

Die Berechnung von zuverlässigen Prognosen und eine entsprechende quantitative Analyse der Prognose-Güte ist eine der wichtigsten Anwendungen der Zeitreihenanalyse. Allgemein geht es bei der Prognose darum eine zukünftige Prozessvariable  $x_{t+h}$  möglichst gut durch eine Funktion

$$\hat{x}_{t,h} = g(x_t, x_{t-1}, \dots)$$

der beobachteten Werte bis zur Gegenwart  $t$  zu approximieren. Dabei ist  $h > 0$  der sogenannte *Prognosehorizont*. Um das Problem exakt zu formulieren, muss man die Funktionenklasse, d. h. die Menge derartiger Prognosefunktionen  $g(\cdot)$ , sowie ein Maß für die Güte der Approximation angeben. Wir diskutieren hier ein spezielles Prognoseproblem, die sogenannte lineare Kleinst-Quadrate-Prognose. Das heißt, wir beschränken uns auf *lineare* (genauer gesagt *affine*) Prognosefunktionen, d. h. auf Funktionen der Form

$$\hat{x}_{t,h} = g(x_t, x_{t-1}, \dots) = c_0 + c_1 x_t + c_2 x_{t-1} + \dots$$

und auf den *mittleren, quadratischen Prognosefehler* („mean squared error“, MSE)

$$\mathbf{E}(x_{t+h} - \hat{x}_{t,h})'(x_{t+h} - \hat{x}_{t,h})$$

als Gütekriterium. Hier betrachten wir ein idealisiertes Problem, bei dem wir annehmen, dass wir die Eigenschaften des zugrundeliegenden Prozesses (Erwartungswert und Kovarianzfunktion) exakt kennen. Dieses idealisierte Prognoseproblem lässt sich einfach und elegant mit Hilfe des Projektionssatzes behandeln. Für eine „echte“ Prognose müssen die Populationsmomente zuerst aus Daten geschätzt werden.

Ohne die Einschränkung auf lineare Funktionen ist der bedingte Erwartungswert  $\mathbf{E}[x_{t+h} | x_t, x_{t-1}, \dots]$  die optimale Kleinst-Quadrate Approximation von  $x_{t+h}$  durch die vergangenen Werte. Die Berechnung dieses bedingten Erwartungswertes benötigt aber

i. Allg. die gemeinsame Verteilung der betrachteten Zufallsvariablen und ist daher in der Praxis oft nur schwer zu berechnen bzw. zu schätzen.

Obwohl die Einschränkung auf lineare Prognosen und das quadratische Gütemaß eine Restriktion bedeutet, ist diese Prognose doch die häufigst verwendete. In Fällen in denen man Fehlprognosen Kosten zuordnen kann und diese Kosten bei Unter- bzw. Überschätzung um den gleichen Betrag sehr unterschiedlich sind, sind andere (nicht symmetrische) Verlustfunktionen zur Optimierung der Prognose angezeigt.

Im ersten Abschnitt dieses Kapitels werden die letzten  $k$  beobachteten Werte für die Prognose verwendet. Man spricht daher auch von der Prognose aus der endlichen Vergangenheit. Der Übergang  $k \rightarrow \infty$ , der entsprechend Prognose aus der unendlichen Vergangenheit genannt wird, führt dann zu der sogenannten Wold-Zerlegung von stationären Prozessen. Siehe [27, 44]. Diese Wold-Zerlegung ist wichtig für die Prognose und darüber hinaus zentral für das Verständnis der Struktur von stationären Prozessen.

Das Prognoseproblem für skalare stationäre Prozesse wurde von Kolmogorov (siehe z. B. [39, Kapitel 1–3]) vollständig gelöst. Für den mehrdimensionalen Fall verweisen wir auf [39] und [17]<sup>1</sup>.

## 2.1 Prognose aus der endlichen Vergangenheit

Wir wollen nun die optimale lineare  $h$ -Schritt-Prognose aus der endlichen Vergangenheit konstruieren. Das entsprechende Optimierungsproblem

$$\begin{aligned} \mathbf{E}(x_{t+h} - \check{x})'(x_{t+h} - \check{x}) &\longrightarrow \min \\ \check{x} &= c_0 + c_1 x_t + \cdots + c_k x_{t+1-k} \end{aligned}$$

kann in  $n$ -unabhängige Teilprobleme zerlegt werden

$$\begin{aligned} \mathbf{E}(x_{i,t+h} - \check{x}_i)^2 &\longrightarrow \min \\ \check{x}_i &= c_{i0} + c_{i1} x_t + \cdots + c_{ik} x_{t+1-k}, \end{aligned}$$

wobei  $c_{i0} \in \mathbb{R}$  das  $i$ -te Element von  $c_0$  und  $c_{ij} \in \mathbb{R}^{1 \times n}$  die  $i$ -te Zeile von  $c_j \in \mathbb{R}^{n \times n}$  bezeichnet ( $i = 1, \dots, n$ ). Äquivalent dazu ist folgendes Problem im Hilbert-Raum der quadratisch integrierbaren Zufallsvariablen  $\mathbb{L}_2(\Omega, \mathcal{A}, \mathbf{P})$

$$\begin{aligned} \|x_{i,t+h} - \check{x}_i\| &\longrightarrow \min \\ \check{x}_i &\in \text{sp}\{1, x_t, \dots, x_{t+1-k}\} =: \mathbb{M} \subset \mathbb{L}_2. \end{aligned}$$

<sup>1</sup> Edward J. Hannan (1921–1994). Australischer Statistiker. Einer der Pioniere der modernen Zeitreihenanalyse.

Die 1 ist hier wieder als Zufallsvariable ( $\omega \mapsto 1$ ) zu interpretieren. Die Lösung folgt unmittelbar aus dem Projektionssatz 1.2

$$\hat{x}_{i,t,h,k} = P_{\mathbb{M}} x_{i,t+h}. \quad (2.1)$$

Die optimale Prognose für  $x_{i,t+h}$  ist also die Projektion von  $x_{i,t+h}$  auf den Unterraum  $\mathbb{M}$ . Verwenden wir die im Abschn. 1.3 eingeführte Konvention für Zufallsvektoren, dann können wir auch schreiben:

$$\hat{x}_{t,h,k} = P_{\mathbb{M}} x_{t+h}. \quad (2.2)$$

Wir zeigen nun, dass man sich im Wesentlichen auf zentrierte Prozesse (d. h. Erwartungswert  $\mathbf{E}x_t = 0$ ) und lineare Prognosen (d. h.  $c_0 = 0$ ) beschränken kann. Dazu definieren wir den mittelwertbereinigten Prozess ( $\tilde{x}_t = x_t - \mu$ ) mit  $\mu = \mathbf{E}x_t$ . Der Unterraum  $\mathbb{M} = \text{sp}\{1, x_t, \dots, x_{t+1-k}\}$  ist die direkte Summe aus den beiden *orthogonalen* Unterräumen  $\text{sp}\{1\}$  und  $\tilde{\mathbb{M}} := \text{sp}\{\tilde{x}_1, \dots, \tilde{x}_{t+1-k}\}$ , da  $\langle 1, \tilde{x}_{is} \rangle = \mathbf{E}\tilde{x}_{is} = 0$ . Die Projektion auf  $\mathbb{M}$  ist daher gleich der Summe der Projektionen auf  $\text{sp}\{1\}$  und der Projektion auf  $\tilde{\mathbb{M}}$ , d. h.  $P_{\mathbb{M}} = P_{\text{sp}\{1\}} + P_{\tilde{\mathbb{M}}}$ . Mit der Linearität der Projektionsoperatoren folgt nun weiter

$$\begin{aligned} P_{\mathbb{M}} x_{t+h} &= (P_{\text{sp}\{1\}} + P_{\tilde{\mathbb{M}}})(\tilde{x}_{t+h} + \mu) \\ &= P_{\text{sp}\{1\}} \tilde{x}_{t+h} + P_{\text{sp}\{1\}} \mu + P_{\tilde{\mathbb{M}}} \tilde{x}_{t+h} + P_{\tilde{\mathbb{M}}} \mu = \mu + P_{\tilde{\mathbb{M}}} \tilde{x}_{t+h}, \end{aligned}$$

da  $\tilde{x}_{i,t+h} \perp \text{sp}\{1\}$ ,  $\mu_i 1 \perp \tilde{\mathbb{M}}$  und  $\mu_i 1 \in \text{sp}\{1\}$ . Dies zeigt:

- (1) Die optimale Prognose des zentrierten Prozesses ist linear (d. h.  $c_0 = 0$ ):

$$\hat{\tilde{x}}_{t,h,k} = (P_{\text{sp}\{1\}} + P_{\tilde{\mathbb{M}}})\tilde{x}_{t+h} = P_{\tilde{\mathbb{M}}} \tilde{x}_{t+h} = c_1 \tilde{x}_t + \dots + c_k \tilde{x}_{t+1-k}.$$

- (2) Die Prognose für  $x_{t+h}$  erhält man einfach, indem man zur Prognose des zentrierten Prozesses noch den Erwartungswert  $\mu = \mathbf{E}x_{t+h}$  addiert:

$$\hat{x}_{t,h,k} = \mu + \hat{\tilde{x}}_{t,h,k} = (I_n - c_1 - \dots - c_k)\mu + c_1 x_t + \dots + c_k x_{t+1-k}.$$

Ähnliche Überlegungen gelten auch für die Prognose aus der unendlichen Vergangenheit. Wir werden daher im Folgenden o. B. d. A. annehmen, dass der betrachtete Prozess schon zentriert ist und daher auch nur lineare Prognosen ( $c_0 = 0$ ) betrachten. Entsprechend bezeichnet  $\mathbb{M}$  ab jetzt den Unterraum  $\mathbb{M} = \text{sp}\{x_t, \dots, x_{t+1-k}\}$ .

Der Projektionssatz liefert ein lineares Gleichungssystem, um die Prognosekoeffizienten zu bestimmen. Eine Linearkombination  $c_{i1}x_t + \dots + c_{ik}x_{t+1-k} \in \mathbb{M}$  ist dann und nur dann gleich der Projektion von  $x_{i,t+h}$  auf den Raum  $\mathbb{M}$  wenn der Fehler  $x_{i,t+h} - (c_{i1}x_t + \dots + c_{ik}x_{t+1-k})$  orthogonal auf diesen Raum ist, d. h. dann und nur dann, wenn

$$\langle x_{i,t+h} - c_{i1}x_t - \dots - c_{ik}x_{t+1-k}, x_{j,t+1-l} \rangle = 0 \text{ für } 1 \leq j \leq n \text{ und } 1 \leq l \leq k. \quad (2.3)$$

Die Gleichungen (2.3) können wir mit  $\hat{x}_{t,h,k} = (c_1, \dots, c_k)x_t^k$  zusammenfassen zu

$$\mathbf{E}[(x_{t+h} - (c_1, c_2, \dots, c_k)x_t^k)(x_t^k)'] = 0$$

und erhalten damit folgende „Prognosegleichungen“

$$(\gamma(h), \gamma(h+1), \dots, \gamma(h+k-1)) = (c_1, \dots, c_k)\Gamma_k \quad (2.4)$$

zur Bestimmung der Koeffizienten  $(c_1, \dots, c_k)$ .

Der (optimale) Prognosefehler  $u_{t,h,k} = x_{t+h} - \hat{x}_{t,h,k} = x_{t+h} - (c_1, \dots, c_k)x_t^k$  hat Erwartungswert gleich null. Die Varianz  $\Sigma_{h,k}$  des Fehlers kann folgendermaßen bestimmt werden. Da  $\hat{x}_{j,t,h,k} \in \mathbb{M}$  folgt  $\langle u_{i,t,h,k}, \hat{x}_{j,t,h,k} \rangle = \mathbf{Cov}(u_{i,t,h,k}, \hat{x}_{j,t,h,k}) = 0$  und daher  $\mathbf{Var}(x_{t+h}) = \mathbf{Var}(\hat{x}_{t,h,k} + u_{t,h,k}) = \mathbf{Var}(\hat{x}_{t,h,k}) + \mathbf{Var}(u_{t,h,k})$ . Somit haben wir

$$\begin{aligned} \Sigma_{h,k} &:= \mathbf{E}u_{t,h,k}u_{t,h,k}' = \mathbf{Var}(x_{t+h}) - \mathbf{Var}(\hat{x}_{t,h,k}) \\ &= \gamma(0) - (c_1, \dots, c_k)\Gamma_k(c_1, \dots, c_k)'. \end{aligned} \quad (2.5)$$

Der mittlere, quadratische Fehler (MSE) der optimalen Prognose ist gleich

$$\mathbf{E}[u_{t,h,k}'u_{t,h,k}] = \text{tr}(\Sigma_{h,k}).$$

Aus dem Projektionssatz können wir folgende Schlüsse ziehen: Die (optimale) Prognose  $\hat{x}_{t,h,k}$  und damit auch der entsprechende Prognosefehler  $u_{t,h,k}$  (und dessen Varianz  $\Sigma_{h,k}$ ) sind (fast sicher) eindeutig. Die Prognosegleichungen (2.4) sind immer lösbar, auch wenn die Matrix  $\Gamma_k$  singulär ist.

Wenn  $\Gamma_k > 0$  positiv definit ist, dann hat (2.4) eine eindeutige Lösung

$$(c_1, \dots, c_k) = (\gamma(h), \dots, \gamma(h+k-1))\Gamma_k^{-1} \quad (2.6)$$

und die Prognosefehlervarianz lässt sich berechnen mit

$$\Sigma_{h,k} = \gamma(0) - (\gamma(h), \dots, \gamma(h+k-1))\Gamma_k^{-1}(\gamma(h), \dots, \gamma(h+k-1))'. \quad (2.7)$$

Falls  $\Gamma_k$  singulär ist, dann existieren unendlich viele Lösungen. Die Zufallsvariablen  $\{x_{1t}, \dots, x_{nt}, x_{1,t-1}, \dots, x_{n,t+1-k}\}$  sind in diesem Fall linear abhängig und bilden daher keine Basis für  $\mathbb{M}$ , vergleiche auch Aufgabe „Projektion“ in Abschn. 1.3.

Aufgrund der Linearität des Projektors können wir auch sofort die optimale Prognose für beliebige Linearkombinationen  $cx_{t+h}$ ,  $c \in \mathbb{R}^{1 \times n}$  angeben:

$$\mathbf{P}_{\mathbb{M}}(cx_{t+h}) = c(\mathbf{P}_{\mathbb{M}} x_{t+h}) = c\hat{x}_{t,h}.$$

Daher folgt auch

$$\Sigma_{h,k} \leq \mathbf{E}(x_{t+h} - \tilde{x})(x_{t+h} - \tilde{x})' \quad (2.8)$$

für alle  $\tilde{x}$  der Form  $\tilde{x} = \tilde{c}_0 + \tilde{c}_1 x_t + \dots + \tilde{c}_k x_{t+1-k}$ ,  $\tilde{c}_0 \in \mathbb{R}^{n \times 1}$ ,  $\tilde{c}_i \in \mathbb{R}^{n \times n}$   $i = 1, \dots, k$ . Die Prognose ist also auch optimal bezüglich der Halb-Ordnung „ $\geq$ “. Eine weitere Folgerung ist

$$\Sigma_{h,k} \leq \Sigma_{h,k-1}, \quad (2.9)$$

d. h. die Prognose wird i. Allg. besser (zumindest kann sie nicht schlechter werden) je mehr Information zur Prognose zur Verfügung steht. (Diese Ungleichung folgt aus (2.8), wenn man  $\tilde{x} = \hat{x}_{t,h,k-1}$  setzt.)

Wenn  $\Sigma_{h,k}$  singulär ist, dann gibt es Linearkombinationen  $c x_{t+h}$ , die perfekt (d. h. ohne Fehler) prognostiziert werden. In diesem Fall ist die (Block-)Toeplitz-Matrix  $\Gamma_{k+h}$  auch singulär.

Im Folgenden werden wir diese Prognose(n) als lineare Kleinst-Quadrate- (KQ)-Prognose(n) bezeichnen. Analoge Überlegungen gelten auch für nicht-stationäre (aber quadratisch integrierbare) Prozesse. Der einzige Unterschied ist, dass die Varianz-Kovarianz-Matrix  $\text{Var}(x_t^k)$  keine (Block-)Toeplitz-Struktur mehr haben muss und dass die Prognosekoeffizienten und die Varianz der Prognosefehler im Allgemeinen auch von  $t$  abhängen.

### Aufgabe

Sei  $(x_t)$  ein zentrierter, stationärer Prozess. Wir betrachten nun die Prognose für  $x_{t+1}$  aus den Werten  $x_1, \dots, x_t$  für  $t \in \mathbb{N}_0$ . Die entsprechenden Prognosen und Prognosefehler bezeichnen wir hier mit  $\hat{x}_{t+1|t}$  und  $u_{t+1|t}$ . Für  $t = 0$  setzen wir  $x_{1|0} = 0$  und  $u_{1|0} = x_1$ . Zeigen Sie nun

- (1)  $\dim(\text{sp}\{x_t, \dots, x_1\}) = \text{rg } \Gamma_t$ .
- (2)  $\text{sp}\{x_t, \dots, x_1\} = \text{sp}\{u_{t|t-1}\} \oplus \text{sp}\{x_{t-1}, \dots, x_1\} = \text{sp}\{u_{t|t-1}\} \oplus \text{sp}\{u_{t-1|t-2}\} \oplus \dots \oplus \text{sp}\{u_{1|0}\}$ , wobei die Teilräume zueinander orthogonal sind, d. h. z. B.  $\text{sp}\{u_{t|t-1}\} \perp \text{sp}\{x_{t-1}, \dots, x_1\}$ .
- (3)

$$\text{rg}(\Gamma_{t+1}) = \text{rg}(\Gamma_t) + \text{rg}(\Sigma_{1,t}) \quad (2.10)$$

$$(\Gamma_{t+1} > 0) \iff ((\Gamma_t > 0) \text{ und } (\Sigma_{1,t} > 0)) \quad (2.11)$$

$$(\det(\Gamma_t) = 0) \implies (\det(\Sigma_{1,t-1}) = 0) \implies (\det(\Sigma_{1,t}) = 0). \quad (2.12)$$

### Aufgabe (Fortsetzung der obigen Aufgabe)

Wir nehmen jetzt zusätzlich an, dass der Prozess skalar ist ( $n = 1$ ) und dass die Toeplitz-Matrix  $\Gamma_k > 0$  regulär ist. Betrachten Sie nun die *Cholesky-Zerlegung* der Toeplitz-Matrix  $\Gamma_k$

$$\Gamma_k = DSD',$$

wobei  $D = (d_{ij})_{i,j=1,\dots,k} \in \mathbb{R}^{k \times k}$  eine obere Dreiecksmatrix ( $d_{ii} = 1$  und  $d_{ij} = 0$  für  $i > j$ ) und  $S \in \mathbb{R}^{k \times k}$  eine Diagonalmatrix ist. Die Inverse von  $D$  bezeichnen wir mit  $C = D^{-1} = (c_{ij})_{i,j=1,\dots,k}$ . Es gilt  $c_{ii} = 1$  und  $c_{ij} = 0$  für  $i > j$ . Zeigen Sie:

$$(u_{k|k-1}, u_{k-1|k-2}, \dots, u_{1|0})' = C(x_k, x_{k-1}, \dots, x_1)'$$

und  $S = \text{diag}(\sigma_{1,k-1}^2, \sigma_{1,k-2}^2, \dots, \sigma_{1,0}^2)$ , wobei  $\sigma_{t,l-1}^2 = \mathbf{E}u_{t|t-1}^2$ . Daher folgt auch für  $1 \leq l < k$

$$\hat{x}_{t,1,l} = -(c_{k-l,k-l+1}x_t + c_{k-l,k-l+2}x_{t-1} + \dots + c_{k-l,k}x_{t+1-l}).$$

**Aufgabe**

Gegeben sei der Prozess  $(x_t = \cos(\lambda t) \mid t \in \mathbb{N})$ , wobei  $\lambda$  eine auf  $[-\pi, \pi]$  gleichverteilte Zufallsvariable ist. In der Aufgabe am Ende von Abschn. 1.2 sollte man zeigen, dass  $\mathbf{E}x_t = 0$  und  $\gamma(k) = \mathbf{E}x_{t+k}x_t = 0$  für  $t+k > t \geq 0$  gilt. Die optimale, *lineare* Prognose ist also gleich null, d. h.  $\hat{x}_{t,h,k} = 0$  für  $1 \leq k < t$ . Dieser Prozess erlaubt aber eine perfekte, *nichtlineare* Prognose. Zeigen Sie:

$$\begin{aligned} x_{t+1} &= 2x_t x_1 - x_{t-1} \\ x_t &= 2x_{t-1} x_1 - x_{t-2} \\ \text{und damit } x_{t+1} &= \frac{x_t^2 - x_{t-1}^2 + x_t x_{t-2}}{x_{t-1}}. \end{aligned}$$

**Aufgabe**

Wir betrachten den skalaren AR(1)-Prozess  $x_t = ax_{t-1} + \epsilon_t$ , mit  $|a| < 1$  und  $(\epsilon_t) \sim \text{WN}(\sigma^2)$ , siehe auch (1.16) und (1.17). Zeigen Sie mithilfe der Gleichungen (2.4) und (2.5), dass für  $k \geq 1$

$$\hat{x}_{t,h,k} = a^h x_t \quad \text{und} \quad \sigma_{h,k}^2 = \sigma^2 \frac{(1 - a^{2h})}{(1 - a^2)}.$$

**Aufgabe**

Betrachten Sie den MA(1)-Prozess  $x_t = \epsilon_t - \epsilon_{t-1}$ , wobei  $(\epsilon_t) \sim \text{WN}(\sigma^2)$  weißes Rauschen mit Varianz  $\mathbf{E}\epsilon_t^2 = \sigma^2$  ist. Beweisen Sie folgende Formeln für die Einschrittprognose  $\hat{x}_{t,1,k}$  aus  $k$  vergangenen Werten und den entsprechenden Prognosefehler  $u_{t,1,k} = x_{t+1} - \hat{x}_{t,1,k}$ :

$$\begin{aligned} \hat{x}_{t,1,k} &= \frac{-1}{k+1} (kx_t + (k-1)x_{t-1} + \cdots + 2x_{t+2-k} + 1x_{t+1-k}) \\ u_{t,1,k} &= \frac{1}{k+1} ((k+1)\epsilon_{t+1} - \epsilon_t - \epsilon_{t-1} - \cdots - \epsilon_{t+1-k} - \epsilon_{t-k}) \\ \sigma_{1,k}^2 &= \mathbf{E}(u_{t,1,k}^2) = \sigma^2 + \frac{\sigma^2}{k+1} = \sigma^2 \frac{k+2}{k+1}. \end{aligned}$$

Zeigen Sie auch, dass der Einschrittprognosefehler für  $k \rightarrow \infty$  gegen  $\epsilon_{t+1}$  konvergiert, d. h.

$$\lim_{k \rightarrow \infty} u_{t,1,k} = \epsilon_{t+1}.$$

Wie schon oben erwähnt kann man mit der selben Strategie auch die Kleinst-Quadrat-Prognose für nicht stationäre (aber quadratisch integrierbare) Prozesse bestimmen.

**Aufgabe**

Sei  $(x_t)$  ein Prozess der Form  $x_t = \mu_t + y_t$ , wobei  $\mu_t$  eine deterministische Funktion der Zeit und  $(y_t)$  ein stationärer, zentrierter Prozess ist. Überzeugen Sie sich, dass

$$\begin{aligned} \hat{x}_{t,h,k} &= \mu_{t+h} + \hat{y}_{t,h,k} = \mu_{t+h} + c_1 y_t + \cdots + c_k y_{t+1-k} \\ &= (\mu_{t+h} - c_1 \mu_t - \cdots - c_k \mu_{t+1-k}) + c_1 x_t + \cdots + c_k x_{t+1-k} \end{aligned}$$

die beste affine Prognose für  $x_{t+h}$  aus  $k$  vergangenen Werten ist. Hier bezeichnet  $\hat{y}_{t,h,k} = c_1 y_t + \cdots + c_k y_{t+1-k}$  die  $h$ -Schrittprognose für  $y_{t+h}$ . Für den Prognosefehler gilt  $x_{t+h} - \hat{x}_{t,h,k} = y_{t+h} - \hat{y}_{t,h,k}$ .

**Aufgabe**

Sei  $(y_t)$  ein zentrierter stationärer Prozess und  $(x_t \mid t \in \mathbb{N}_0)$  der durch  $x_t = x_0 + \sum_{j=1}^t y_j$  definierte integrierte Prozess. Wir nehmen an, dass der Startwert  $x_0$  quadratisch integrierbar und unkorreliert zu  $y_s$ ,  $s \geq 1$  ist. Zeigen Sie, dass

$$\hat{x}_{t+h} = x_t + \hat{y}_{t,1,t} + \hat{y}_{t,2,t} + \cdots + \hat{y}_{t,h,t}$$

die beste Prognose für  $x_{t+h}$  aus den Werten  $x_0, \dots, x_t$  ist.

Ist  $(y_t) \sim \text{WN}(\sigma^2)$  weißes Rauschen (d.h.  $(x_t)$  eine Irrfahrt), dann ist die naive Prognose  $\hat{x}_{t+h} = x_t$  die optimale.

**2.2 Prognose aus der unendlichen Vergangenheit**

In diesem Abschnitt betrachten wir den Grenzwert der Prognosen  $\hat{x}_{t,h,k}$  für  $k \rightarrow \infty$ , d.h. wir verwenden die gesamte Informationen aus der Vergangenheit für die Prognose. Die Prognose aus der unendlichen Vergangenheit zeigt gewisse strukturelle Eigenschaften des zugrunde liegenden Prozesses auf, wie im nächsten Abschnitt gezeigt wird. Wir betrachten wieder nur den Fall von zentrierten Prozessen, d.h.  $\mathbf{E}x_t = 0$ . Alle Resultate lassen sich aber ohne weiteres auf den allgemeinen Fall übertragen. Es gilt folgender Satz.

**Satz 2.1** Die Folge  $(\hat{x}_{t,h,k} \mid k \in \mathbb{N})$  konvergiert (im quadratischen Mittel) und der Grenzwert ist die Projektion von  $x_{t+h}$  auf den Raum  $\mathbb{H}_t(x) = \overline{\text{sp}}\{x_{is} \mid 1 \leq i \leq n, s \leq t\} = \overline{\text{sp}}\{x_s \mid s \leq t\}$ , also

$$\lim_{k \rightarrow \infty} \hat{x}_{t,h,k} = \mathbf{P}_{\mathbb{H}_t(x)} x_{t+h} =: \hat{x}_{t,h}.$$

Die Varianz des entsprechenden Fehlers  $u_{t,h} = x_{t+h} - \hat{x}_{t,h}$  ist

$$\Sigma_h := \text{Var}(u_{t,h}) = \mathbf{E}u_{t,h}u_{t,h}' = \lim_{k \rightarrow \infty} \Sigma_{h,k}.$$

*Beweis* Sei  $\mathbb{H}_{t,k}(x) = \text{sp}\{x_s \mid t+1-k \leq s \leq t\}$ . Aus  $\mathbb{H}_{t,k}(x) \subset \mathbb{H}_t(x)$  folgt dann  $\mathbf{P}_{\mathbb{H}_{t,k}(x)} = \mathbf{P}_{\mathbb{H}_{t,k}(x)} \mathbf{P}_{\mathbb{H}_t(x)}$  und somit  $\hat{x}_{t,h,k} = \mathbf{P}_{\mathbb{H}_{t,k}(x)} \hat{x}_{t,h}$  wobei  $\hat{x}_{t,h} = \mathbf{P}_{\mathbb{H}_t(x)} x_{t+h}$ . Andererseits gilt  $\hat{x}_{t,h} = \text{l.i.m.}_k x^{(k)}$  für eine geeignet gewählte Folge von Zufallsvektoren  $x^{(k)} \in (\mathbb{H}_{t,k}(x))^n$ , da  $\hat{x}_{t,h}$  ein Grenzwert von endlichen Summen ist. Aus den Eigenschaften der Projektion folgt schließlich

$$\mathbf{E}(\hat{x}_{t,h} - \hat{x}_{t,h,k})(\hat{x}_{t,h} - \hat{x}_{t,h,k})' \leq \mathbf{E}(\hat{x}_{t,h} - x^{(k)})(\hat{x}_{t,h} - x^{(k)})'$$

und somit die Konvergenz von  $\hat{x}_{t,h,k}$  gegen  $\hat{x}_{t,h}$  für  $k \rightarrow \infty$ . □

Mit der Ein-Schritt-Prognose aus der unendlichen Vergangenheit erhalten wir eine Zerlegung des Prozesses der Form

$$x_{t+1} = \hat{x}_{t,1} + u_{t,1},$$

wobei  $\hat{x}_{t,1}$  der Teil von  $x_{t+1}$  ist, der aus der Vergangenheit bestimmt ist, und  $u_{t,1}$  ist der „nicht vorhersehbare“ Anteil. Daher nennt man die Ein-Schritt-Prognosefehler aus der unendlichen Vergangenheit die *Innovationen* des Prozesses.

**Satz 2.2** Die Innovationen  $(u_t = u_{t-1,1} \mid t \in \mathbb{Z})$  eines stationären Prozesses sind weißes Rauschen.

*Beweis* Klarerweise sind die Innovationen (schwach) stationär und der Erwartungswert ist gleich null ( $\mathbf{E}u_{t,1} = 0$ ). Es gilt  $u_{t-1,1} \in \mathbb{H}_t(x)$  und  $u_{t-1,1} \perp \mathbb{H}_{t-1}(x)$ . Daher gilt  $u_{t-1,1} \perp u_{s-1,1} \in \mathbb{H}_s(x) \subset \mathbb{H}_{t-1}(x)$  für alle  $s \leq t-1$  und damit also auch  $u_{t,1} \perp u_{s,1}$  für alle  $s \neq t$ .  $\square$

Die Ungleichung

$$\Sigma_{h+1} \geq \Sigma_h \quad (2.13)$$

folgt unmittelbar aus  $\mathbb{H}_{t-1}(x) \subset \mathbb{H}_t(x)$  und  $\hat{x}_{t-1,h+1} = \mathbf{P}_{\mathbb{H}_{t-1}(x)} x_{t+h}$  und  $\hat{x}_{t,h} = \mathbf{P}_{\mathbb{H}_t(x)} x_{t+h}$ .

Das hier diskutierte Prognoseverfahren basiert auf der Kovarianzfunktion des Prozesses. Hat man ein parametrisches Modell für den Prozess (wie z. B. ein AR-Modell oder ein Zustandsraummodell) zur Verfügung, dann kann die Prognose erheblich vereinfacht werden. Wir werden das in den entsprechenden Kapiteln noch diskutieren.

## 2.3 Reguläre und singuläre Prozesse und die Wold-Zerlegung

Die Wold-Zerlegung eines stationären Prozesses teilt den Prozess in einen „deterministischen“ und einen „regulären“ Anteil. Hier bedeutet „deterministisch“, dass die Zukunft komplett durch die Vergangenheit bestimmt ist, während „regulär“ bedeutet, dass die unendlich ferne Vergangenheit für die Zukunft keine Rolle spielt. Ist die Wold-Zerlegung eines Prozesses bekannt, so erhält man eine einfache und explizite Darstellung der Prognose des regulären Teils (für beliebige  $h > 0$ ).

### Definition

Ein stationärer Prozess  $(x_t)$  ist

- *regulär* („regular“, „purely non-deterministic“), wenn  $\text{l.i.m.}_{h \rightarrow \infty} \hat{x}_{t,h} = 0$  (und daher  $\lim_{h \rightarrow \infty} \Sigma_h = \mathbf{E}x_t x_t'$ ) gilt. Der Erwartungswert eines regulären Prozesses muss null sein ( $\mathbf{E}x_t = 0$ ).
- *singulär* („singular“, „deterministic“), wenn  $\Sigma_h = 0$  für ein  $h > 0$  (und daher auch für alle  $h > 0$ ) gilt. Einen singulären Prozess nennt man auch *deterministisch*, weil die Zukunft aus der Vergangenheit bestimmt ist.



**Aufgabe**

Zeigen Sie, dass  $\Sigma_{\bar{h}} = 0$  für ein  $\bar{h} > 0$  auch  $\Sigma_h = 0$  für alle  $h > 0$  impliziert.

**Aufgabe**

Zeigen Sie, dass harmonische Prozesse (siehe Gleichung (1.19)) singulär sind. Hinweis: Beweisen Sie  $\hat{x}_{t,1,k} = x_{t+1}$  für  $k \geq K$ .

**Beispiele**

(1) MA( $q$ ) Prozesse sind regulär, da  $\hat{x}_{t,h} = \lim_{k \rightarrow \infty} \hat{x}_{t,h,k} = 0$  für  $h > q$ .

(2) Prozesse mit einer kausalen MA( $\infty$ )-Darstellung sind regulär:

Sei  $x_t = \sum_{k \geq 0} b_k \epsilon_{t-k}$  eine kausale MA( $\infty$ )-Darstellung für den Prozess  $(x_t)$ . Da  $x_s \in \mathbb{H}_t(\epsilon)$   $\forall s \leq t$  folgt  $\mathbb{H}_t(x) \subset \mathbb{H}_t(\epsilon)$ . Wir betrachten nun die Projektion  $\tilde{x}_{t,h} = P_{\mathbb{H}_t(\epsilon)} x_{t+h}$  von  $x_{t+h}$  auf  $\mathbb{H}_t(\epsilon)$ . Da  $(\epsilon_t)$  weißes Rauschen ist, ist diese Projektion sehr einfach zu berechnen:

$$\tilde{x}_{t,h} = P_{\mathbb{H}_t(\epsilon)} x_{t+h} = \sum_{k \geq 0} b_k P_{\mathbb{H}_t(\epsilon)} \epsilon_{t+h-k} = \sum_{k \geq h} b_k \epsilon_{t+h-k}.$$

Für  $h \rightarrow \infty$  konvergiert  $\tilde{x}_{t,h}$  gegen null, d. h. l.i.m. $_{h \rightarrow \infty} \tilde{x}_{t,h} = 0$ . Da  $\mathbb{H}_t(x)$  ein Teil-Hilbert-Raum von  $\mathbb{H}_t(\epsilon)$  ist, folgt auch  $\mathbf{E}(x_{t+h} - \hat{x}_{t,h})(x_{t+h} - \hat{x}_{t,h})' \geq \mathbf{E}(x_{t+h} - \tilde{x}_{t,h})(x_{t+h} - \tilde{x}_{t,h})'$  bzw.  $\mathbf{E}(\hat{x}_{t,h})(\hat{x}_{t,h})' \leq \mathbf{E}(\tilde{x}_{t,h})(\tilde{x}_{t,h})'$ . Zusammen ergibt das l.i.m. $_{h \rightarrow \infty} \hat{x}_{t,h} = 0$  wie behauptet. Das folgende Theorem (Wold-Zerlegung) wird zeigen, dass umgekehrt jeder reguläre Prozess eine kausale MA( $\infty$ )-Darstellung besitzt.

**Satz 2.3 (Wold-Zerlegung)**

(1) Jeder stationäre Prozess  $(x_t)$  besitzt eine eindeutige Zerlegung  $x_t = y_t + z_t$  mit folgenden Eigenschaften:

(a)  $(y_t)$  ist regulär und  $(z_t)$  ist singulär.

(b) Die Prozesse  $(y_t)$  und  $(z_t)$  sind zueinander orthogonal ( $\mathbf{E}y_t z_s' = 0$  für alle  $t, s \in \mathbb{Z}$ ).

(c)  $y_t \in \text{sp}\{1\} + \mathbb{H}_x(t)$  und  $z_t \in \text{sp}\{1\} + \mathbb{H}_x(t)$ .

(2) Der reguläre Prozess  $(y_t)$  besitzt eine kausale MA( $\infty$ )-Darstellung

$$y_t = \sum_{j \geq 0} b_j \epsilon_{t-j}, \text{ mit } b_0 = I \text{ und } \sum_{j \geq 0} \|b_j\|^2 < \infty, \quad (2.14)$$

wobei  $(\epsilon_t)$  ein weißes Rauschen ist. Es gilt  $\mathbb{H}_\epsilon(t) = \mathbb{H}_y(t)$  und die  $\epsilon_t$ s sind sowohl die Innovationen von  $(x_t)$  als auch von  $(y_t)$ .

**Beweis** Um den Beweis etwas zu vereinfachen, nehmen wir an, dass  $\mathbf{E}x_t = 0$  gilt. Für den allgemeinen Fall betrachtet man einfach die Wold-Zerlegung  $\tilde{x}_t = \tilde{y} + \tilde{z}_t$  des zentrierten Prozesses  $\tilde{x}_t = x_t - \mathbf{E}x_t$  und setzt  $y_t = \tilde{y}_t$  und  $z_t = \tilde{z}_t + \mathbf{E}x_t$ .

Zunächst definieren wir die Innovationen

$$\epsilon_t = x_t - P_{\mathbb{H}_{t-1}(x)} x_t \quad (2.15)$$

des Prozesses  $(x_t)$  und merken an, dass

$$\begin{aligned}\epsilon_t &\in (\mathbb{H}_t(x))^n \\ \mathbb{H}_t(\epsilon) &\subset \mathbb{H}_t(x) \\ \epsilon_t &\perp \mathbb{H}_{t-1}(x).\end{aligned}\tag{2.16}$$

Die Prozesse  $(y_t)$  und  $(z_t)$  definieren wir durch

$$y_t = P_{\mathbb{H}_t(\epsilon)} x_t \tag{2.17}$$

$$z_t = x_t - y_t. \tag{2.18}$$

Der Prozess  $(\epsilon_t)$  ist weißes Rauschen (siehe Satz 2.2) und daher besitzt  $y_t \in \mathbb{H}_t(\epsilon)$  eine kausale MA( $\infty$ )-Darstellung

$$y_t = \sum_{j \geq 0} b_j \epsilon_{t-j}$$

mit quadratisch summierbaren Koeffizienten  $(b_j)_{j \geq 0}$ . Wegen

$$b_0 \epsilon_t = P_{\text{sp}\{\epsilon_t\}} x_t = P_{\text{sp}\{\epsilon_t\}} \epsilon_t + P_{\text{sp}\{\epsilon_t\}} P_{\mathbb{H}_{t-1}(x)} x_t = \epsilon_t$$

können wir o. B. d. A.  $b_0 = I$  setzen. Wir sehen auch, dass

$$\begin{aligned}y_t &\in (\mathbb{H}_t(\epsilon))^n \\ \mathbb{H}_t(y) &\subset \mathbb{H}_t(\epsilon) \subset \mathbb{H}_t(x) \\ z_t &\in (\mathbb{H}_t(x))^n \\ \mathbb{H}_t(z) &\subset \mathbb{H}_t(x) \\ z_t &\perp \mathbb{H}_t(\epsilon).\end{aligned}$$

Die Prozesse  $(z_t)$  und  $(\epsilon_t)$  sind orthogonal zueinander, d. h.

$$z_t \perp \epsilon_s \text{ für alle } s, t \in \mathbb{Z}.$$

Für  $t \geq s$  folgt diese Behauptung aus  $z_t \perp \mathbb{H}_t(\epsilon)$ ,  $\epsilon_s \in (\mathbb{H}_t(\epsilon))^n$  und für  $s > t$  aus  $\epsilon_s \perp \mathbb{H}_t(x)$ ,  $z_t \in (\mathbb{H}_t(x))^n$ . Da  $\mathbb{H}_t(y) \subset \mathbb{H}_t(\epsilon)$  sind auch die Prozesse  $(y_t)$  und  $(z_t)$  orthogonal zueinander. Es gilt  $x_t = y_t + z_t$  und daher  $\mathbb{H}_t(x) \subset \mathbb{H}_t(y) \oplus \mathbb{H}_t(z)$ . Andererseits haben wir aber auch  $\mathbb{H}_t(y) \oplus \mathbb{H}_t(z) \subset \mathbb{H}_t(x)$  wegen  $\mathbb{H}_t(y) \subset \mathbb{H}_t(x)$  und  $\mathbb{H}_t(z) \subset \mathbb{H}_t(x)$ . Der Unterraum  $\mathbb{H}_t(x)$  ist also die Summe von zwei orthogonalen Unterräumen

$$\mathbb{H}_t(x) = \mathbb{H}_t(y) \oplus \mathbb{H}_t(z).$$

Wegen  $\mathbb{H}_t(\epsilon) \subset \mathbb{H}_t(x) = \mathbb{H}_t(y) \oplus \mathbb{H}_t(z)$ ,  $\mathbb{H}_t(\epsilon) \perp \mathbb{H}_t(z)$  und  $\mathbb{H}_t(y) \subset \mathbb{H}_t(\epsilon)$  folgt auch

$$\mathbb{H}_t(\epsilon) = \mathbb{H}_t(y).$$

Wir betrachten nun die Prognose von  $y_{t+h}$  aus der eigenen, unendlichen Vergangenheit. Dazu zerlegen wir  $y_{t+h}$  in

$$y_{t+h} = \underbrace{\sum_{j=0}^{h-1} b_j \epsilon_{t+h-j}}_{\perp \mathbb{H}_t(\epsilon) = \mathbb{H}_t(y)} + \underbrace{\sum_{j \geq h} b_j \epsilon_{t+h-j}}_{\in \mathbb{H}_t(\epsilon) = \mathbb{H}_t(y)}.$$

Der zweite Teil der rechten Seite ist in  $\mathbb{H}_t(y) = \mathbb{H}_t(\epsilon)$  enthalten und der erste ist orthogonal auf diesen Raum. Daher ist die Prognose gleich

$$\hat{y}_{t,h} = P_{\mathbb{H}_t(y)} y_{t+h} = \sum_{j \geq h} b_j \epsilon_{t+h-j} \quad (2.19)$$

und wir sehen, dass der Prozess  $(y_t)$  regulär ist, da

$$\text{l.i.m}_{h \rightarrow \infty} \hat{y}_{t,h} = \text{l.i.m}_{h \rightarrow \infty} \sum_{j \geq h} b_j \epsilon_{t+h-j} = 0.$$

Der Fehler der Ein-Schritt-Prognose für  $y_{t+1}$  ist  $b_0 \epsilon_{t+1} = \epsilon_{t+1}$ . Das heißt, die  $\epsilon_t$ 's sind auch die Innovationen von  $(y_t)$ .

Aufgrund der Orthogonalitätsbeziehung  $\epsilon_t \perp \mathbb{H}_{t-1}(x)$ , siehe (2.16) und (2.15), können wir den Raum  $\mathbb{H}_t(x)$  auch folgendermaßen in eine Summe von orthogonalen Räumen zerlegen

$$\begin{aligned} \mathbb{H}_t(x) &= \text{sp}\{\epsilon_t\} \oplus \mathbb{H}_{t-1}(x) \\ &= \text{sp}\{\epsilon_t\} \oplus \text{sp}\{\epsilon_{t-1}\} \oplus \mathbb{H}_{t-2}(x) \\ &\vdots \\ &= \text{sp}\{\epsilon_t\} \oplus \cdots \oplus \text{sp}\{\epsilon_{t+1-k}\} \oplus \mathbb{H}_{t-k}(x). \end{aligned}$$

Der Zufallsvektor  $z_t$  ist orthogonal auf  $\mathbb{H}_t(\epsilon)$ , d. h. auf alle  $\epsilon_s$ ,  $s \leq t$ . Daher folgt aus der obigen Zerlegung von  $\mathbb{H}_t(x)$  (zusammen mit  $z_t \in \mathbb{H}_t(x)$ ), dass

$$z_t \in \mathbb{H}_s(x) \text{ für alle } s \leq t.$$

Insbesondere gilt  $z_{t+1} \in \mathbb{H}_t(x) = \mathbb{H}_t(y) \oplus \mathbb{H}_t(z)$  und wegen  $z_{t+1} \perp \mathbb{H}_t(y)$  auch  $z_{t+1} \in \mathbb{H}_t(z)$ . Das bedeutet aber

$$\hat{z}_{t,1} = P_{\mathbb{H}_t(z)} z_{t+1} = z_{t+1}$$

und wir haben somit gezeigt, dass  $(z_t)$  ein singulärer Prozess ist.

Es bleibt nur noch die Eindeutigkeit dieser Wold-Zerlegung zu beweisen. Sei also  $x_t = y_t + z_t$  eine (beliebige) Zerlegung, die die Bedingungen (a)–(c) erfüllt. Damit folgt  $\mathbb{H}_t(x) = \mathbb{H}_t(y) \oplus \mathbb{H}_t(z)$ ,  $\mathbb{H}_t(y) \perp \mathbb{H}_t(z)$  und daher

$$\begin{aligned} \lim_{s \rightarrow -\infty} \text{l.i.m. } P_{\mathbb{H}_s(x)} x_t &= \text{l.i.m.}(P_{\mathbb{H}_s(y)} + P_{\mathbb{H}_s(z)})(y_t + z_t) \\ &= \underbrace{\text{l.i.m. } P_{\mathbb{H}_s(y)} y_t}_{=0} + \underbrace{\text{l.i.m. } P_{\mathbb{H}_s(z)} z_t}_{=z_t} \\ &= z_t. \end{aligned}$$

Das heißt  $z_t$  (und damit natürlich auch  $y_t$ ) ist eindeutig.  $\square$

Der Beweis zeigt auch, dass die  $h$ -Schritt-Prognose für  $x_{t+h}$  (aus der unendlichen Vergangenheit) gegeben ist durch

$$\hat{x}_{t,h} = \hat{y}_{t,h} + z_{t+h}.$$

Folgendes Korollar ist eine unmittelbare Folgerung des obigen Satzes bzw. dessen Beweises.

**Folgerung 2.4** *Ein stationärer Prozess  $(x_t)$  ist dann und nur dann regulär, wenn er eine kausale  $MA(\infty)$ -Darstellung  $x_t = \sum_{j \geq 0} b_j \epsilon_{t-j}$  mit  $b_0 = I$  und  $\mathbb{H}_t(x) = \mathbb{H}_t(\epsilon)$  besitzt. Die  $\epsilon_t$ 's sind die Innovationen des Prozesses  $(x_t)$  und für die  $h$ -Schritt-Prognose gilt*

$$\begin{aligned} \hat{x}_{t,h} &= \sum_{j \geq h} b_j \epsilon_{t+h-j} \\ u_{t,h} &= \sum_{j=0}^{h-1} b_j \epsilon_{t+h-j} \\ \Sigma_h &= \sum_{j=0}^{h-1} b_j \Sigma b_j'. \end{aligned}$$

Eine Konsequenz von Satz 2.3 ist, dass man die Prognose für  $x_{t+h}$  (aus der unendlichen Vergangenheit) dadurch erhält, dass man den regulären und den singulären Teil getrennt aus ihrer jeweiligen Vergangenheit prognostiziert und die Prognosen dann addiert. Ist  $(z_t)$  ein harmonischer Prozess, so kann die Prognose auf Basis der Formel (1.19) erfolgen. Die Extraktion des harmonischen Teils kann durch Regression auf harmonische Funktionen erfolgen. Zur Prognose des regulären Teiles benötigt man zunächst die Koeffizienten  $b_j$  in der Wold-Zerlegung. Wie man aus den zweiten Momenten von  $(y_t)$  die

Wold-Darstellung (2.14) und damit den Prädiktor erhält, ist relativ allgemein in [39, Kapitel 2] beschrieben. Die praktisch wichtigsten Fälle, AR-Prozesse bzw. ARMA-Prozesse und Prozesse, die mit Zustandsraumsystemen modelliert werden, werden in den entsprechenden Kap. 5 bzw. 6 und 7 diskutiert.

Die beste (i. Allg. nicht lineare) Kleinst-Quadrate-Prognose ist der bedingte Erwartungswert  $\mathbf{E}[x_{t+h} | x_t, x_{t-1}, \dots]$ . Wir betrachten nun einen regulären Prozess  $(x_t)$  wie in der obigen Folgerung, nehmen aber zusätzlich an, dass die Innovationen  $(\epsilon_t)$  eine *Martingal-Differenzenfolge* sind, d. h. dass  $\mathbf{E}[\epsilon_{t+h} | \epsilon_t, \epsilon_{t-1}, \dots] = 0$  für alle  $t$  und  $h > 0$  gilt. Die von  $\{x_s | s \leq t\}$  und  $\{\epsilon_s | s \leq t\}$  erzeugten  $\sigma$ -Algebren sind wegen  $\mathbb{H}_t(x) = \mathbb{H}_t(\epsilon)$  gleich und daher folgt

$$\begin{aligned} \mathbf{E}[x_{t+h} | x_t, x_{t-1}, \dots] &= \hat{x}_{t,h} + \mathbf{E}[u_{t,h} | x_t, x_{t-1}, \dots] \\ &= \hat{x}_{t,h} + \sum_{j=0}^{h-1} b_j \underbrace{\mathbf{E}[\epsilon_{t+h-1} | \epsilon_t, \epsilon_{t-1}, \dots]}_{=0} = \hat{x}_{t,h}. \end{aligned}$$

Der bedingte Erwartungswert ist hier also linear und somit bedeutet die Beschränkung auf lineare Funktionen keinen Verlust der Prognosegüte. Der Innovationsprozess ist insbesondere dann eine Martingal-Differenz, wenn  $(\epsilon_t)$  ein IID (unabhängig, identisch, verteilt) Prozess ist, oder noch stärker, wenn  $(\epsilon_t)$  bzw.  $(x_t)$  ein Gauß-Prozess ist.

#### Aufgabe (Charakterisierung von MA(q)-Prozessen)

Zeigen Sie, dass ein Prozess  $(x_t)$  mit  $\gamma(q) \neq 0$  und  $\gamma(k) = 0$  für  $|k| > q > 0$  ein MA(q)-Prozess ist. Hinweis: Zeigen Sie zunächst  $\hat{x}_{t,h} = 0$  für  $h > q$  und verwenden Sie dann die Wold-Zerlegung von  $(x_t)$  und die entsprechende Darstellung der  $(q + 1)$ -Schritt-Prognose  $\hat{x}_{t,q+1}$ .

#### Aufgabe

Sei  $(\epsilon_t) \sim \text{WN}(\sigma^2)$  ein skalares weißes Rauschen,  $z$  eine quadratisch integrierbare Zufallsvariable (mit  $\mathbf{E}z\epsilon_t = 0 \ \forall t \in \mathbb{Z}$ ) und  $(y_t = \epsilon_t - b\epsilon_{t-1})$ ,  $b \in \mathbb{R}$  ein MA(1)-Prozess. Zeigen Sie

- (1) Der Prozess  $(y_t)$  ist regulär und der Prozess  $(z_t = z)$  ist singulär.
- (2) Der Prozess  $(x_t = y_t + z)$  ist weder singulär noch regulär. Der Prozess  $(y_t)$  ist der reguläre Teil und  $(z_t = z)$  der singuläre Teil von  $(x_t)$ . Siehe Punkte (1) im Satz 2.3.

Wir werden in Abschn. 4.5 zeigen, dass die Darstellung  $y_t = \epsilon_t - b\epsilon_{t-1}$  dann und nur dann der Wold-Darstellung (2.14) entspricht, wenn  $|b| \leq 1$  gilt. Nur in diesem Fall sind die  $\epsilon_t$ s also die Innovationen von  $(y_t)$  und von  $(x_t)$ .

Modelle der Zeitreihenanalyse

Deistler, M.; Scherrer, W.

2018, X, 159 S. 10 Abb., 9 Abb. in Farbe., Softcover

ISBN: 978-3-319-68663-9

A product of Birkhäuser Basel