

3 Lineare Modelle

3.1 Lineare Filter

Lineare Filter

3.1

Definition 11 (Backshift-Operator B). Die Transformation B mit

$$BY_t = Y_{t-1}.$$

bezeichnen wir als **Backshift-Operator**.

Die Transformation B verschiebt den stochastischen Prozess um eine Zeiteinheit. Hintereinanderausführungen von B schreiben wir als Potenzen; allgemein ist

$$B^2 Y_t = B(BY_t) = BY_{t-1} = Y_{t-2},$$

und damit rekursiv

$$B^p Y_t = Y_{t-p}.$$

Lineare Filter

3.2

Definition 12 (Linearer Filter). Eine lineare Transformation eines stochastischen Prozesses (X_t) in einen anderen (Y_t) gemäß

$$Y_t = \sum_{u=-r}^s a_u X_{t-u}$$

wird als **linearer Filter** bezeichnet. Ein Filter wird durch seine Gewichte

$$(a_u)_{u=-r, \dots, s}$$

eindeutig definiert. Alternativ kann er als gewichtete Summe von Shiftoperatoren dargestellt werden

$$A(B) = \sum_{u=-r}^s a_u B^u.$$

Dann schreiben wir für Y_t

$$Y_t = A(B) X_t.$$

Lineare Filter

3.3

Wir bezeichnen die Anwendung eines Filters auf einen stochastischen Prozess (X_t) als Filtration des Prozesses, (X_t) als Input und den gefilterten Prozess (Y_t) als Output des Filters.

Definition 13 (*kausaler Filter*). Ein linearer Filter mit der Gewichtsfolge (a_u) heißt *kausal*, wenn $a_u = 0$ für $u < 0$ ist, d.h. wenn

$$A(B) = \sum_{u=0}^{\infty} a_u B^u.$$

Definition 14 (*absolut summierbarer Filter*). Ein Filter heißt *absolut summierbar*, wenn seine Gewichtsfolge (a_u) absolut summierbar ist.

Lineare Filter

3.4

Theorem 15 (Unendliche lineare Filter). Sei $(X_t)_{t \in \mathbb{Z}}$ ein stochastischer Prozess mit der Eigenschaft

$$\mathbb{E}[X_t^2] \leq K \quad \text{für alle } t \in \mathbb{Z},$$

wobei K eine endliche Konstante ist. Sei (a_u) ein absolut summierbarer Filter. Dann existiert ein stochastischer Prozess $(Y_t)_{t \in \mathbb{Z}}$, so dass für alle $t \in \mathbb{Z}$

$$\lim_{n \rightarrow \infty} \mathbb{E} \left[\left(Y_t - \sum_{u=-n}^n a_u X_{t-u} \right)^2 \right] = 0, \quad (3.1)$$

$$\mathbb{E}[Y_t^2] < \infty \quad \text{und} \quad \mathbb{E}[Y_t] = \lim_{n \rightarrow \infty} \sum_{u=-n}^n a_u \mathbb{E}[X_{t-u}].$$

Weiter benutzen wir für den Limes (3.1) die Notation:

$$Y_t = \sum_{u=-\infty}^{\infty} a_u X_{t-u}.$$

Lineare Filter

3.5

Unendliche lineare Filter mit absolut summierbarer Folge von Gewichten können unbekümmert im Sinne des vorigen Satzes gehandhabt werden.

- Bspw. ist die Vertauschung von Summation und Erwartungswertoperator erlaubt.
- Ist der Input (X_t) stationär, so ist auch der Output (Y_t) stationär.

Lineare Filter ■ Allgemeine lineare Prozesse

3.6

In Kapitel 1 haben wir den allgemeinen linearen Prozess

$$Y_t = \sum_{u=-\infty}^{\infty} a_u \varepsilon_{t-u}$$

mit $(\varepsilon_t)_{t \in \mathbb{Z}}$ White-Noise und (a_u) absolut summierbar eingeführt. Der allgemeine lineare Prozess wird durch Filtration des White-Noise-Prozesses $(\varepsilon_t)_{t \in \mathbb{Z}}$

$$Y_t = A(B) \varepsilon_t$$

mit dem unendlichen linearen Filter

$$A(B) = \sum_{u=-\infty}^{\infty} a_u B^u$$

erzeugt.

3.1.1 MA-Prozesse

Lineare Filter ■ MA-Prozesse

3.7

Definition 16 (Moving-Average-Prozess der Ordnung q , $MA[q]$). Ein stochastischer Prozess (Y_t) ,

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

heißt **Moving-Average-Prozess der Ordnung q** , kurz $MA[q]$ -Prozess. Dabei ist (ε_t) ein White-Noise-Prozess. Im Fall $q = \infty$ sprechen wir von einem unendlichen Moving-Average-Prozess ($MA[\infty]$).

Ein MA-Prozess entsteht durch Filtration des White-Noise-Prozesses (ε_t)

$$Y_t = \beta(B) \varepsilon_t$$

mit dem kausalen linearen Filter

$$\beta(B) = 1 - \beta_1 B - \beta_2 B^2 - \dots - \beta_q B^q.$$

Lineare Filter ■ MA-Prozesse – Theoretische Momente

3.8

Sei $\mathbb{E}[\varepsilon_t] = \mu_\varepsilon$, $\text{Var}[\varepsilon_t] = \sigma_\varepsilon^2$ und $\beta_0 = -1$.

$MA[q]$ -Prozesse (Y_t) sind *stationäre* Prozesse mit den Momenten:

1. $\mu_Y = \mathbb{E}[Y_t] = \mu_\varepsilon \sum_{u=0}^q (-\beta_u)$
2. $\gamma_Y(\tau) = \sum_{u=0}^q \sum_{v=0}^q \beta_u \beta_v \gamma_\varepsilon(\tau + u - v) = \begin{cases} 0, & \tau > q \\ \sigma_\varepsilon^2 \sum_{u=0}^{q-\tau} \beta_u \beta_{u+\tau}, & 0 \leq \tau \leq q \\ \gamma_Y(-\tau), & \tau < 0 \end{cases}$

Lineare Filter ■ MA-Prozesse – Theoretische Momente

3.9

$$3. \sigma_Y^2 = \text{Var}[Y_t] = \sigma_\varepsilon^2 \sum_{u=0}^q \beta_u^2$$

$$4. \rho_Y(\tau) = \frac{\gamma_Y(\tau)}{\gamma_Y(0)} = \begin{cases} 0, & \tau > q \\ \sum_{u=0}^{q-\tau} \beta_u \beta_{u+\tau} / \sum_{u=0}^q \beta_u^2, & 0 < \tau \leq q \\ 1, & \tau = 0 \\ \rho_Y(-\tau), & \tau < 0 \end{cases}$$

Die ACF eines $MA[q]$ -Prozesses verschwindet für Lags $\tau > q$.

3.1.2 AR-Prozesse

Lineare Filter ■ AR-Prozesse

3.10

Definition 17 (Autoregressiver Prozess der Ordnung p , $AR[p]$). Ein stochastischer Prozess (Y_t) ,

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

heißt **Autoregressiver Prozess der Ordnung p** , kurz $AR[p]$ -Prozess. Dabei ist (ε_t) ein White-Noise-Prozess. Im Fall $p = \infty$ heißt (Y_t) ein unendlicher AR-Prozess ($AR[\infty]$).

Alternative Darstellung:

$$\varepsilon_t = Y_t - \alpha_1 Y_{t-1} - \dots - \alpha_p Y_{t-p}.$$

Der White-Noise-Prozess (ε_t) ist der Output des kausalen linearen Filters

$$\alpha(B) = 1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_p B^p$$

mit dem Input dem AR-Prozess (Y_t)

$$\varepsilon_t = \alpha(B) Y_t.$$

3.1.3 ARMA-Prozesse

Lineare Filter ■ ARMA-Prozesse

3.11

Definition 18 (Autoregressiver-Moving-Average-Prozess der Ordnung p, q , $ARMA[p, q]$). Ein stochastischer Prozess (Y_t) ,

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \dots + \alpha_p Y_{t-p} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \dots - \beta_q \varepsilon_{t-q}$$

mit (ε_t) ein White-Noise-Prozess heißt **Autoregressiver Moving-Average-Prozess der Ordnung p, q** , kurz $ARMA[p, q]$ -Prozess.

Alternative Darstellung:

$$Y_t - \alpha_1 Y_{t-1} - \alpha_2 Y_{t-2} - \dots - \alpha_p Y_{t-p} = \varepsilon_t - \beta_1 \varepsilon_{t-1} - \beta_2 \varepsilon_{t-2} - \dots - \beta_q \varepsilon_{t-q},$$

oder kurz:

$$\alpha(B) Y_t = \beta(B) \varepsilon_t$$

mit den kausalen linearen Filtern

$$\alpha(B) = 1 - \alpha_1 B - \alpha_2 B^2 - \dots - \alpha_p B^p$$

$$\beta(B) = 1 - \beta_1 B - \beta_2 B^2 - \dots - \beta_q B^q.$$

3.2 Invertierbare lineare Filter

Invertierbare lineare Filter

3.12

Problem 19. Gegeben sei der lineare Filter $A(B)$,

$$A(B) = 1 - a_1 B - a_2 B^2 - \dots - a_p B^p. \quad (3.2)$$

Bestimmen Sie einen zu $A(B)$ inversen linearen Filter $A^{-1}(B)$ mit Gewichten (c_u) mit den Eigenschaften

- $A^{-1}(B)$ ist kausal

und

- $A^{-1}(B)$ ist absolut summierbar, falls $A^{-1}(B)$ ein unendlicher Filter ist,

so dass für jeden stochastischen Prozess (X_t) gilt

$$Y_t = \sum_{u=0}^p a_u X_{t-u} \Leftrightarrow X_t = \sum_u c_u Y_{t-u}.$$

Invertierbare lineare Filter

3.13

Die Funktion

$$A(z) = 1 - a_1 z - a_2 z^2 - \dots - a_p z^p, \quad z \in \mathbb{C}$$

heißt das charakteristische Polynom des Filters

$$A(B) = 1 - a_1 B - a_2 B^2 - \dots - a_p B^p.$$

Jeder lineare Filter $A(B)$ ist eindeutig durch sein charakteristisches Polynom $A(z)$ bestimmt. Daher werden wir das weiter oben beschriebene Problem für $A(z)$ anstatt für $A(B)$ lösen.

D.h. wir suchen ein zu $A(z)$ inverses charakteristisches Polynom

$$A^{-1}(z) = c_0 + c_1 z + c_2 z^2 + \dots$$

mit absolut summierbarer Koeffizientenfolge (c_u) , falls $A^{-1}(z)$ ein unendliches Polynom ist.

Invertierbare lineare Filter

3.14

Theorem 20. Für den linearen Filter

$$A(B) = 1 - a_1 B - a_2 B^2 - \dots - a_p B^p$$

existiert dann und nur dann ein kausaler, absolut summierbarer, inverser Filter

$$A^{-1}(B) = c_0 + c_1 B + c_2 B^2 + \dots,$$

wenn alle Nullstellen des charakteristischen Polynoms

$$A(z) = 1 - a_1 z - a_2 z^2 - \dots - a_p z^p$$

außerhalb des Einheitskreises liegen.

Invertierbare lineare Filter ■ Eindeutige Beschreibung mit Hilfe der ACF

3.15

Hinweis: Ein stochastischer Prozess Y_t

$$Y_t = A(B) X_t$$

mit invertierbarem linearem Filter

$$A(B) = 1 - a_1 B - a_2 B^2 - \dots - a_p B^p$$

und kausalem, absolut summierbarem, inversem Filter $A^{-1}(B)$ ist eindeutig durch seine Autokorrelationsfunktion definiert.

3.2.1 Stationäre AR-Prozesse

Stationäre AR-Prozesse ■ Stationaritätsbedingung

3.16

Theorem 21. Ein $AR[p]$ -Prozess $Y_t = \alpha_1 Y_{t-1} + \dots + \alpha_p Y_{t-p} + \varepsilon_t$ kann dann und nur dann als $MA[\infty]$ -Prozess

$$Y_t = \sum_{u=0}^{\infty} \beta_u \varepsilon_{t-u}$$

mit absolut summierbarer Koeffizientenfolge (β_u) dargestellt werden, wenn alle Nullstellen des charakteristischen Polynoms

$$1 - \alpha_1 z - \alpha_2 z^2 - \dots - \alpha_p z^p = 0$$

außerhalb des Einheitskreises liegen. Speziell ist (Y_t) dann stationär.

Stationäre AR-Prozesse ■ Stationaritätsbedingung

3.17

- Stationaritätsbedingung für AR[1]-Prozesse $Y_t = \alpha Y_{t-1} + \varepsilon_t$

$$|\alpha| < 1$$

- Stationaritätsbedingungen für AR[2]-Prozesse $Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \varepsilon_t$

$$\begin{cases} \alpha_1 + \alpha_2 < 1 \\ \alpha_2 - \alpha_1 < 1 \\ \alpha_2 > -1 \end{cases}$$

3.18

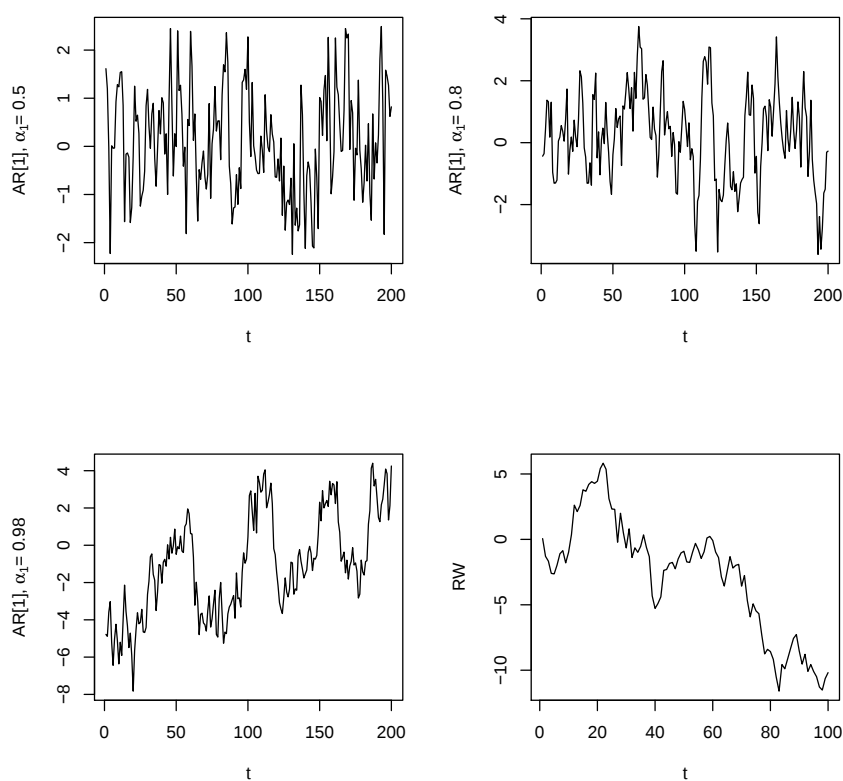


Abb. 3.1: Realisationen von AR-Prozessen und einem Random-Walk

3.19

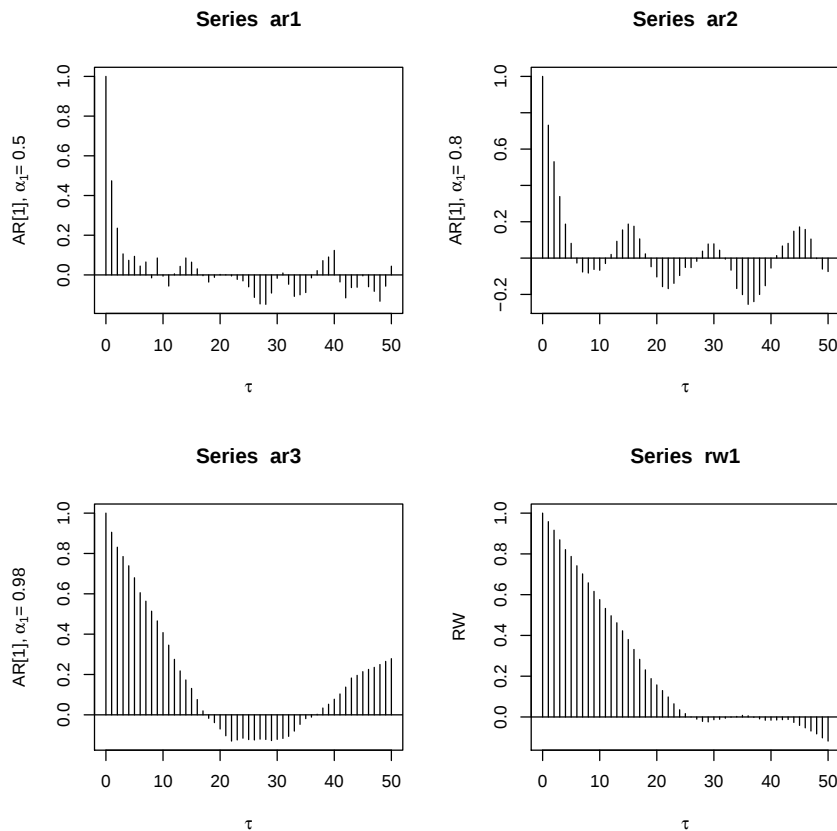


Abb. 3.2: Die empirischen ACF

Stationäre AR-Prozesse ■ Eindeutige Beschreibung mit Hilfe der ACF

3.20

Stationäre AR-Prozesse sind eindeutig durch ihre ACF definiert. Berechnung der ACF für einen stationären AR-Prozess Y_t mit $\mathbb{E}[Y_t] = 0$:

$$\begin{aligned}
 Y_t &= \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} \dots + \alpha_p Y_{t-p} + \varepsilon_t \\
 &\Downarrow \\
 Y_t Y_{t-\tau} &= \alpha_1 Y_{t-1} Y_{t-\tau} + \alpha_2 Y_{t-2} Y_{t-\tau} \dots + \alpha_p Y_{t-p} Y_{t-\tau} + \varepsilon_t Y_{t-\tau} \\
 &\Downarrow \\
 \mathbb{E}[Y_t Y_{t-\tau}] &= \alpha_1 \mathbb{E}[Y_{t-1} Y_{t-\tau}] + \alpha_2 \mathbb{E}[Y_{t-2} Y_{t-\tau}] \dots + \alpha_p \mathbb{E}[Y_{t-p} Y_{t-\tau}] + \mathbb{E}[\varepsilon_t Y_{t-\tau}] \\
 &\Downarrow \\
 \gamma(\tau) &= \alpha_1 \gamma(\tau-1) + \alpha_2 \gamma(\tau-2) \dots + \alpha_p \gamma(\tau-p) \\
 &\Downarrow \\
 \rho(\tau) &= \alpha_1 \rho(\tau-1) + \alpha_2 \rho(\tau-2) + \dots + \alpha_p \rho(\tau-p)
 \end{aligned}$$

Stationäre AR-Prozesse ■ Eindeutige Beschreibung mit Hilfe der ACF

3.21

Das System der so genannten Yule-Walker-Gleichungen:

$$\rho(\tau) = \alpha_1 \rho(\tau - 1) + \alpha_2 \rho(\tau - 2) + \dots + \alpha_p \rho(\tau - p), \quad 1 \leq \tau \leq p \quad (3.3)$$

bzw. in der Matrixschreibweise:

$$\begin{pmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{pmatrix} = \begin{pmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{pmatrix}.$$

Dieses System enthält p Gleichungen in p Unbekannten ρ_τ , $1 \leq \tau \leq p$.

Es ist eindeutig lösbar für stationäre AR-Prozesse.

Die Yule-Walker-Gleichungen

3.22

Das empirische Analogon zu den Yule-Walker-Gleichungen:

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

Hier sind $\hat{\alpha}_i$, $1 \leq i \leq p$, die Unbekannten. Die Werte r_τ stellen die empirischen Autokorrelationen dar. Durch Lösen des Gleichungssystems erhalten wir die Parameterschätzungen $\hat{\alpha}_i$. Diese sind die so genannten Yule-Walker-Schätzer des AR-Modells.

Die Levinson-Durbin-Rekursion ■ Schätzung eines AR[3]-Modells

3.23

Gegeben seien die Yule-Walker-Gleichungen für den AR[2]-Prozess:

$$\begin{pmatrix} 1 & r_1 \\ r_1 & 1 \end{pmatrix} \begin{pmatrix} \hat{\alpha}_1 \\ \hat{\alpha}_2 \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \end{pmatrix}$$

mit bekannten Lösungen $\hat{\alpha}_1$ und $\hat{\alpha}_2$. Wir berechnen die Lösungen $\hat{\theta}_1$, $\hat{\theta}_2$ und $\hat{\theta}_3$ der Yule-Walker-Gleichungen für den AR[3]-Prozess:

$$\begin{pmatrix} 1 & r_1 & r_2 \\ r_1 & 1 & r_1 \\ r_2 & r_1 & 1 \end{pmatrix} \begin{pmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \\ \hat{\theta}_3 \end{pmatrix} = \begin{pmatrix} r_1 \\ r_2 \\ r_3 \end{pmatrix}.$$

Die Levinson-Durbin-Rekursion ■ Schätzung eines AR[3]-Modells

3.24

Wir erhalten die Lösungen:

$$\begin{cases} \hat{\theta}_1 = \hat{\alpha}_1 - \hat{\alpha}_2 \hat{\theta}_3 \\ \hat{\theta}_2 = \hat{\alpha}_2 - \hat{\alpha}_1 \hat{\theta}_3 \\ \hat{\theta}_3 = \frac{r_3 - r_2 \hat{\alpha}_1 - r_1 \hat{\alpha}_2}{1 - r_2 \hat{\alpha}_2 - r_1 \hat{\alpha}_1} \end{cases}$$

Die Levinson-Durbin-Rekursion

3.25

Theorem 22 (Die Levinson-Durbin-Rekursion). *Gegeben seien*

- die Startwerte: $\hat{\alpha}_1(1) = r_1$, $\hat{Q}(1) = 1 - r_1^2$
- die Lösungen $\hat{\alpha}_1(p-1), \hat{\alpha}_2(p-1), \dots, \hat{\alpha}_{p-1}(p-1)$ für die Yule-Walker-Gleichungen der Ordnung $p-1$
- der Wert $\hat{Q}(p-1) = 1 - \hat{\alpha}_1(p-1)r_1 - \dots - \hat{\alpha}_{p-1}(p-1)r_{p-1}$.

Dann ergeben sich die Lösungen der Yule-Walker-Gleichungen der Ordnung p

$$\hat{\alpha}_p(p) = \Delta(p) / \hat{Q}(p-1)$$

$$\hat{\alpha}_i(p) = \hat{\alpha}_i(p-1) - \hat{\alpha}_p(p) \hat{\alpha}_{p-i}(p-1), \quad i = 1, \dots, p-1$$

wobei $\hat{Q}(p) = \hat{Q}(p-1)(1 - \hat{\alpha}_p^2(p))$ und $\Delta(p) = r_p - (\hat{\alpha}_1(p-1)r_{p-1} - \dots - \hat{\alpha}_{p-1}(p-1)r_1)$.

Die Levinson-Durbin-Rekursion

3.26

Bei der Anwendung der Levinson-Durbin-Rekursion für die Berechnung der Yule-Walker-Schätzer $\hat{\alpha}_i$, $1 \leq i \leq p$ eines AR[p]-Modells werden in einem Durchgang auch

- die Yule-Walker-Schätzer für alle AR-Modelle niedrigerer Ordnung 1 bis $p-1$

und

- die empirische partielle Autokorrelationsfunktion ($\hat{\pi}_\tau$) des AR[p]-Modells

berechnet.

Die empirischen partiellen Autokorrelationen $\hat{\pi}_p$ sind gerade die Koeffizienten $\hat{\alpha}_p(p)$, die sich bei der Levinson-Durbin-Rekursion ergeben.

Die partielle Autokorrelationsfunktion

3.27

Die partielle Autokorrelation zweier Zeitreihenvariablen Y_t und $Y_{t-\tau}$ ist die Autokorrelation von Y_t und $Y_{t-\tau}$ unter Konstanthaltung / Ausschaltung der dazwischen liegenden Zufallsvariablen $Y_{t-\tau+1}, Y_{t-\tau+2}, \dots, Y_{t-1}$. Die Konstanthaltung / Ausschaltung der dazwischen liegenden Variablen erfolgt, indem man die mit Hilfe der auszuschaltenden Variablen $Y_{t-\tau+1}, Y_{t-\tau+2}, \dots, Y_{t-1}$ erstellten Prognosen $\hat{Y}_t$ und $\hat{Y}_{t-\tau}$ jeweils abzieht. Die partielle Autokorrelation ist dann die Kor-

relation zwischen $Y_t - \hat{Y}_t$ und $Y_{t-\tau} - \hat{Y}_{t-\tau}$.

Die partielle Autokorrelationsfunktion

3.28

Definition 23 (Die partielle Autokorrelationsfunktion (PACF)). Sei (Y_t) ein stationärer stochastischer Prozess. Die **partielle Autokorrelationsfunktion** (π_τ) , kurz PACF, ist die Folge der Korrelationen von $Y_t - \hat{Y}_t$ und $Y_{t-\tau} - \hat{Y}_{t-\tau}$ wobei $\hat{Y}_t$ und $\hat{Y}_{t-\tau}$ die optimalen linearen Prognosen von Y_t und $Y_{t-\tau}$ durch $Y_{t-\tau+1}, Y_{t-\tau+2}, \dots, Y_{t-1}$ sind. Wir definieren $\pi_0 = 1$ und $\pi_1 = \rho_1$.

Die partielle Autokorrelationsfunktion

3.29

Wir betrachten einen AR[1]-Prozess (Y_t)

$$\begin{array}{ccccccc} Y_{t-2} & \xrightarrow{\alpha} & Y_{t-1} & \xrightarrow{\alpha} & Y_t & \xrightarrow{\alpha} & Y_{t+1} & \xrightarrow{\alpha} & Y_{t+2} & \xrightarrow{\alpha} & \dots \\ \uparrow & & \uparrow & & \uparrow & & \uparrow & & \uparrow & & \\ \varepsilon_{t-2} & & \varepsilon_{t-1} & & \varepsilon_t & & \varepsilon_{t+1} & & \varepsilon_{t+2} & & \end{array}$$

Die PACF ist

$$\pi_\tau = \begin{cases} 1, & \tau = 0 \\ \rho_1, & \tau = 1 \\ 0, & \tau > 1 \end{cases}$$

Die partielle Autokorrelationsfunktion

3.30

Theorem 24. (Y_t) ist dann und nur dann ein AR[p]-Prozess $Y_t = \alpha_1 Y_{t-1} + \dots + \alpha_p Y_{t-p} + \varepsilon_t$ mit $\alpha_p \neq 0$, wenn $\pi_p \neq 0$ und $\pi_\tau = 0$ für $\tau > p$ gilt.

Die PACF eines AR[p]-Prozesses verschwindet für Lags $\tau > p$. Sie hat die folgende approximative Verteilung für hinreichend großes N :

$$\hat{\pi}_N(\tau) \approx N \left(0, \frac{1}{N} \right) \text{ für } \tau > p$$

Die partielle Autokorrelationsfunktion ■ Schwankungsintervall

3.31

Sei (Y_t) ein AR[p]-Prozess. Für hinreichend großes N erhalten wir das approximative Schwankungsintervall für die empirische partielle Autokorrelation $\hat{\pi}_N(\tau)$ mit $\tau > p$

$$\pm 1.96 \frac{1}{\sqrt{N}}.$$

Dabei ist 1.96 rund das 0.975-Quantil der Standardnormalverteilung.

- Realisationen von $\hat{\pi}_N(\tau)$ außerhalb des Schwankungsintervalls können in lediglich 5% der Fälle vorkommen, wenn der datengenerierende Prozess (Y_t) ein $AR[p]$ -Prozess ist (Irrtumswahrscheinlichkeit 5%).

Modellspezifikation

3.32

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 aus.

Modellspezifikation

3.33

- Verhaltensmuster der ACF: Die ACF $\rho(\tau)$ eines stationären $AR[p]$ -Prozesses klingt exponentiell ab und konvergiert gegen Null für $\tau \rightarrow \infty$.
- Verhaltensmuster der PACF: Die PACF $\pi(\tau)$ eines stationären $AR[p]$ -Prozesses verschwindet für Lags $\tau > p$.

Für genügend lange Reihen können diese Verhaltensmuster anhand der empirischen Kennfunktionen überprüft werden.

Simulationen

3.34

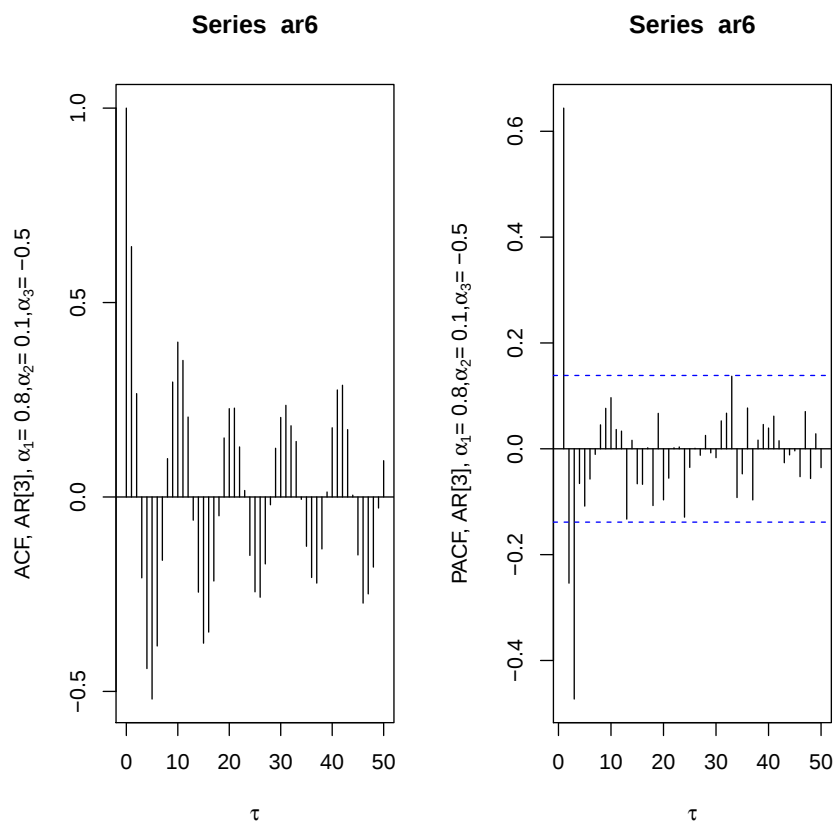


Abb. 3.3:

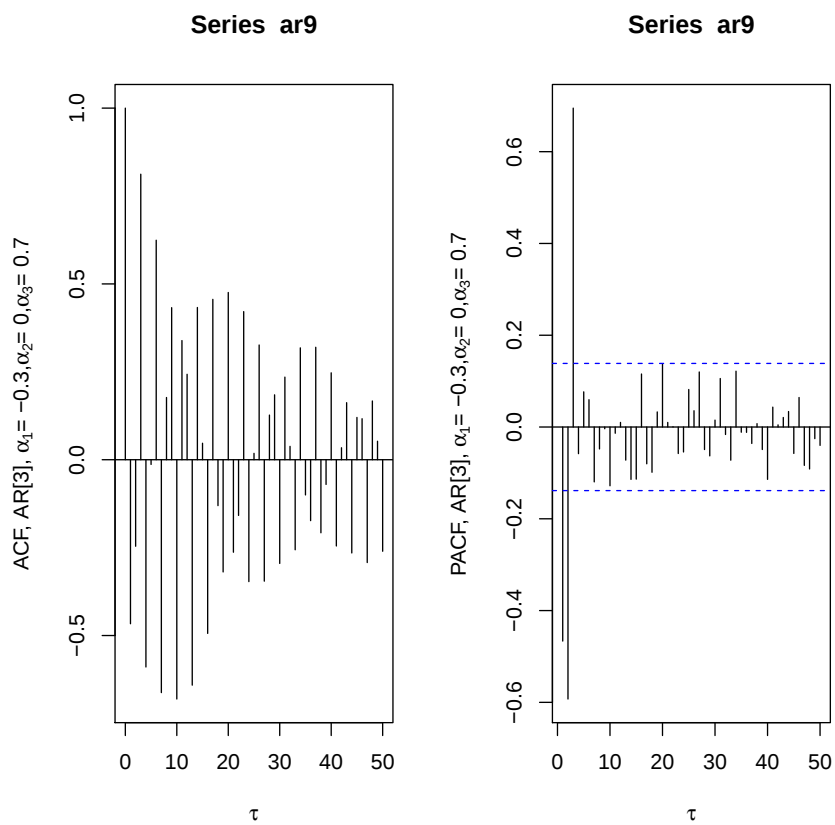


Abb. 3.4:

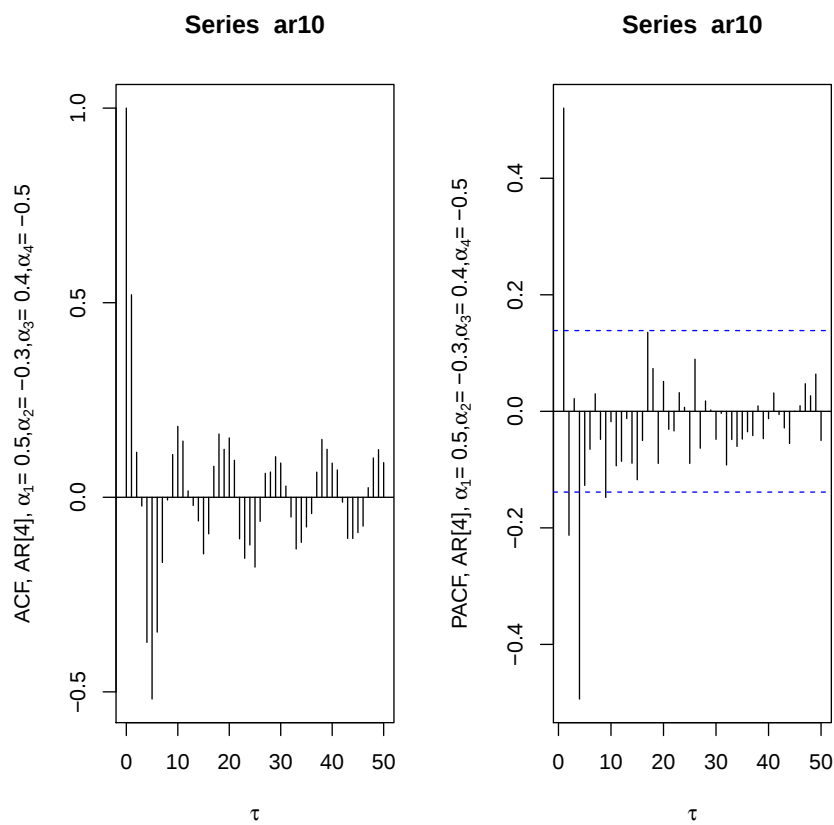


Abb. 3.5:

Simulationen

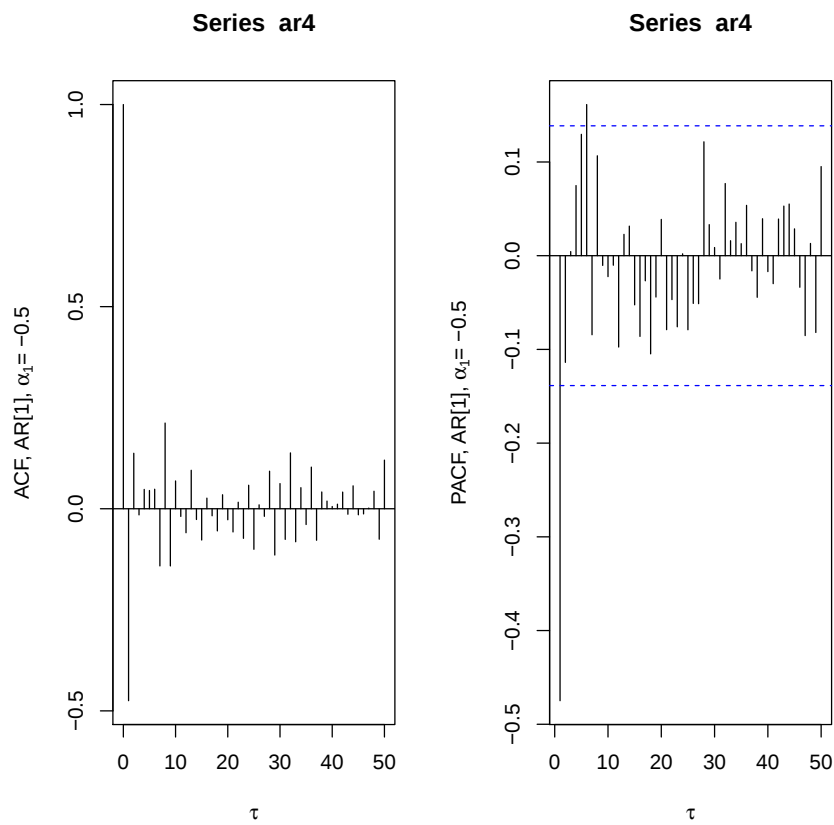


Abb. 3.6:

Simulationen

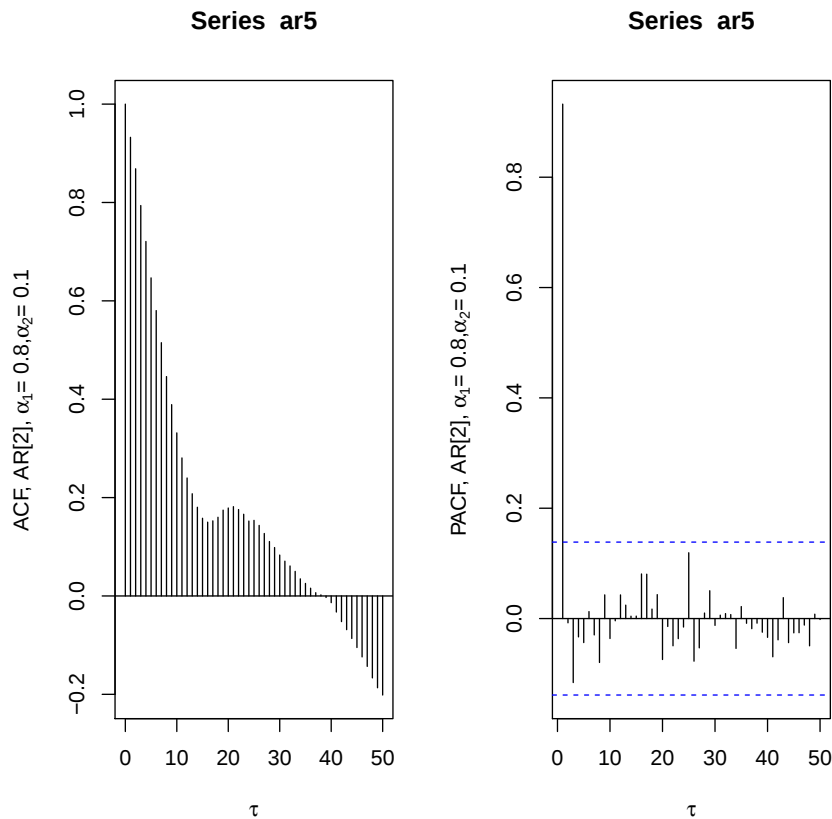


Abb. 3.7:

3.2.2 Invertierbare MA-Prozesse

Invertierbare MA-Prozesse

3.39

Definition 25. Ein $MA[q]$ -Prozess $Y_t = \beta(B) \varepsilon_t$ heißt **invertierbar**, wenn alle Nullstellen des charakteristischen Polynoms

$$\beta(z) = 1 - \beta_1 z - \beta_2 z^2 - \dots - \beta_q z^q$$

außerhalb des Einheitskreises liegen.

Invertierbare $MA[q]$ -Prozesse können in der Form $\varepsilon_t = \beta^{-1}(B) Y_t$ geschrieben werden, wobei

$$\beta^{-1}(B) = c_0 + c_1 B + c_2 B^2 + \dots$$

ein kausaler linearer Filter mit absolut summierbarer Koeffizientenfolge ist. Dies entspricht einer Darstellung des invertierbaren MA-Prozesses als unendlicher AR-Prozess.

Eindeutige Beschreibung mit Hilfe der ACF

3.40

Invertierbare MA-Prozesse sind eindeutig durch ihre ACF definiert. Die ACF ist

$$\rho_Y(\tau) = \frac{\gamma_Y(\tau)}{\gamma_Y(0)} = \begin{cases} 0, & \tau > q \\ \sum_{u=0}^{q-\tau} \beta_u \beta_{u+\tau} / \sum_{u=0}^q \beta_u^2, & 0 < \tau \leq q \\ 1, & \tau = 0 \\ \rho_Y(-\tau), & \tau < 0 \end{cases}$$

Die ACF eines MA[q]-Prozesses verschwindet für Lags $\tau > q$.

Eindeutige Beschreibung mit Hilfe der ACF

3.41

Das Gleichungssystem

$$\rho(\tau) = \sum_{u=0}^{q-\tau} \beta_u \beta_{u+\tau} / \sum_{u=0}^q \beta_u^2, \quad 0 < \tau \leq q,$$

d.h.:

$$\begin{cases} \rho_1 = \frac{\beta_0 \beta_1 + \beta_1 \beta_2 + \beta_2 \beta_3 + \dots + \beta_{q-1} \beta_q}{\beta_0^2 + \beta_1^2 + \beta_2^2 + \dots + \beta_q^2} \\ \rho_2 = \frac{\beta_0 \beta_2 + \beta_1 \beta_3 + \beta_2 \beta_4 + \dots + \beta_{q-2} \beta_q}{\beta_0^2 + \beta_1^2 + \beta_2^2 + \dots + \beta_q^2} \\ \vdots \\ \rho_{q-1} = \frac{\beta_0 \beta_{q-1} + \beta_1 \beta_q}{\beta_0^2 + \beta_1^2 + \beta_2^2 + \dots + \beta_q^2} \\ \rho_q = \frac{\beta_0 \beta_q}{\beta_0^2 + \beta_1^2 + \beta_2^2 + \dots + \beta_q^2} \end{cases}$$

hat q Gleichungen in q Unbekannten. Es ist eindeutig lösbar für invertierbare MA-Prozesse.

Eindeutige Beschreibung mit Hilfe der ACF

3.42

Für die empirische Autokorrelation $\hat{\rho}_N(\tau)$ von einem MA[q]-Prozess, $\tau > q$, gilt für hinreichend großes N :

$$\mathbb{E}[\hat{\rho}_N(\tau)] \approx 0 \quad \text{und} \quad \text{Var}[\hat{\rho}_N(\tau)] \approx \frac{1}{N} (1 + 2\rho_1^2 + 2\rho_2^2 + \dots + 2\rho_q^2)$$

und

$$\hat{\rho}_N(\tau) \approx N \left(0, \frac{1}{N} (1 + 2\rho_1^2 + 2\rho_2^2 + \dots + 2\rho_q^2) \right).$$

Eindeutige Beschreibung mit Hilfe der ACF ■ Schwankungsintervall

3.43

Für hinreichend großes N erhalten wir das approximative Schwankungsintervall für die empirische Autokorrelation $\hat{\rho}_N(\tau)$ von einem $\text{MA}[q]$ -Prozess, $\tau > q$

$$\pm 1.96 \sqrt{(1 + 2\rho_1^2 + 2\rho_2^2 + \dots + 2\rho_q^2) / N}.$$

Dabei ist 1.96 rund das 0.975-Quantil der Standardnormalverteilung.

In praktischen Anwendungen verwenden wir die so genannten Bartlett-Grenzen

$$\pm 1.96 \sqrt{(1 + 2\hat{\rho}_1^2 + 2\hat{\rho}_2^2 + \dots + 2\hat{\rho}_q^2) / N}.$$

- Realisationen von $\hat{\rho}_N(\tau)$ außerhalb des Schwankungsintervalls können in lediglich 5% der Fälle vorkommen, wenn der datengenerierende Prozess ein $\text{MA}[q]$ -Prozess ist (Irrtumswahrscheinlichkeit 5%).

Modellspezifikation

3.44

- Verhaltensmuster der ACF: Die ACF $\rho(\tau)$ eines $\text{MA}[q]$ -Prozesses verschwindet für Lags $\tau > q$.
- Verhaltensmuster der PACF: Die PACF $\pi(\tau)$ eines $\text{MA}[q]$ -Prozesses klingt für $\tau > q$ exponentiell ab und konvergiert gegen Null für $\tau \rightarrow \infty$.

Für genügend lange Reihen können diese Verhaltensmuster anhand der empirischen Kennfunktionen überprüft werden.

Simulationen

3.45

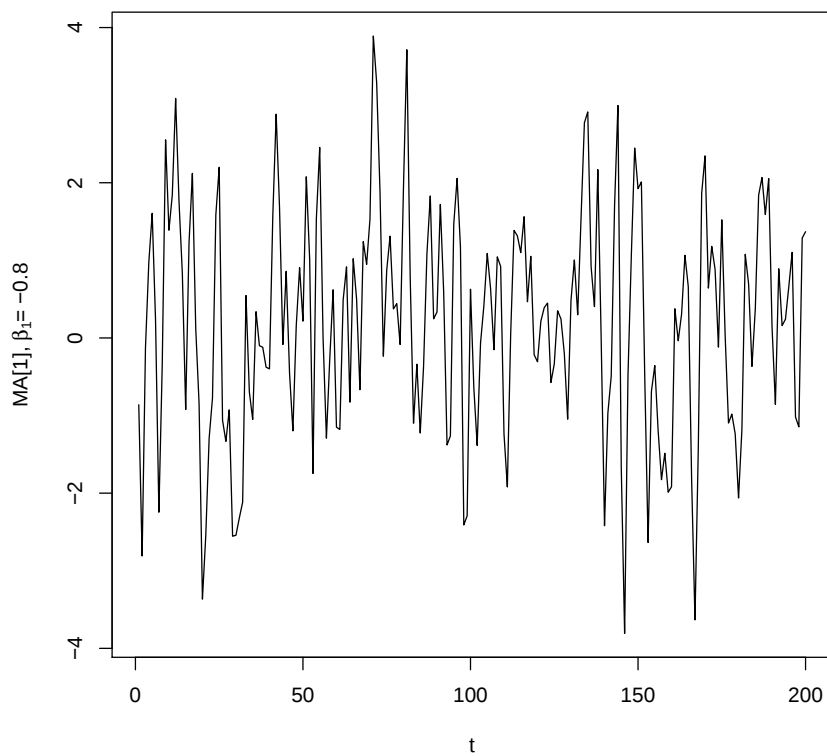


Abb. 3.8: Realisation eines MA-Prozesses

Simulationen

3.46

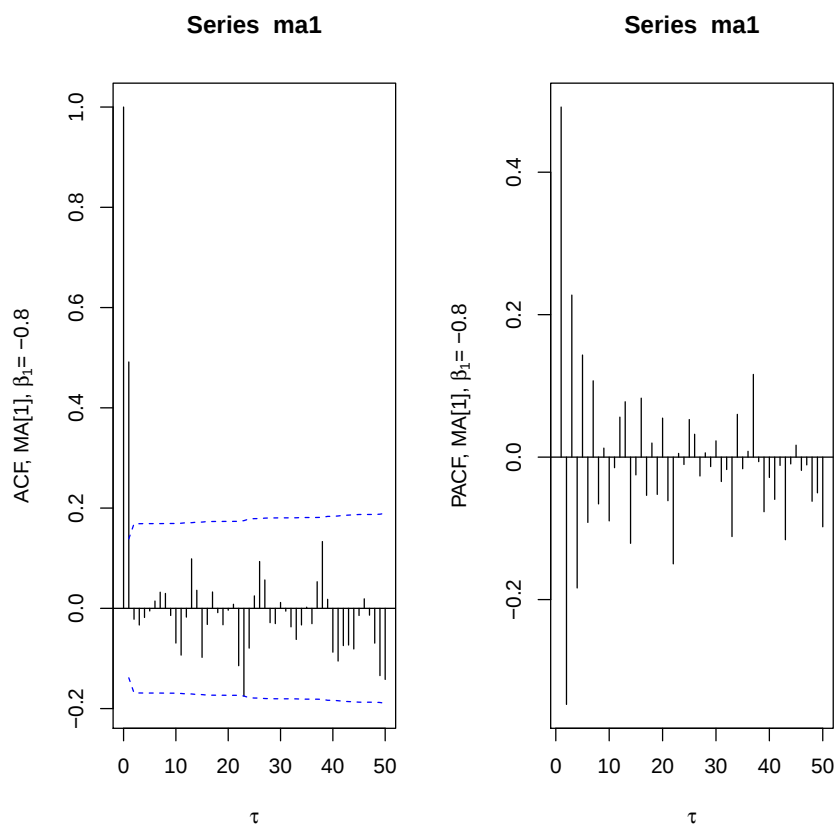


Abb. 3.9:

Simulationen

3.47

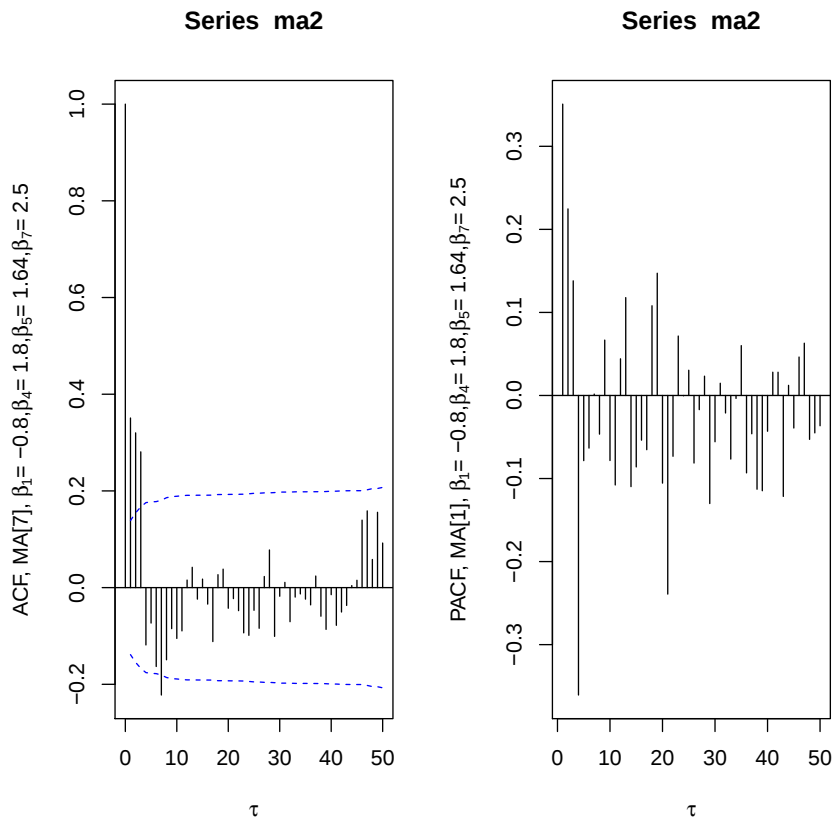


Abb. 3.10:

3.2.3 Stationäre und invertierbare ARMA-Prozesse

Stationäre und invertierbare ARMA-Prozesse

3.48

Theorem 26. 1. Ein ARMA-Prozess $\alpha(B)Y_t = \beta(B)\varepsilon_t$ kann als $MA[\infty]$ -Prozess dargestellt werden

$$Y_t = \alpha(B)^{-1} \beta(B) \varepsilon_t,$$

wenn alle Nullstellen des charakteristischen Polynoms des AR-Teils

$$1 - \alpha_1 z - \alpha_2 z^2 - \dots - \alpha_p z^p = 0$$

außerhalb des Einheitskreises liegen.

Speziell ist (Y_t) dann ein stationärer ARMA-Prozess.

Stationäre und invertierbare ARMA-Prozesse

3.49

Theorem 27. 2. Ein ARMA-Prozess $\alpha(B) Y_t = \beta(B) \varepsilon_t$ kann als $AR[\infty]$ -Prozess dargestellt werden

$$\beta(B)^{-1} \alpha(B) Y_t = \varepsilon_t$$

wenn alle Nullstellen des charakteristischen Polynoms des MA-Teils

$$1 - \beta_1 z - \beta_2 z^2 - \dots - \beta_q z^q = 0$$

außerhalb des Einheitskreises liegen.

Wir sprechen von einem invertierbaren ARMA-Prozess (Y_t) .

Modellspezifikation

3.50

- Verhaltensmuster der ACF: Die ACF $\rho(\tau)$ eines stationären $ARMA[p, q]$ -Prozesses klingt exponentiell ab und konvergiert gegen Null für $\tau \rightarrow \infty$.
- Verhaltensmuster der PACF: Die PACF $\pi(\tau)$ eines stationären $ARMA[p, q]$ -Prozesses klingt exponentiell ab und konvergiert gegen Null für $\tau \rightarrow \infty$.

Die Spezifikation gemischter $ARMA[p, q]$ -Modelle mit Hilfe der ACF und PACF ist eine schwierige Aufgabe.

Simulationen

3.51

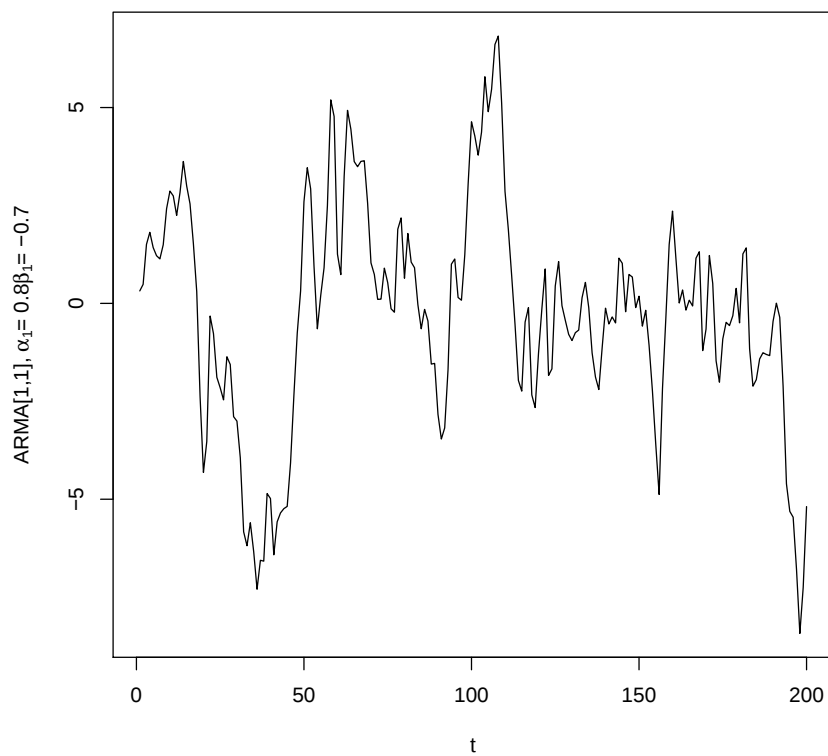


Abb. 3.11: Realisation eines ARMA-Prozesses

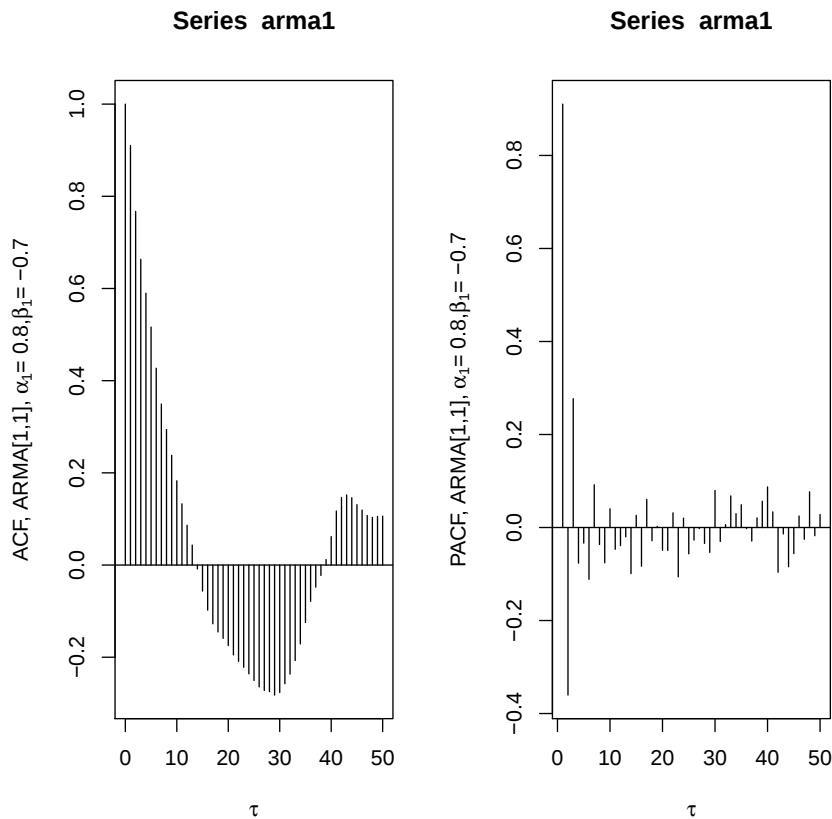


Abb. 3.12:

Ergodizität von ARMA-Prozessen ■ Wiederholung aus Kapitel 1

3.53

Theorem 28. $(Y_t)_{t \in \mathbb{Z}}$ sei ein stationärer Prozess mit Erwartungswert μ und Autokovarianzfunktion (γ_τ) . Ist (γ_τ) absolut summierbar, so ist (Y_t) mittelwertergodisch.

Das bedeutet, dass μ durch $\bar{Y}_N$ konsistent geschätzt werden kann, wenn die Abhängigkeit von weit auseinanderliegenden Zeitpunkten genügend klein ist.

Ergodizität von ARMA-Prozessen ■ Wiederholung aus Kapitel 1

3.54

Theorem 29. Sei (Y_t) ein stationärer Normalprozess mit absolut summierbarer Autokovarianzfunktion (γ_τ) .

alternativ:

Sei (Y_t) ein allgemeiner linearer Prozess $Y_t - \mu = \sum_{u=-\infty}^{\infty} \beta_u \varepsilon_{t-u}$ mit $\mathbb{E}[\varepsilon_t^4] < \infty$ und $\sum |\beta_u| < \infty$. Dann ist (Y_t) kovarianzergodisch.

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 $\mathbf{0}$ 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

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

Modellspezifikation ■ Überprüfung der Parametersignifikanz

3.76

Testhypothesen:

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

Teststatistik:

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

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 $\alpha \in (0, 1)$:

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

- 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 $\alpha^{-1}(B)$ und $\beta^{-1}(B)$. 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