4 Prognose

4.1 Die optimale Prognose aus der endlichen Vergangenheit

4.1

Die optimale Prognose aus der endlichen Vergangenheit

Sei (Y_t) ein stochastischer Prozess mit $\mathbb{E}[Y_t^2] < \infty$ und $(y_t)_{1 \leq t \leq N}$ die zur Verfügung stehende Zeitreihe.

- Gesucht ist eine Prognose für den Wert y_{N+h} auf der Grundlage der Zeitreihe $(y_t)_{1 \leq t \leq N}$.
- Für einen positiven Prognosehorizont $h > 0$ liegt y_{N+h} in der Zukunft, für einen negativen Prognosehorizont $h < 0$ – in der Vergangenheit.
- Eine Prognose wird mit Hilfe einer Prognosefunktion $\widehat{Y}_{N,h}$ der Zeitreihenvariablen des Prozesses

  $$\widehat{Y}_{N,h}\left(Y_1, Y_2, \ldots, Y_N\right)$$

  berechnet.

Die optimale Prognose aus der endlichen Vergangenheit 4.2

Prognosen sind *Schätzungen*, Prognosefunktionen – *Schätzfunktionen* für eine Realisation y_{N+h} des Prozesses.

Die optimale Prognose aus der endlichen Vergangenheit 4.3

Wir betrachten den Prognosefehler

$$Y_{N+h} - \widehat{Y}_{N,h}.$$

Gesucht ist die optimale/beste Prognosefunktion für y_{N+h}, d.h. die Prognosefunktion, welche den mittleren quadratischen Prognosefehler $MSE\left[\widehat{Y}_{N,h}\right]$ minimiert

$$\mathrm{MSE}\left[\widehat{Y}_{N,h}\right] = \mathbb{E}\left[\left(Y_{N+h} - \widehat{Y}_{N,h}\right)^2\right].$$

Die optimale Prognose aus der endlichen Vergangenheit 4.4

Theorem 40 (Optimale Prognose)**.** *Sei (Y_t) ein stochastischer Prozess mit $\mathbb{E}[Y_t^2] < \infty$. Die optimale Prognose von y_{N+h} mit Hilfe einer Prognosefunktion der Zeitreihenvariablen $Y_1, Y_2, \ldots, Y_N$ ist der bedingte Erwartungswert*

$$\widehat{Y}_{N+h}\left(Y_1, Y_2, \ldots, Y_N\right) = \mathbb{E}\left[Y_{N+h} \left| Y_1 = y_1, Y_2 = y_2, \ldots, Y_N = y_N\right.\right].$$

Die optimale lineare Prognose aus der endlichen Vergangenheit 4.5

Theorem 41 (Optimale lineare Prognose). *Sei (Y_t) ein stationärer stochastischer Prozess mit Autokorrelationsfunktion (ρ_τ). Die optimale Prognose von y_{N+h} mit Hilfe einer linearen Prognosefunktion*

$$\widehat{Y}_{N,h}(Y_1, Y_2, \ldots, Y_N) = \sum_{u=0}^{N-1} a_u Y_{N-u}$$

benutzt die Gewichte $(a_u)_{0 \leq u \leq N-1}$, welche das System der Yule-Walker-Gleichungen löst

$$\begin{pmatrix} 1 & \rho_1 & \ldots & \rho_{N-1} \\ \rho_1 & 1 & \ldots & \rho_{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{N-1} & \rho_{N-2} & \ldots & 1 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{N-1} \end{pmatrix} = \begin{pmatrix} \rho_h \\ \rho_{h+1} \\ \vdots \\ \rho_{h+N-1} \end{pmatrix}.$$

Die optimale lineare Prognose aus der endlichen Vergangenheit 4.6

Die optimale lineare Prognose $\widehat{Y}_{N,h}$ aus der endlichen Vergangenheit $\{y_1, y_2, \ldots, y_N\}$ kann mit Hilfe der R-Funktion *TrenchForecasts* aus dem Paket *ltsa* (Linear Time Series Analysis) implementiert werden.

- Dabei werden die Autokovarianzen mit Lags $\tau \in \{0, 1, 2, \ldots, h + N - 1\}$ benötigt. Werden die $h + N$ Werte der Autokovarianzfunktion aus der Zeitreihe $\{y_1, y_2, \ldots, y_N\}$ geschätzt, so können lediglich Prognosen mit negativem Prognosehorizont $h < 0$ erstellt werden. Für Prognosen in die Zukunft muss die Autokovarianzfunktion bekannt sein oder mit einem alternativen Verfahren geschätzt werden.

Für ARIMA-Modelle kann die optimale lineare Prognose $\widehat{Y}_{N,h}$ aus der endlichen Vergangenheit auch mit Hilfe der Funktion *predict* berechnet werden.

4.2 Die optimale lineare Prognose aus der unendlichen Vergangenheit

Die optimale lineare Prognose aus der unendlichen Vergangenheit ▪ Minimierung des mittleren quadratischen Prognosefehlers 4.7

Wir betrachten im Folgenden lineare Prognosen aus der unendlichen Vergangenheit

$$\widehat{Y}_{t,h}(Y_t, Y_{t-1}, \ldots) = \sum_{u=0}^{\infty} a_u Y_{t-u}.$$

Der Prognosefehler ist

$$Y_{t+h} - \widehat{Y}_{t,h}.$$

Gesucht ist die optimale/beste Prognosefunktion für y_{t+h}, d.h. die Prognosefunktion, welche den mittleren quadratischen Prognosefehler $\mathrm{MSE}\left[\widehat{Y}_{t,h}\right]$ minimiert

$$\mathrm{MSE}\left[\widehat{Y}_{t,h}\right] = \mathbb{E}\left[\left(Y_{t+h} - \widehat{Y}_{t,h}\right)^2\right].$$

Die optimale lineare Prognose aus der unendlichen Vergangenheit ▪ Minimierung des mittleren quadratischen Prognosefehlers 4.8

Diese Minimierungsaufgabe führt zu einem unendlichen System von Yule-Walker-Gleichungen mit den Unbekannten (a_u)

$$\begin{pmatrix} 1 & \rho_1 & \dots \\ \rho_1 & 1 & \dots \\ \vdots & \vdots & \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ \vdots \end{pmatrix} = \begin{pmatrix} \rho_h \\ \rho_{h+1} \\ \vdots \end{pmatrix}.$$

Die optimale lineare Prognose aus der unendlichen Vergangenheit ▪ Prognose für MA$[\infty]$-Prozesse 4.9

Sei (Y_t) ein MA$[\infty]$-Prozess

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

wobei $\beta_0 = -1$ gesetzt wurde.

Die optimale lineare Prognose für den Wert y_{t+h} ist

$$\widehat{Y}_{t,h} = -\sum_{u=0}^{\infty} \beta_{h+u} \varepsilon_{t-u}.$$

Die optimale lineare Prognose aus der unendlichen Vergangenheit ▪ Unkorreliertheit zwischen dem Prognosefehler und den Zeitreihenvariablen 4.10

Alternativ kann die Folge (a_u) mit Hilfe des folgenden Satzes bestimmt werden.

Theorem 42. *Sei (Y_t) ein rein nicht-deterministischer Prozess. Die lineare Prognose aus der unendlichen Vergangenheit*

$$\widehat{Y}_{t,h}\left(Y_t, Y_{t-1}, \dots\right) = \sum_{u=0}^{\infty} a_u Y_{t-u}$$

ist dann und nur dann optimal, wenn der Prognosefehler $Y_{t+h} - \widehat{Y}_{t,h}$ unkorreliert mit allen Zeitreihenvariablen $Y_t, Y_{t-1}, \dots$ ist.

Die optimale lineare Prognose aus der unendlichen Vergangenheit ▪ Unkorreliertheit zwischen dem Prognosefehler und den Zeitreihenvariablen 4.11

Wir betrachten den Prognosefehler zum Zeitpunkt $t+h$

$$Y_{t+h} - \widehat{Y}_{t,h}$$

und den mittleren quadratischen Prognosefehler

$$\mathrm{MSE}\left[\widehat{Y}_{t,h}\right] = \mathbb{E}\left[\left(Y_{t+h} - \widehat{Y}_{t,h}\right)^2\right].$$

Sei $\Delta(h)$ der mittlere quadratische Prognosefehler für die optimale lineare Prognose

$$\Delta(h) := \min_{(a_u)} \mathrm{MSE}\left[\widehat{Y}_{t,h}\right].$$

Die optimale lineare Prognose aus der unendlichen Vergangenheit ▪ Unkorreliertheit zwischen dem Prognosefehler und den Zeitreihenvariablen 4.12

Definition 43 (deterministischer Prozess)**.** Ein stationärer Prozess (Y_t) heißt **deterministisch** (singulär, exakt vorhersagbar), wenn gilt

$$\Delta(h) = 0 \quad \forall h > 0.$$

Definition 44 (rein nicht-deterministischer Prozess)**.** Ein stationärer Prozess (Y_t) heißt rein nicht-deterministisch, wenn gilt

$$\lim_{h\to\infty} \Delta(h) = \gamma_0.$$

Ein stationärer Prozess ist dann und nur dann rein nicht-deterministisch, wenn er sich als MA[∞]-Prozess

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

mit quadratisch summierbaren Gewichten $\sum_{u=0}^{\infty} \beta_u^2 < \infty$ darstellen lässt.

Die optimale lineare Prognose aus der unendlichen Vergangenheit ▪ Unkorreliertheit zwischen dem Prognosefehler und den Zeitreihenvariablen 4.13

Theorem 45 (Wold'scher Zerlegungssatz)**.** *Jeder stationäre Prozess (Y_t) kann eindeutig als Summe zweier unkorrelierter stationärer Prozesse, nämlich eines deterministischen und eines rein nicht-deterministischen Prozesses dargestellt werden.*

Weiter werden wir uns nur mit der Prognose rein nicht-deterministischer Prozesse beschäftigen.

Theorem 46. *Sei (Y_t) ein rein nicht-deterministischer Prozess. Die lineare Prognose aus der unendlichen Vergangenheit*

$$\widehat{Y}_{t,h}\left(Y_t, Y_{t-1}, \ldots\right) = \sum_{u=0}^{\infty} a_u Y_{t-u}$$

ist dann und nur dann optimal, wenn der Prognosefehler $Y_{t+h} - \widehat{Y}_{t,h}$ unkorreliert mit allen Zeitreihenvariablen $Y_t, Y_{t-1}, \ldots$ ist.

4.2.1 Der Box-Jenkins-Ansatz

Der Box-Jenkins-Ansatz 4.14

Der Box-Jenkins-Ansatz ist ein *rekursives* Berechnungsverfahren der optimalen linearen Prognose aus der unendlichen Vergangenheit für allgemeine ARIMA-Prozesse.

Der Box-Jenkins-Ansatz ▪ Prognose für MA$[q]$-Prozesse 4.15

Sei (Y_t) ein MA$[q]$-Prozess

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

wobei $\beta_0 = -1$ gesetzt wurde. Wir können (Y_t) als MA$[\infty]$-Prozess mit Parametern $\beta_u = 0$ für $u > q$ darstellen.

Dann lautet die optimale lineare Prognose für den Wert y_{t+h}

$$\widehat{Y}_{t,h} = \begin{cases} -\sum_{u=0}^{q-h} \beta_{h+u} \varepsilon_{t-u}, & h \leq q \\ 0, & h > q \end{cases}$$

Der Box-Jenkins-Ansatz ▪ Prognose für MA$[q]$-Prozesse 4.16

Die Störungen $(\varepsilon_u)_{t+h-q \leq u \leq t}$ entsprechen den Prognosefehlern der 1-Schritt-Prognosen

$$\begin{aligned} \varepsilon_t &= Y_t - \widehat{Y}_{t-1,1} \\ \varepsilon_{t-1} &= Y_{t-1} - \widehat{Y}_{t-2,1} \\ \varepsilon_{t-2} &= Y_{t-2} - \widehat{Y}_{t-3,1} \\ &\vdots \\ \varepsilon_{t-h-q} &= Y_{t+h-q} - \widehat{Y}_{t+h-q-1,1} \end{aligned}$$

Sie werden *rekursiv* berechnet.

Der Box-Jenkins-Ansatz ▪ Prognose für MA[q]-Prozesse 4.17

Wir schreiben die Definitionsgleichung für Y_{t+h} fort

$$Y_{t+h} = \varepsilon_{t+h} - \beta_1 \varepsilon_{t+h-1} - \ldots - \beta_h \varepsilon_t - \beta_{h+1} \varepsilon_{t-1} - \ldots - \beta_q \varepsilon_{t+h-q}.$$

Die Prognose $\widehat{Y}_{t,h}$ wird aus Y_{t+h}, wie folgt, berechnet

$$\widehat{Y}_{t,h} = \varepsilon_{t+h} - \beta_1 \varepsilon_{t+h-1} - \ldots - \beta_h \varepsilon_t - \beta_{h+1} \varepsilon_{t-1} - \ldots - \beta_q \varepsilon_{t+h-q}.$$

- Die Störungen $\varepsilon_{t+h}, \varepsilon_{t+h-1}, \ldots, \varepsilon_{t+1}$ werden mit ihrem Erwartungswert von Null ersetzt.
- Die Störungen $\varepsilon_t, \varepsilon_{t-1}, \ldots$ entsprechen den Prognosefehlern $Y_t - \widehat{Y}_{t-1,1}, Y_{t-1} - \widehat{Y}_{t-2,1}, \ldots$ der 1-Schritt-Prognosen.

Der Box-Jenkins-Ansatz ▪ Prognose für AR[p]-Prozesse 4.18

Sei (Y_t) ein AR[p]-Prozess

$$Y_t = \alpha_1 Y_{t-1} + \alpha_2 Y_{t-2} + \ldots + \alpha_p Y_{t-p} + \varepsilon_t.$$

Die optimalen linearen Prognosen sind

$$\begin{aligned} \widehat{Y}_{t,1} &= \alpha_1 Y_t + \alpha_2 Y_{t-1} + \ldots + \alpha_p Y_{t+1-p} \\ \widehat{Y}_{t,2} &= \alpha_1 \widehat{Y}_{t,1} + \alpha_2 Y_t + \alpha_3 Y_{t-1} + \ldots + \alpha_p Y_{t+2-p} \end{aligned}$$

Für einen beliebigen Prognosehorizont h

$$\begin{aligned} \widehat{Y}_{t,h} &= \alpha_1 \widehat{Y}_{t,h-1} + \alpha_2 \widehat{Y}_{t,h-2} + \ldots + \alpha_{h-1} \widehat{Y}_{t,1} + \\ &\quad \alpha_h Y_t + \alpha_{h+1} Y_{t-1} + \ldots + \alpha_p Y_{t+h-p} \end{aligned}$$

Der Box-Jenkins-Ansatz ▪ Prognose für AR[p]-Prozesse 4.19

Wir schreiben die Definitionsgleichung für Y_{t+h} fort

$$Y_{t+h} = \alpha_1 Y_{t+h-1} + \alpha_2 Y_{t+h-2} + \ldots + \alpha_p Y_{t+h-p} + \varepsilon_{t+h}.$$

Die Prognose $\widehat{Y}_{t,h}$ wird aus Y_{t+h}, wie folgt, berechnet

$$\widehat{Y}_{t,h} = \alpha_1 Y_{t+h-1} + \alpha_2 Y_{t+h-2} + \ldots + \alpha_p Y_{t+h-p} + \varepsilon_{t+h}.$$

- Die unbekannten Variablen $Y_{t+h-1}, \ldots, Y_{t+1}$ werden mit den optimalen linearen Prognosen $\widehat{Y}_{t,h-1}, \ldots, \widehat{Y}_{t,1}$ ersetzt.

- Für die Variablen $Y_t, Y_{t-1}, \ldots$ setzen wir ihre Beobachtungen ein.
- Die Störung ε_{t+h} wird mit ihrem Erwartungswert von Null ersetzt.

Der Box-Jenkins-Ansatz ▪ Prognose für ARMA$[p,q]$-Prozesse 4.20

Sei (Y_t) ein ARMA$[p,q]$-Prozess

$$Y_t = \alpha_1 Y_{t-1} + \ldots + \alpha_p Y_{t-p} + \varepsilon_t - \beta_1 \varepsilon_{t-1} - \ldots - \beta_q \varepsilon_{t-q}.$$

Wir schreiben die Definitionsgleichung für Y_{t+h} fort

$$Y_{t+h} = \alpha_1 Y_{t+h-1} + \ldots + \alpha_p Y_{t+h-p} + \varepsilon_{t+h} - \beta_1 \varepsilon_{t+h-1} - \ldots - \beta_q \varepsilon_{t+h-q}.$$

Der Box-Jenkins-Ansatz ▪ Prognose für ARMA$[p,q]$-Prozesse 4.21

Die Prognose $\widehat{Y}_{t,h}$ wird aus Y_{t+h}, wie folgt, berechnet

$$\widehat{Y}_{t,h} = \alpha_1 Y_{t+h-1} + \ldots + \alpha_p Y_{t+h-p} + \varepsilon_{t+h} - \beta_1 \varepsilon_{t+h-1} - \ldots - \beta_q \varepsilon_{t+h-q}.$$

- Die unbekannten Variablen $Y_{t+h-1}, \ldots, Y_{t+1}$ werden mit den optimalen linearen Prognosen $\widehat{Y}_{t,h-1}, \ldots, \widehat{Y}_{t,1}$ ersetzt.
- Für die Variablen $Y_t, Y_{t-1}, \ldots$ setzen wir ihre Beobachtungen ein.
- Die Störungen $\varepsilon_{t+h}, \varepsilon_{t+h-1}, \ldots, \varepsilon_{t+1}$ werden mit ihrem Erwartungswert von Null ersetzt.
- Die Störungen $\varepsilon_t, \varepsilon_{t-1}, \ldots$ entsprechen den Prognosefehlern $Y_t - \widehat{Y}_{t-1,1}, Y_{t-1} - \widehat{Y}_{t-2,1}, \ldots$ der 1-Schritt-Prognosen.

Der Box-Jenkins-Ansatz ▪ Prognose für ARMA$[p,q]$-Prozesse 4.22

Wir schreiben den ARMA$[p,q]$-Prozess (Y_t), wie folgt, auf:

$$Y_t = c\,(\mathrm{B})\,\varepsilon_t$$

mit

$$c\,(\mathrm{B}) = \alpha^{-1}\,(\mathrm{B})\,\beta\,(\mathrm{B})\,.$$

Die Varianz des Prognosefehlers ist

$$\mathrm{Var}\Big[Y_{t+h} - \widehat{Y}_{t,h}\Big] = \sigma_\varepsilon^2 \sum_{u=0}^{h-1} c_u^2$$

mit (c_u) die Gewichte des Filters $c\,(\mathrm{B})$.

Der Box-Jenkins-Ansatz ▪ Prognose für ARMA$[p,q]$-Prozesse 4.23

Die optimale lineare Prognose eines ARMA$[p,q]$-Modells hat die folgenden Eigenschaften:

$$\lim_{h\to\infty} \operatorname{Var}\left[Y_{t+h} - \widehat{Y}_{t,h}\right] = \operatorname{Var}[Y_t]$$

$$\lim_{h\to\infty} \widehat{Y}_{t,h} = \mathbb{E}[Y_t] = 0$$

Daher ist sie *unbrauchbar für großen Prognosehorizont* h.

Der Box-Jenkins-Ansatz ▪ Prognose für ARMA$[p,q]$-Prozesse 4.24

Sei (Y_t) ein Gaußscher ARMA$[p,q]$-Prozess. Dann gilt

$$Y_{t+h} - \widehat{Y}_{t,h} \sim N\left(0, \sigma_\varepsilon^2 \sum_{u=0}^{h-1} c_u^2\right).$$

Prognoseintervall für Y_{t+h}:

$$\widehat{Y}_{t,h} \pm z_{1-\alpha/2}\,\sigma_\varepsilon \sqrt{\sum_{u=0}^{h-1} c_u^2}$$

mit $z_{1-\alpha/2}$ das $1-\alpha\,/2$-Quantil der Standardnormalverteilung.

Der Box-Jenkins-Ansatz ▪ Prognose für angepasste ARMA$[p,q]$-Prozesse 4.25

Der Box-Jenkins-Ansatz ist ein rekursives Berechnungsverfahren für die optimale lineare Prognose $\widehat{Y}_{t,h}$, wenn

- die Parameter $\boldsymbol{\alpha} = \begin{pmatrix} \alpha_1 & \alpha_2 & \dots & \alpha_p \end{pmatrix}$ und $\boldsymbol{\beta} = \begin{pmatrix} \beta_1 & \beta_2 & \dots & \beta_q \end{pmatrix}$ bekannt sind,
- eine unendliche Zeitreihe $y_{-\infty}, \dots, y_t$ zur Verfügung steht (unendliche Vergangenheit).

Der Box-Jenkins-Ansatz ▪ Prognose für angepasste ARMA$[p,q]$-Prozesse 4.26

In praktischen Anwendungen verwenden wir eine angepasste Prognosefunktion auf der Grundlage von $\widehat{Y}_{t,h}$. Dabei setzen wir die folgenden Werte ein:

- die Parameterschätzungen $\widehat{\boldsymbol{\alpha}} = \begin{pmatrix} \widehat{\alpha}_1 & \widehat{\alpha}_2 & \dots & \widehat{\alpha}_p \end{pmatrix}$ und $\widehat{\boldsymbol{\beta}} = \begin{pmatrix} \widehat{\beta}_1 & \widehat{\beta}_2 & \dots & \widehat{\beta}_q \end{pmatrix}$ nach einer der Methoden CLS, ULS oder ML,
- geeignete Startwerte $\widetilde{\varepsilon}_0, \widetilde{\varepsilon}_{-1}, \dots, \widetilde{Y}_0, \widetilde{Y}_{-1}, \dots$

- Schätzungen (Approximationen) $\widetilde{\varepsilon}_t, \ldots \widetilde{\varepsilon}_1$ für die Störungen $\varepsilon_t, \ldots \varepsilon_1$.

Diese Prognosefunktion bezeichnen wir im Folgenden mit $\widetilde{Y}_{t,h}$.

Der Box-Jenkins-Ansatz ▪ Prognose für angepasste ARMA$[p, q]$-Prozesse 4.27

Theorem 47. *Sei (Y_t) ein stationärer und invertierbarer ARMA$[p, q]$-Prozess, $\alpha(\mathrm{B}) Y_t = \beta(\mathrm{B}) \varepsilon_t$, mit White-Noise-Prozess (ε_t). Sei $\widetilde{Y}_{t,h}$ die entsprechend angepasste Prognosefunktion auf der Grundlage der optimalen linearen Prognose aus der unendlichen Vergangenheit $\widehat{Y}_{t,h}$. Dann gilt für $\varepsilon > 0$ beliebig*

$$\lim_{N\to\infty} \mathbb{P}\left[\sqrt{N}\left|\widetilde{Y}_{t,h} - \widehat{Y}_{t,h}\right| > \varepsilon\right] = 0.$$

- Die Anwendung der Prognosefunktion $\widetilde{Y}_{t,h}$ anstatt von $\widehat{Y}_{t,h}$ hat einen vernachlässigbaren Effekt auf die Prognosegüte.

Literaturhinweise 4.28

Schlittgen und Streitberg (1999), Zeitreihenanalyse: Kapitel 4.1.1, 4.2.1, 4.3.1

Schlittgen (2015), Angewandte Zeitreihenanalyse mit `R`: Kapitel 5.2

4.3 Die exponentielle Glättung

Die exponentielle Glättung 4.29

Die exponentielle Glättung ist ein Extrapolationsverfahren für die Berechnung von Prognosen. Es basiert auf das Bilden rekursiver Fortschreibungen der Form

$$\widehat{y}_{N,1} = (1-\alpha)\widehat{y}_{N-1,1} + \alpha y_N.$$

Durch rekursives Einsetzen der Prognosen erhalten wir die Darstellung als exponentiell gewichtete Summe von Zeitreihenwerten

$$\widehat{y}_{N,1} = (1-\alpha)^{N-1} y_1 + \alpha \sum_{u=0}^{N-2} (1-\alpha)^u y_{N-u}.$$

Dabei wurde das Verfahren mit $\widehat{y}_{1,1} = y_1$ initialisiert.

Die exponentielle Glättung 4.30

Alternative Darstellung:

$$\widehat{y}_{N,1} = \widehat{y}_{N-1,1} + \alpha(y_N - \widehat{y}_{N-1,1})$$

Eine neue Prognose wird durch Aufdatierung der letzten Prognose erstellt. Dabei erfolgt eine Korrektur um den letzten Porgnosefehler.

Literaturhinweise 4.31

Schlittgen (2015), Angewandte Zeitreihenanalyse mit R: Kapitel 5.1

4.4 Auswertung von Prognoseergebnissen

Auswertung von Prognoseergebnissen 4.32

Prognoseverfahren:

- optimale lineare Prognose $\widehat{Y}_{N,h}$ aus der endlichen Vergangenheit $\{y_1, y_2, \ldots, y_N\}$. Berechnung durch Lösen der Yule-Walker-Gleichungen mit Hilfe des Levinson-Durbin-Algorithmus.
- optimale lineare Prognose $\widehat{Y}_{t,h}$ aus der unendlichen Vergangenheit $\{y_t, y_{t-1}, y_{t-2} \ldots\}$. Speziell für ARIMA-Modelle kann $\widehat{Y}_{t,h}$ mit Hilfe des Box-Jenkins-Ansatzes berechnet werden.

Auswertung von Prognoseergebnissen 4.33

Gegeben sei eine Zeitreihe (y_t) der Länge N, $t \in \{1, 2, \ldots, N\}$. Wir unterscheiden zwischen

- in-sample Prognose (innerhalb der Stichprobe)**:** wir schätzen das Zeitreihenmodell (ZR-Modell) aus der Zeitreihe (y_t) und erstellen Prognosen ebenfalls für (y_t).
- **out-of-sample Prognose (außerhalb der Stichprobe):** wir schätzen das ZR-Modell aus der Zeitreihe (y_t). Die Prognosen werden für einen unterschiedlichen, zukünftigen Zeitraum $s \in \{N+1, N+2, \ldots, N+H\}$ erstellt.

Auswertung von Prognoseergebnissen 4.34

Gegeben sei eine Zeitreihe (y_t) der Länge N. Wir erstellen Prognosen $(\widehat{y}_s)$ für den Prognosezeitraum $s \in \{N+1, N+2, \ldots, N+H\}$. Wir untersuchen die Prognosegüte von $(\widehat{y}_s)$ ex-post, d.h. durch Vergleiche zwischen den Prognosen $(\widehat{y}_s)$ und den tatsächlichen Realisationen (y_s).

Fragestellung 1: Wir berechnen $(\widehat{y}_s)$. Wir möchten die Prognosegüte von $(\widehat{y}_s)$ beurteilen.

Fragestellung 2: Wir erstellen Prognosen $\left(\widehat{y}_s^A\right)$ und $\left(\widehat{y}_s^B\right)$ mit zwei unterschiedlichen ZR-Modellen oder zwei unterschiedlichen Prognoseverfahren A und B. Wir möchten bestimmen, welches Modell/Verfahren die beste Prognose geliefert hat.

Auswertung von Prognoseergebnissen ▪ Empirische Maßzahlen 4.35

Wir berechnen

- den empirischen mittleren absoluten Prognosefehler

$$\widehat{MAE}\left(\widehat{y}_s\right) = \frac{1}{H} \sum_{s=N+1}^{N+H} \left|y_s - \widehat{y}_s\right|$$

- den empirischen mittleren quadratischen Prognosefehler

$$\widehat{MSE}\left(\widehat{y}_s\right) = \frac{1}{H} \sum_{s=N+1}^{N+H} \left(y_s - \widehat{y}_s\right)^2 .$$

Niedrige Werte dieser empirischen Maßzahlen sprechen für eine gute Prognose $\left(\widehat{y}_s\right)$.

Auswertung von Prognoseergebnissen ▪ Mincer-Zarnowitz-Regression 4.36

Wir führen eine einfache lineare Regression der Realisationen (y_s) auf die Prognosen $\left(\widehat{y}_s\right)$

$$y_s = \beta_0 + \beta_1 \widehat{y}_s + u_s, \quad s = N+1, N+2, \ldots N+H$$

durch.

- Eine gute Prognose $\left(\widehat{y}_s\right)$ ist durch ein hohes Bestimmtheitsmaß R^2 gekennzeichnet.
- Auch können wir einen F-Test mit den Testhypothesen:

$$H_0 : \beta_0 = 0 \wedge \beta_1 = 1 \quad \text{gegen} \quad H_1 : \beta_0 \neq 0 \vee \beta_1 \neq 1$$

  durchführen. Die Beibehaltung der H_0-Hypothese weist auf eine gute Prognose $\left(\widehat{y}_s\right)$ hin.

Achtung! Der Störterm (u_s) ist nicht White-Noise. Dazu noch mehr in Kapitel „Regressionsmodelle für Zeitreihen“.

Auswertung von Prognoseergebnissen 4.37

Wir betrachten die Prognosen $\left(\widehat{y}_s^A\right)$ und $\left(\widehat{y}_s^B\right)$ mit zwei unterschiedlichen ZR-Modellen/Prognoseverfahren A und B. Wir können die beiden Prognosen bzgl.

- der empirischen Maßzahlen

- der Güte der jeweiligen Mincer-Zarnowitz-Regressionen

miteinander vergleichen.

Diese Resultate werden auf der Grundlage einer Realisation (y_s) ermittelt und sind dem Zufall überlassen. Neue Realisationen (y_s) führen möglicherweise zu unterschiedlichen Resultaten. Wir brauchen Methoden, womit wir einen statistisch signifikanten Unterschied zwischen $\left(\widehat{y}_s^A\right)$ und $\left(\widehat{y}_s^B\right)$ belegen können.

Auswertung von Prognoseergebnissen ▪ Diebold-Mariano-Test 4.38

Wir möchten die Prognosen $\left(\widehat{y}_s^A\right)$ und $\left(\widehat{y}_s^B\right)$ mit zwei ZR-Modellen/Prognoseverfahren A und B bzgl. ihrer Genauigkeit vergleichen. Wir betrachten

- den Abstand zwischen ihren quadrierten Prognosefehlern zum Zeitpunkt s

$$d_s = \left(y_s - \widehat{y}_s^A\right)^2 - \left(y_s - \widehat{y}_s^B\right)^2$$

- den Abstand zwischen ihren empirischen MSE-Werten

$$\begin{aligned} \overline{d} &= \widehat{MSE}\left(\widehat{y}_s^A\right) - \widehat{MSE}\left(\widehat{y}_s^B\right) \\ &= \frac{1}{H}\sum_{s=N+1}^{N+H}\left(y_s - \widehat{y}_s^A\right)^2 - \frac{1}{H}\sum_{s=N+1}^{N+H}\left(y_s - \widehat{y}_s^B\right)^2. \end{aligned}$$

Auswertung von Prognoseergebnissen ▪ Diebold-Mariano-Test 4.39

Theorem 48 (Prognosen mit gleicher Prognosegenauigkeit)**.** *Zwei Prognosen $\left(\widehat{y}_s^A\right)$ und $\left(\widehat{y}_s^B\right)$ haben die gleiche Genauigkeit dann und nur dann wenn*

$$E\left[d_s\right] = 0$$

für alle Prognosezeitpunkte s gilt.

Auswertung von Prognoseergebnissen ▪ Diebold-Mariano-Test 4.40

Testhypothesen:

$$H_0 : E\left[d_s\right] = 0 \quad \forall s \quad \text{gegen} \quad H_1 : \exists s \quad E\left[d_s\right] = 0$$

Teststatistik:

$$\frac{\sqrt{H}\,\overline{d}}{\sqrt{\widehat{\omega}}} \xrightarrow{d} N\left(0,\ 1\right)|H_0$$

mit $\widehat{\omega}$ einem asymptotischen konsistenten Schätzer der Größe $Var\left[\sqrt{H}\,\overline{d}\right]$.

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

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

  Dann haben die Prognosen $\left(\widehat{y}_s^A\right)$ und $\left(\widehat{y}_s^B\right)$ unterschiedliche Prognosegenauigkeit (Irrtumswahrscheinlichkeit α).