzverfahren basieren auf der Minimierung der Residuenquadratsumme:

1. CLS (Conditional Least Squares) – Für die Terme $\varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{-(q-1)}$ wird der Erwartungswert Null eingesetzt.

2. ULS (Unconditional Least Squares) – Das MA-Modell wird zuerst mit dem CLS-Verfahren geschätzt und es werden die Prognosen in die Vergangenheit $\hat{\varepsilon}_0, \hat{\varepsilon}_{-1}, \dots, \hat{\varepsilon}_{-(q-1)}$ berechnet. Dann wird die Residuenquadratsumme minimiert, wobei für die Terme $\varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{-(q-1)}$ ihre Prognosen eingesetzt werden.

Der CLS- und die ULS-Schätzer sind *nichtlineare* Schätzfunktionen, die mit Hilfe von *numerischen* Optimierungsverfahren berechnet werden. Alternativ:

3. ML (Maximum Likelihood)

3.3.3 Anpassung von Autoregressiven-Moving-Average-Modellen

Schätzung von ARMA-Modellen

3.82

1. CLS (Conditional Least Squares) – Für die Terme $y_0, y_{-1}, \dots, y_{-(p-1)}, \varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{-(q-1)}$ wird jeweils 0 eingesetzt.
2. ULS (Unconditional Least Squares) – Das ARMA-Modell wird zuerst mit dem CLS-Verfahren geschätzt und es werden die Prognosen in die Vergangenheit $\hat{y}_0, \hat{y}_{-1}, \dots, \hat{y}_{-(p-1)}, \hat{\varepsilon}_0, \hat{\varepsilon}_{-1}, \dots, \hat{\varepsilon}_{-(q-1)}$ berechnet. Dann wird die Residuenquadratsumme minimiert, wobei für die Terme $y_0, y_{-1}, \dots, y_{-(p-1)}, \varepsilon_0, \varepsilon_{-1}, \dots, \varepsilon_{-(q-1)}$ ihre Prognosen eingesetzt werden.

Der CLS- und die ULS-Schätzer sind *nichtlineare* Schätzfunktion, die mit Hilfe von *numerischen* Optimierungsverfahren berechnet werden. Alternativ:

3. ML (Maximum Likelihood)

Schätzung von ARMA-Modellen

3.83

Theorem 35 (Asymptotische Eigenschaften der Parameterschätzer eines ARMA-Prozesses). Sei (Y_t) ein stationärer und invertierbarer ARMA $[p, q]$ -Prozess $\alpha(B) Y_t = \beta(B) \varepsilon_t$ mit $\mathbb{E}[\varepsilon_t] = 0$ und $\text{Var}[\varepsilon_t] = \sigma^2$. Ferner sei unterstellt, dass die charakteristischen Polynome der Filter $\alpha(B)$ und $\beta(B)$ keine gemeinsame Nullstellen haben. Dann gilt für die nach einer der drei Methoden berechneten Schätzer $\hat{\alpha} = (\hat{\alpha}_1 \ \hat{\alpha}_2 \ \dots \ \hat{\alpha}_p)$ und $\hat{\beta} = (\hat{\beta}_1 \ \hat{\beta}_2 \ \dots \ \hat{\beta}_q)$:

- $\hat{\alpha}$ und $\hat{\beta}$ konvergieren fast sicher (mit Wahrscheinlichkeit 1) gegen die wahren Parametervektoren α und β wenn $N \rightarrow \infty$

$$\mathbb{P}\left(\lim_{N \rightarrow \infty} \hat{\alpha}(N) = \alpha\right) = 1 \quad \text{und} \quad \mathbb{P}\left(\lim_{N \rightarrow \infty} \hat{\beta}(N) = \beta\right) = 1$$

Schätzung von ARMA-Modellen

3.84

Theorem 36 (Asymptotische Eigenschaften der Parameterschätzer eines ARMA-Prozesses). ■
 die Zufallsgrößen $\sqrt{N}(\hat{\alpha} - \alpha)$, $\sqrt{N}(\hat{\beta} - \beta)$ sind asymptotisch gemeinsam multivariat normalverteilt mit Erwartungswertvektoren $\mathbf{0}$ und Kovarianzmatrix

$$\mathbf{V} = \begin{bmatrix} \Phi_{aa} & -\Phi_{ab} \\ -\Phi'_{ab} & \Phi_{bb} \end{bmatrix}^{-1}.$$

Seien (a_t) und (b_t) die Folgen der Gewichte der inversen Filter $\alpha^{-1}(\mathbf{B})$ und $\beta^{-1}(\mathbf{B})$. Die Kovarianzmatrix $\mathbf{V}$ wird mit Hilfe der unendlichen Reihen

$$\phi_{ab}(\tau) = \sum_{t=0}^{\infty} a_t b_{t+\tau}$$

und analog dazu $\phi_{aa}(\tau)$ und $\phi_{bb}(\tau)$ für $\tau \geq 0$ definiert.

Schätzung von ARMA-Modellen

3.85

Theorem 37 (Asymptotische Eigenschaften der Parameterschätzer eines ARMA-Prozesses). Die Matrix $\mathbf{V}$ hat die Partitionen:

$$\Phi_{aa} = \begin{pmatrix} \phi_{aa}(0) & \phi_{aa}(1) & \dots & \phi_{aa}(p-1) \\ \phi_{aa}(1) & \phi_{aa}(0) & & \vdots \\ \vdots & & \ddots & \phi_{aa}(1) \\ \phi_{aa}(p-1) & \dots & \phi_{aa}(1) & \phi_{aa}(0) \end{pmatrix}$$

$$\Phi_{bb} = \begin{pmatrix} \phi_{bb}(0) & \phi_{bb}(1) & \dots & \phi_{bb}(q-1) \\ \phi_{bb}(1) & \phi_{bb}(0) & & \vdots \\ \vdots & & \ddots & \phi_{bb}(1) \\ \phi_{bb}(q-1) & \dots & \phi_{bb}(1) & \phi_{bb}(0) \end{pmatrix}$$

Schätzung von ARMA-Modellen

3.86

Theorem 38 (Asymptotische Eigenschaften der Parameterschätzer eines ARMA-Prozesses).

$$\Phi_{ab} = \begin{pmatrix} \phi_{ab}(0) & \phi_{ab}(1) & \dots & \phi_{ab}(q-1) \\ \phi_{ab}(1) & \phi_{ab}(0) & & \phi_{ab}(q-2) \\ \vdots & & \ddots & \vdots \\ \phi_{ab}(p-1) & \phi_{ab}(p-2) & \dots & \phi_{ab}(0) \end{pmatrix}$$

Literaturhinweise

3.87

Schlittgen und Streitberg (1999), Zeitreihenanalyse: Kapitel 6.1.1, 6.1.7

Schlittgen (2015), Angewandte Zeitreihenanalyse mit R: Kapitel 3

3.4 Saisonale ARIMA-Modelle

Saisonale ARIMA-Modelle ■ Saisonale Differenzen

3.88

Im Rahmen der Datentransformation kann

- der Trend durch das d -fache Bilden von einfachen Differenzen

$$\Delta^d Y_t = (1 - B)^d Y_t$$

- die Saisonkomponente durch das D -fache Bilden von saisonalen Differenzen

$$\Delta_s Y_t = Y_t - Y_{t-s} = Y_t - B^s Y_t = (1 - B^s) Y_t$$

$$\Delta_s^D Y_t = (1 - B^s)^D Y_t$$

eliminiert werden.

Saisonale ARIMA-Modelle ■ Saisonale Differenzen

3.89

Einen Hinweis auf das Vorliegen einer Saisonkomponente liefern die ACF und PACF, wenn sie signifikante Werte für kleine Lags sowie für die Lags um die Saisonperiode $s, 2s, \dots$ aufweisen. Gleichzeitig sind die dazwischen liegenden Lags nicht signifikant von Null verschieden.

Saisonale ARIMA-Modelle ■ Saisonale Differenzen

3.90

Oft ist es nötig, beide Trend und Saisonkomponente zu bereinigen

$$(1 - B)^d (1 - B^s)^D Y_t.$$

Wir entscheiden über die Anzahl d und D der Differenzenbildungen mit Hilfe der Methode der *variablen Differenzen*. Wir wählen das Paar (d, D) mit der kleinsten empirischen Varianz der transformierten Zeitreihe $(1 - B)^d (1 - B^s)^D y_t$ aus.

Saisonale ARIMA-Modelle

3.91

Definition 39 (SARIMA $[p, d, q] \times [P, D, Q]_s$ -Prozess). Ein stochastischer Prozess (Y_t) heißt **saisonaler Autoregressiver-Integrierter-Moving-Average-Prozess der Ordnungen $[p, d, q]$ und $[P, D, Q]_s$** , kurz SARIMA $[p, d, q] \times [P, D, Q]_s$ -Prozess, wenn er der Gleichung

$$\alpha(B) \eta(B^s) (1 - B)^d (1 - B^s)^D Y_t = \beta(B) \theta(B^s) \varepsilon_t$$

mit (ε_t) ein White-Noise-Prozess genügt.

Literaturhinweise

3.92

Schlittgen (2015), Angewandte Zeitreihenanalyse mit R: Abschnitt *Saisonale ARIMA-Modelle* in Kapitel 3.4 und Abschnitt *Lineare Filterung von Zeitreihen* in Kapitel 2.5

3.5 Differenzen- vs. trendstationäre lineare Modelle**Differenzen- vs. trendstationäre lineare Modelle**

3.93

Wir unterscheiden zwischen

- einem *deterministischen* Trend, z.B. dem polynomialen Trend. Ein lineares Modell mit deterministischem Trend bezeichnen wir als *trendstationär*.

und

- einem *stochastischen* Trend, welcher im Rahmen eines instationären linearen Modells der Klasse ARIMA entsteht. Ein lineares Modell mit stochastischem Trend bezeichnen wir als *differenzenstationär*.

Differenzen- vs. trendstationäre lineare Modelle

3.94

Sei (ε_t) ein White-Noise-Prozess und (y_t) die zur Verfügung stehende Zeitreihe. Wir betrachten das lineare Modell

$$Y_t = \theta_0 + \theta_1 t + Z_t \quad \text{mit} \quad (1 - \alpha B) Z_t = \varepsilon_t.$$

1. $|\alpha| < 1$, (Z_t) ist stationär, (Y_t) ist trendstationär.

Die fehlerhafte Behandlung der Zeitreihe (y_t) als Realisation eines differenzenstationären Prozesses hat vertretbare Konsequenzen für die Modellschätzung und für das Testen.

2. $\alpha = 1$, (Z_t) ist nicht stationär, (Y_t) ist (zumindest) differenzenstationär.

Die fehlerhafte Behandlung der Zeitreihe (y_t) als Realisation eines trendstationären und nicht differenzenstationären Prozesses hat gravierende Konsequenzen für die Modellschätzung und für das Testen.

- (3.) $|\alpha| > 1$, ökonomisch nicht relevant.

Differenzen- vs. trendstationäre lineare Modelle ■ Der DF-Test

3.95

Der Einheitswurzeltest (Unit-Root-Test) von Dickey und Fuller erfolgt im Modell

$$\Delta Y_t = \mu + \beta t + \pi Y_{t-1} + \varepsilon_t$$

mit

$$\begin{cases} \mu = \theta_0 (1 - \alpha) + \alpha \theta_1 \\ \beta = \theta_1 (1 - \alpha) \\ \pi = \alpha - 1 \end{cases}.$$

Differenzen- vs. trendstationäre lineare Modelle ■ Der ADF-Test

3.96

Der erweiterte (*augmented*) Einheitswurzeltest von Dickey und Fuller erfolgt im Modell

$$\Delta Y_t = \mu + \beta t + \pi Y_{t-1} + \sum_{j=1}^k \gamma_j \Delta Y_{t-j} + \varepsilon_t$$

mit

$$\begin{cases} \mu = \theta_0 (1 - \alpha) + \alpha \theta_1 \\ \beta = \theta_1 (1 - \alpha) \\ \pi = \alpha - 1 \end{cases}.$$

Differenzen- vs. trendstationäre lineare Modelle ■ Der ADF-Test

3.97

Modell: $\Delta Y_t = \mu + \beta t + \pi Y_{t-1} + \sum_{j=1}^k \gamma_j \Delta Y_{t-j} + \varepsilon_t$ **Testhypothesen:**

$$H_0 : \beta = \pi = 0 \quad \text{gegen} \quad H_1 : \beta \neq 0 \vee \pi \neq 0$$

Teststatistik: Die Teststatistik ist die übliche F -Teststatistik. Ihre Verteilung unter H_0 wurde von Dickey und Fuller (1981) tabelliert. **Testentscheidung** (unter Berücksichtigung der jeweiligen Fehlerwahrscheinlichkeiten):

- H_0 beibehalten: die Hypothese eines differenzenstationären und nicht-trendstationären Modells kann nicht abgelehnt werden.
- H_0 ablehnen: die Hypothese eines differenzenstationären Modells wird abgelehnt. In diesem Fall können wir mit dem Ausgangsmodell wieder arbeiten und die Trendstationarität des Modells mit einem Standardtestverfahren überprüfen.

Differenzen- vs. trendstationäre lineare Modelle ■ Der ADF-Test

3.98

Modell: $Y_t = \theta_0 + \theta_1 t + Z_t$, mit $\alpha(B) Z_t = \varepsilon_t$, Z_t stationär. **Testhypothesen:**

$$H_0 : \theta_1 = 0 \quad \text{gegen} \quad H_1 : \theta_1 \neq 0$$

Teststatistik:

$$z = \frac{\hat{\theta}_1}{\hat{\sigma}_{\hat{\theta}_1}} \sim t_{N-2} | H_0$$

Testentscheidung (unter Berücksichtigung der jeweiligen Fehlerwahrscheinlichkeiten):

- H_0 beibehalten: die Hypothese eines nicht-trendstationären Modells kann nicht abgelehnt werden. **Insgesamt:** die Hypothese eines *stationären Modells* kann nicht abgelehnt werden.
- H_0 ablehnen: die Hypothese eines nicht-trendstationären Modells wird abgelehnt. **Insgesamt:** Hinweis auf ein *trendstationäres und nicht-differenzenstationäres Modell*.

Achtung! Der Störterm (Z_t) ist autokorreliert. Dazu noch mehr in Kapitel „Regressionsmodelle für Zeitreihen“.

Literaturhinweise

3.99

Schlittgen (2015), Angewandte Zeitreihenanalyse mit R: Kapitel 4

Literaturhinweise ■ Weiterführende Literatur

3.100

Kapitel 3.1, 3.2, 5.1 in

Pfaff, B. (2008), *Analysis of integrated and cointegrated time series with R*, 2. Auflage, Springer Verlag, New York