7 Vorhersage bei stationären Zeitreihen

Vorhersageproblem: Aus einem „Abschnitt" $X_1, \ldots, X_n$ einer reellen stationären Zeitreihe sollen die „zukünftigen" Werte X_t $(t = n+1, n+2, \ldots)$ vorhergesagt werden.

Optimalitätskriterium: Bestimme $\hat{X}_t = \hat{X}_t(X_1, \ldots, X_n)$ derart, dass

$$\|\hat{X}_t - X_t\| = \{E|\hat{X}_t - X_t|^2\}^{1/2} \overset{!}{=} \min \qquad (t \geq n+1)\,.$$

Wird $\hat{X}_t$ in einem (geeigneten) abgeschlossenen Teilraum $\mathcal{M}$ von $\mathcal{L}^2(P)$ bestimmt, so bedeutet dies:

$$(7.1) \qquad \hat{X}_t = P_{\mathcal{M}}\, X_t \qquad (t \geq n+1)\,,$$

wobei $P_{\mathcal{M}}$ die Projektion auf $\mathcal{M}$ bezeichnet.

Projektionen in Hilbert-Räumen

Sei $\mathcal{H}$ Hilbertraum mit Norm $\|\cdot\|$ und Skalarprodukt $\langle\cdot,\cdot\rangle$ und sei $\mathcal{M}$ ein (bzgl. $\|\cdot\|$) abgeschlossener Teilraum. Dann gilt:

1) $\mathcal{M}^\perp = \{x \in \mathcal{H} \mid \langle x, y\rangle = 0 \;\; \forall\, y \in \mathcal{M}\}$ ist ein abgeschlossener Teilraum (gilt für beliebige Teilmengen $\mathcal{M}$ von $\mathcal{H}$);

2) $\forall\, x \in \mathcal{H}$ $\exists_1$ (eindeutig) $\hat{x} \in \mathcal{M}$: $\|x - \hat{x}\| = \min\limits_{y \in \mathcal{M}} \|x - y\|$;

   $\hat{x} =: P_{\mathcal{M}}\, x$ heißt Orthogonalprojektion von x auf $\mathcal{M}$ und ist charakterisiert durch die Bedingungen:

   $$(7.2) \qquad \hat{x} \in \mathcal{M} \quad \text{und} \quad x - \hat{x} \in \mathcal{M}^\perp\,;$$

3) Eigenschaften der Projektion:

   a) $P_{\mathcal{M}}: \; \mathcal{H} \to \mathcal{M}$ ist linear und stetig;

   b) $\|x\|^2 = \|\hat{x}\|^2 + \|x - \hat{x}\|^2$;

   c) $x \in \mathcal{M} \iff P_{\mathcal{M}}\, x = x$,
      $x \in \mathcal{M}^\perp \iff P_{\mathcal{M}}\, x = 0$;

   d) $\mathcal{M}_1 \subset \mathcal{M}_2 \iff P_{\mathcal{M}_1} P_{\mathcal{M}_2}\, x = P_{\mathcal{M}_1}\, x \quad \forall\, x \in \mathcal{H}$.

Aus praktischen Gründen beschränkt man sich beim Vorhersageproblem i.A. darauf, beste lineare Vorhersagen zu finden, d.h. man finde

$$\hat{X}_t = \hat{c}_0 + \sum_{j=1}^{n} \hat{c}_j X_j \qquad \text{derart, dass}$$

$$\left\| \hat{X}_t - X_t \right\|_{\mathcal{L}^2(P)} \stackrel{!}{=} \min \qquad (t \geq n)\,.$$

Hier also: $\mathcal{H} = \mathcal{L}^2(P)$, $\langle X, Y \rangle = EXY$, $\|X\|^2 = EX^2$,
$\mathcal{M} = [\{1, X_1, \ldots, X_n\}]$ (abgeschlossen).

Sei nun $\{X_n\}$ relle stationäre Zeitreihe mit E.W. m und acv.f. $\gamma(\boldsymbol{\cdot})$
$\Longrightarrow$ $\{X_n - m\}$ zentriert und stationär mit acv.f. $\gamma(\boldsymbol{\cdot})$.

Es gilt: $P_{[\{1,X_1,\ldots,X_n\}]}\, X_{n+h} = m + P_{[\{X_1-m,\ldots,X_n-m\}]}\,(X_{n+h} - m)$

Daher o.E. im Folgenden: $\boxed{m = EX_n = 0}$

1-Schritt-Vorhersage $(h = 1)$

Es gilt: $\hat{X}_{n+1} = c_{n,1} X_n + \ldots + c_{n,n} X_1$, wobei wegen (7.2) gelten muss:

$$(7.3) \qquad \Big\langle \underbrace{\sum_{j=1}^{n} c_{n,j}\, X_{n+1-j}}_{\hat{X}_{n+1}}, X_{n+1-k} \Big\rangle = \big\langle X_{n+1}, X_{n+1-k} \big\rangle \qquad (k = 1, \ldots, n)\,,$$

d.h.

$$(7.3') \qquad \sum_{j=1}^{n} c_{n,j}\, \gamma(j-k) = \gamma(k) \qquad (k = 1, \ldots, n)\,.$$

Setzt man $\Gamma_n = \big(\!\big(\gamma(j-k)\big)\!\big)_{j,k=1,\ldots,n}$, $\gamma_n = \big(\gamma(1), \ldots, \gamma(n)\big)^\top$ und $c_n = (c_{n,1}, \ldots, c_{n,n})^\top$, so schreibt sich (7.3') in Matrixform als:

$$(7.3'') \qquad \Gamma_n\, c_n = \gamma_n \qquad (\,c_n \text{ nicht notwendig eindeutige Lösung}\,)\,.$$

Falls Γ_n regulär: $\boxed{c_n = \Gamma_n^{-1}\, \gamma_n}$

Satz 7.1. *Sei $\{X_n\}$ zentrierte stationäre Zeitreihe mit acv.f. $\gamma(\cdot)$. Falls $\gamma(0) > 0$ und $\gamma(h) \to 0$ $(h \to \infty)$, so ist die Kovarianzmatrix Γ_n von $(X_1, \ldots, X_n)^\top$ regulär $\forall\, n \in \mathbb{N}$.*

Unter den Voraussetzungen von Satz 7.1:

$$\hat{X}_{n+1} = \sum_{j=1}^{n} c_{n,j}\, X_{n+1-j}$$

mit $c_n = (c_{n,1}, \ldots, c_{n,n})^\top$ gemäß (7.3") ist beste lineare Vorhersage von X_{n+1} mit Vorhersagefehler (m.s.e.)

$$(7.4) \qquad v_n := E(X_{n+1} - \hat{X}_{n+1})^2 = \gamma(0) - \gamma_n^\top\, \Gamma_n^{-1}\, \gamma_n\,.$$

***h*-Schritt-Vorhersage** $(h \geq 1)$

Entsprechend ist

$$\hat{X}_{n+h} \;=\; \sum_{j=1}^{n} c_{nj}^{(h)}\, X_{n+1-j}$$

beste lineare Vorhersage, wobei $c_n^{(h)} = (c_{n,1}^{(h)}, \ldots, c_{n,n}^{(h)})^\top$ Lösung ist von

$$(7.5) \qquad \Gamma_n\, c_n^{(h)} = \gamma_n^{(h)} \quad , \quad \gamma_n^{(h)} = \big(\gamma(h), \gamma(h+1), \ldots, \gamma(n+h+1)\big)^\top.$$

Bemerkung 7.1. Die Linearkoeffizienten c_{nj} $\big($bzw. $c_{nj}^{(h)}\big)$ aus (7.3") (bzw. (7.5)) können rekursiv bestimmt werden, ohne die Matrix Γ_n zu invertieren (großer numerischer Vorteil!):

Durbin-Levinson-Algorithmus:

Sei $\{X_j\}_{j\in\mathbb{Z}}$ zentrierte stationäre Zeitreihe mit acv.f. $\gamma(\cdot)$, für die $\gamma(0) > 0$ und $\gamma(h) \to 0$ $(h \to \infty)$ gilt. Die Koeffizienten c_{nj} aus (7.3") und Vorhersagefehler v_n aus (7.4) genügen den Rekursionsformeln:

$$c_{1,1} \;=\; \frac{\gamma(1)}{\gamma(0)} \quad , \quad v_0 = \gamma(0) \quad ,$$

$$c_{n,n} \;=\; \Big\{\gamma(n) \;-\; \sum_{j=1}^{n-1} c_{n-1,j}\, \gamma(n-j)\Big\} \Big/ v_{n-1} \quad ,$$

$$\begin{bmatrix} c_{n,1} \\ \vdots \\ c_{n,n-1} \end{bmatrix} = \begin{bmatrix} c_{n-1,1} \\ \vdots \\ c_{n-1,n-1} \end{bmatrix} - c_{nn} \begin{bmatrix} c_{n-1,n-1} \\ \vdots \\ c_{n-1,1} \end{bmatrix} ,$$

und

$$v_n \;=\; v_{n-1}\,(1-c_{n,n}^2)\,.$$

(vgl. Brockwell & Davis (1991), Proposition 5.2.1).

Der folgende Algorithmus ist allgemeiner anwendbar und kann auch bei nicht-stationären Zeitreihen eingesetzt werden. Wir setzen dann

$$\gamma(j,k) = Cov(X_j, X_k)\,, \qquad j,k \in \mathbb{Z}\,.$$

Innovations-Algorithmus:

Sei $\{X_j\}_{j\in\mathbb{Z}}$ zentrierte Zeitreihe mit Kovarianzfunktion $\gamma(j,k) = EX_jX_k$ und sei $\big(\!\big(\gamma(j,k)\big)\!\big)_{j,k=1,\ldots,n}$ regulär für alle $n \in \mathbb{N}$.

Dann ergeben sich die 1-Schritt-Vorhersagen $\hat{X}_{n+1}$ und ihre Vorhersagefehler $v_n = E(X_{n+1}-\hat{X}_{n+1})^2$ wie folgt:

$$(7.6) \qquad \hat{X}_{n+1} \;=\; \begin{cases} 0 & ,\quad n=0\;, \\ \sum\limits_{j=1}^{n} d_{n,j}\left(X_{n+1-j}-\hat{X}_{n+1-j}\right) & ,\quad n\geq 1\;, \end{cases}$$

wobei

$$(7.7) \qquad \begin{cases} v_0 &=\; \gamma(1,1)\;, \\ d_{n,n-k} &=\; \dfrac{1}{v_k}\left(\gamma(n+1,k+1) - \sum\limits_{j=0}^{k-1} d_{k,k-j}\,d_{n,n-j}\,v_j\right)\,, \qquad k=0,\ldots,n-1\;, \\ v_n &=\; \gamma(n+1,n+1) - \sum\limits_{j=0}^{n-1} d_{n,n-j}^2\,v_j\;. \end{cases}$$

Beispiel 7.1. (Vorhersage von MA(1)-Reihen)

$X_n = e_n + \beta\, e_{n-1}\,, \quad \{e_n\}_{n\in\mathbb{Z}} \overset{\mathcal{D}}{\sim} WN(0,\sigma^2)\,.$

Hier gilt: $\gamma(j,k)=0 \;\;(|j-k|>1)\,, \;\; \gamma(j,j) = (1+\beta^2)\sigma^2\,, \;\; \gamma(j,j+1) = \beta\sigma^2\,.$

Der Innovations-Algorithmus liefert (rekursiv):

$$\begin{aligned} v_0 &= (1+\beta^2)\,\sigma^2 \;, \\ d_{n,j} &= 0 \;, \qquad j=2,\ldots,n\;, \\ d_{n,1} &= \frac{1}{v_{n-1}}\,\beta\,\sigma^2 \;, \\ v_n &= \left\{1+\beta^2 - \frac{1}{v_{n-1}}\,\beta^2\sigma^2\right\}\sigma^2\;. \end{aligned}$$

Setzt man $r_n = v_n/\sigma^2$, so erhält man die Rekursion

$$\hat{X}_{n+1} = \beta\big(X_n - \hat{X}_n\big)/r_{n-1} \qquad \big[= d_{n,1}(X_n - \hat{X}_n)\big]$$

mit $r_0 = 1 + \beta^2$, $r_n = 1 + \beta^2 - (\beta^2/r_{n-1})$ $(n = 1, 2, \ldots)$.

Numerisches Beispiel: Vorhersage von x_6 aus den Werten $x_1, \ldots, x_5$ einer MA(1)-Reihe mit $\beta = -0.9$, $\sigma^2 = 1$:

n	x_{n+1}	$\hat{x}_{n+1}$	r_n
0	-2.58	0.00	1.810
1	1.62	1.28	1.362
2	-0.96	-0.22	1.215
3	2.62	0.55	1.144
4	-1.36	-1.63	1.102
5		-0.22	1.075
			$\vdots$
			$\downarrow$ $(n \to \infty)$
			1

h-Schritt-Vorhersage $(h \geq 1)$: Setze $\hat{\hat{X}}_{n+h} := P_{\mathcal{M}_n} X_{n+h}$.

$$\text{(7.8)} \qquad \hat{\hat{X}}_{n+h} = \sum_{j=h}^{n+h-1} d_{n+h-1,j}\,\big(X_{n+h-j} - \hat{X}_{n+h-j}\big) \quad \text{mit}$$

$$\text{(7.9)} \qquad v_{n+h} = E\big(X_{n+h} - \hat{\hat{X}}_{n+h}\big)^2 = \gamma(n+h, n+h) - \sum_{j=h}^{n+h-1} d^2_{n+h-1,j}\, v_{n+h-j}\,,$$

wobei die $d_{n,j}$ wieder gemäß (7.7) bestimmt sind.

Rekursive Vorhersage bei kausalen ARMA(p, q)-Reihen

Der Innovations-Algorithmus kann direkt auf ARMA(p, q)-Reihen angewandt werden, aber es ergibt sich noch eine erhebliche Vereinfachung nach geeigneter Transformation. Gelte

$$a(B)\,X_n = b(B)\,e_n\,, \quad \{e_n\} \overset{\mathcal{D}}{\sim} WN(0, \sigma^2)\,, \quad B \text{ Backward Shift-Operator}\,,$$

mit $a(B) = 1 - a_1 B - \cdots - a_p B^p$, $a_p \neq 0$; $b(B) = 1 + b_1 B + \cdots + b_q B^q$, $b_q \neq 0$.

Statt $\{X_n\}$ betrachte man zunächst die transformierte Reihe $\{Y_n\}$, wobei

$$(7.10)\quad \begin{cases} Y_n = \dfrac{1}{\sigma} X_n & , \quad n = 1, \dots, r = \max(p,q)\ ; \\ Y_n = \dfrac{1}{\sigma} a(B)\, X_n & , \quad n = r+1, \dots\ . \end{cases}$$

Es gilt: $\mathcal{M}_n = [\{X_1, \dots, X_n\}] = [\{Y_1, \dots, Y_n\}]$,

denn „ $\supset$ “ direkt aus (7.10)

und „ $\subset$ “ rekursiv aus (7.10), z.B.

$$X_{r+1} = \sigma Y_{r+1} + a_1 X_r + \cdots + a_p X_{r+1-p} \in [\{Y_1, \dots, Y_{r+1}\}] \quad \text{usw.}$$

Sei nun $\hat{X}_{n+1} = P_{\mathcal{M}_n} X_{n+1}$, $\hat{Y}_{n+1} = P_{\mathcal{M}_n} Y_{n+1}$.

Die Autokovarianzfunktion $\gamma_X(\cdot)$ kann gemäß § 5 bestimmt werden. Für die Kovarianzen $\gamma(j,k) = EY_jY_k$ $(j \le k)$ ergibt sich dann:

$$\gamma(j,k) = \begin{cases} \dfrac{1}{\sigma^2}\, \gamma_X(j-k) & , \quad 1 \le j \le k \le r\ ; \\ \dfrac{1}{\sigma^2} \left\{ \gamma_X(j-k) - \sum\limits_{\ell=1}^{p} a_l\, \gamma_X(\ell + j - k) \right\} & , \quad 1 \le j \le r < k \le 2r\ ; \\ \sum\limits_{\ell=0}^{q} b_\ell\, b_{\ell+k-j} & , \quad r < j \le k \quad (b_0 = 1)\ ; \\ 0 & , \quad \text{sonst}\ . \end{cases}$$

[Beachte: $a(B)\, X_n = b(B)\, e_n$, $e_n \overset{\mathcal{D}}{\sim} WN(0, \sigma^2)$.]

Der Innovationsalgorithmus (für $\{Y_j\}$) liefert:

$$(7.11)\quad \begin{cases} \hat{Y}_{n+1} = \sum\limits_{j=1}^{n} d_{n,j} \left(Y_{n+1-j} - \hat{Y}_{n+1-j} \right) & , \quad n = 1, \dots, r-1\ , \\ \hat{Y}_{n+1} = \sum\limits_{j=1}^{q} d_{n,j} \left(Y_{n+1-j} - \hat{Y}_{n+1-j} \right) & , \quad n \ge r\ , \end{cases}$$

mit $d_{n,j}$ und $r_n = E\big(Y_{n+1} - \hat{Y}_{n+1}\big)^2$ gemäß (7.7).

Beachte: $d_{n,j} = 0$, falls $n \ge r$ und $j > q$ wegen (7.7)

und $\gamma(n,j) = 0$, falls $n > r$ und $n - j > q$.

$\left[Y_n = \dfrac{1}{\sigma} \Big(e_n + b_1 e_{n-1} + \cdots + b_q e_{n-q} \Big) \right]$

Die $\hat{X}_n$ lassen sich nun aus den $\hat{Y}_n$ wie folgt bestimmen:

Projektion von (7.10) auf $\mathcal{M}_{n-1}$ liefert:

$$
(7.12)\quad \left.\begin{cases} \hat{Y}_n = \dfrac{1}{\sigma}\hat{X}_n & , \quad n = 1,\dots,r \\ \hat{Y}_n = \dfrac{1}{\sigma}\big(\hat{X}_n - a_1 X_{n-1} - \cdots - a_p X_{n-p}\big) & , \quad n > r \end{cases}\right\} \overset{(7.10)-(7.12)}{\Longrightarrow}
$$

$$
(7.12\,')\qquad X_n - \hat{X}_n = \sigma\big(Y_n - \hat{Y}_n\big) \qquad \forall\, n = 1, 2, \dots\,.
$$

Einsetzen von (7.12)/(7.12') in (7.11) liefert schließlich:

$$
(7.13)\quad \begin{cases} \hat{X}_{n+1} = \sum\limits_{j=1}^{n} d_{n,j}\big(X_{n+1-j} - \hat{X}_{n+1-j}\big) & , \quad 1 \le n < r\,; \\ \hat{X}_{n+1} = a_1 X_n + \cdots + a_p X_{n+1-p} + \sum\limits_{j=1}^{q} d_{n,j}\big(X_{n+1-j} - \hat{X}_{n+1-j}\big) & , \quad n \ge r\,. \end{cases}
$$

Bemerkung 7.2. Zur Vorhersage von $\hat{X}_{n+1}$ $(n \ge r)$ benötigt man nur die p Beobachtungen $X_n, \dots, X_{n-p+1}$ sowie die q Innovationen $X_n - \hat{X}_n, \dots, X_{n+1-q} - \hat{X}_{n+1-q}$ (großer numerischer Vorteil!).

Beispiel 7.2. (Vorhersage von ARMA(1, 1)-Reihen)

Gelte $X_n - \alpha X_{n-1} = e_n + \beta e_{n-1}$, $|\alpha| < 1$ (kausal), $\{e_n\} \overset{\mathcal{D}}{\sim} WN(0, \sigma^2)$.

$$
(7.13) \overset{(r=1)}{\Longrightarrow} \quad \hat{X}_{n+1} = \alpha X_n + d_{n,1}\big(X_n - \hat{X}_n\big), \quad n \ge 1\,.
$$

Bestimme $d_{n,1}$: Aus Formel (5.12) erhält man

$$
\begin{aligned}
\gamma_X(0) - \alpha\,\gamma_X(1) &= \sigma^2(1 + \alpha\beta + \beta^2)\,, \\
\gamma_X(1) - \alpha\gamma_X(0) &= \sigma^2\beta
\end{aligned}
$$

$$
\Longrightarrow \qquad \gamma_X(0)(1 - \alpha^2) = \sigma^2(1 + 2\alpha\beta + \beta^2)
$$

$$
\Longrightarrow \qquad \gamma(j,k) \overset{(r=1)}{=} \begin{cases} \dfrac{1 + 2\alpha\beta + \beta^2}{1 - \alpha^2} & , \quad j = k = 1\,; \\ 1 + \beta^2 & , \quad j = k \ge 2\,; \\ \beta & , \quad |j - k| = 1\,; \\ 0 & , \quad \text{sonst}\,. \end{cases}
$$

$$\overset{(7.7)}{\Longrightarrow} \quad \begin{cases} r_0 &= \dfrac{1+2\alpha\beta+\beta^2}{1-\alpha^2} \quad, \\ d_{n,1} &= \dfrac{\beta}{r_{n-1}} \quad, \\ r_n &= 1+\beta^2-\dfrac{\beta^2}{r_{n-1}} \quad. \end{cases}$$

Diese Gleichungen können leicht nach den r_n und $d_{n,1}$ aufgelöst werden (vgl. Brockwell und Davis (1991), Problem 5.13) .

Vorhersage im Frequenzbereich

Hat man eine Spektraldarstellung einer zentrierten stationären Zeitreihe $\{X_n\}_{n\in\mathbb{Z}}$ mit Spektralfunktion F und zugehörigem Prozess $\{Z(\lambda)\}_{\lambda\in[-\pi,\pi]}$ mit orthogonalen Zuwächsen, so definiert das stochastische Integral

$$\mathcal{I}(g) \;=\; \int\limits_{(-\pi,\pi]} g(\lambda)\,dZ(\lambda)\,, \qquad g\in\mathcal{L}^2(F)\,,$$

eine Isometrie zwischen $\mathcal{L}^2(F)$ und $\overline{\mathcal{H}} = \overline{sp}\{X_n, n\in\mathbb{Z}\}$ mit

$$\mathcal{I}(e^{in\,\bullet}) \;=\; X_n\,, \qquad n\in\mathbb{Z}\,.$$

[<u>Hier</u> : $\overline{sp}\,U$ = kleinster abgeschlossener Teilraum von $\mathcal{L}^2(P) \supset U$.]

Dies macht es möglich, Projektionen (d.h. Vorhersagen) zunächst in $\mathcal{L}^2(F)$ zu bestimmen und dann nach $\mathcal{H}$ zu projezieren , denn es gilt z.B. (wegen der Isometrie) :

(7.14) $\quad P_{\mathcal{M}_n} X_{n+h} \;=\; \mathcal{I}\Big(P_{\overline{sp}\{e^{ij\,\bullet}\,:-\infty<j\le n\}}\, e^{i(n+h)\,\bullet}\Big)\,,$

wobei $\mathcal{M}_n = \overline{sp}\{X_j : -\infty < j \le n\,,\; j\in\mathbb{Z}\}$.

Beispiel 7.3. Sei $\{X_n\}_{n\in\mathbb{Z}}$ eine zentrierte stationäre Zeitreihe mit Spektraldichte f auf $[-\pi,\pi]$, wobei

$$f(\lambda) = 0\,, \quad 1 < |\lambda| \le \pi\,.$$

Wegen $|1-e^{-i\lambda}| < 1$ für $|\lambda| \le 1$ erhält man :

$$\sum_{k=0}^{\infty}\left(1-e^{-i\lambda}\right)^k \;=\; \sum_{k=0}^{\infty}\sum_{j=0}^{k}\binom{k}{j}(-1)^j\,e^{-i\lambda j} \;=\; \frac{1}{1-(1-e^{-i\lambda})} \;=\; e^{i\lambda}\,,$$

wobei die Konvergenz gleichmäßig ist für $\lambda\in[-1,1]$. Daher :

$$e^{i(n+1)\lambda} \;=\; \sum_{k=0}^{\infty}\sum_{j=0}^{k}\binom{k}{j}(-1)^j\,e^{i\lambda(n-j)}\,, \qquad |\lambda| \le 1\,,$$

gleichmäßig in $\lambda \in [-1,1]$, aber auch in $\mathcal{L}^2(F)$, da wegen $|1 - e^{-i\lambda}| < 1$ und der endlich geometrischen Reihe die N-ten Partialsummen gleichmäßig beschränkt und damit integrierbar (bzgl. F) sind, so dass der Satz von Lebesgue auch die $\mathcal{L}^2$-Konvergenz liefert.

Offensichtlich: $e^{i(n+1)\bullet} \in \overline{sp}\{e^{ij\bullet} : -\infty < j \leq n\}$, also

$$P_{\overline{sp}\{e^{ij\bullet}:-\infty<j\leq n\}}\, e^{i(n+1)\bullet} = e^{i(n+1)\bullet},$$

so dass wegen (7.14) folgt:

$$P_{\mathcal{M}_n} X_{n+1} = \mathcal{I}\big(e^{i(n+1)\bullet}\big) = X_{n+1}, \qquad n \in \mathbb{Z}.$$

Eine direkte Vorhersage von X_{n+1} im Zeitbereich (s.o.) wäre wesentlich aufwendiger.

Bemerkung 7.3. In Beispiel 7.3 ist X_{n+1} „perfekt vorhersagbar" durch die Werte der Reihe $\{X_t : -\infty < t \leq n\}$ (bis z.Zt. n). Derartige Prozesse heißen „deterministisch" (s.u.). Es läßt sich zeigen, dass nicht-deterministische stationäre Zeitreihen eine Zerlegung besitzen

$$X_n = U_n + V_n, \qquad n \in \mathbb{Z},$$

wobei $\{U_n\}_{n\in\mathbb{Z}}$ eine MA(∞)-Reihe und $\{V_n\}_{n\in\mathbb{Z}}$ eine „deterministische" Zeitreihe ist, die zu $\{U_n\}_{n\in\mathbb{Z}}$ unkorreliert (orthogonal) ist.

Im Beweis dieser Wold-Zerlegung benutzen wir die folgenden Bezeichnungen:

$$\sigma^2 := E|X_{n+1} - P_{\mathcal{M}_n} X_{n+1}|^2 \qquad \text{(1-Schritt-Vorhersagefehler)},$$

$$\mathcal{M}_{-\infty} := \bigcap_{n=-\infty}^{\infty} \mathcal{M}_n \quad \big(\subset \mathcal{M} := \overline{sp}\,\{X_j : j \in \mathbb{Z}\}\big).$$

Bemerkung 7.4. $\{X_n\}_{n\in\mathbb{Z}}$ heißt deterministisch, falls $\sigma^2 = 0$ bzw. (äquivalent) falls $X_n \in \mathcal{M}_{-\infty}$ $\forall\, n \in \mathbb{Z}$, d.h. $\mathcal{M}_{-\infty} = \mathcal{M}$.

Satz 7.2. (*Wold-Zerlegung*) *Sei* $\{X_n\}_{n\in\mathbb{Z}}$ *zentrierte stationäre Zeitreihe mit* $\sigma^2 > 0$. *Dann gilt:*

$$X_n = \sum_{j=0}^{\infty} c_j\, e_{n-j} + V_n\,, \tag{7.15}$$

wobei

(i) $c_0 = 1\,, \quad \sum_{j=0}^{\infty} c_j^2 < \infty\,;$

(ii) $\{e_n\} \overset{\mathcal{D}}{\sim} WN(0,\sigma^2)\,;$

(iii) $e_n \in \mathcal{M}_n \quad \forall\, n \in \mathbb{Z}\,;$

(iv) $E(e_n V_m) = 0 \quad \forall\, m, n \in \mathbb{Z}\,;$

(v) $V_n \in \mathcal{M}_{-\infty} \quad \forall\, n \in \mathbb{Z}\,;$

(vi) $\{V_n\}$ *deterministisch*.

Ferner sind die Folgen $\{c_n\}, \{e_n\}$ *und* $\{V_n\}$ *eindeutig durch* (7.15) *und die Bedingungen* (i) – (vi) *festgelegt.*

Bemerkung 7.5. Man beachte, dass (v) und (vi) nicht identisch sind, da $\mathcal{M}_{-\infty}$ für $\{X_n\}$, nicht für $\{V_n\}$ definiert ist.

Der Beweis von Satz 7.2 hat noch gezeigt, dass gilt:

Korollar 7.1. *Unter der Voraussetzung von* Satz 7.2 *ergibt sich:*

a) $\overline{sp}\,\{V_j : -\infty < j \le n\} = \mathcal{M}_{-\infty} \quad \forall\, n \in \mathbb{Z}\,;$

b) $\mathcal{M}_n = \overline{sp}\,\{e_j : -\infty < j \le n\} \oplus \mathcal{M}_{-\infty}\,;$

c) $\mathcal{M}_{-\infty}^{\perp} = \overline{sp}\,\{e_n : n \in \mathbb{Z}\}\,;$

d) $\overline{sp}\,\{U_j : -\infty < j \le n\} = \overline{sp}\,\{e_j : -\infty < j \le n\}\,,$ *wobei* $U_n = \sum_{j=0}^{\infty} c_j\, e_{n-j}\,.$

Interpretation von (7.15): Man erhält eine Zerlegung der Zeitreihe in zwei orthogonale Anteile gemäß der Zerlegung in b), d.h. $\mathcal{M}_n = \overline{sp}\,\{e_j : -\infty < j \le n\} \oplus \mathcal{M}_{-\infty}$.

Bemerkung 7.6. Eine stationäre Zeitreihe heißt rein nicht-deterministisch, wenn $\mathcal{M}_{-\infty} = \{0\}$. In diesem Fall gilt:

$$X_n = \sum_{j=0}^{\infty} c_j \, e_{n-j} \,, \qquad \underline{\text{also eine MA}(\infty)\text{-Darstellung}}$$

(vgl. z.B. kausale ARMA(p,q)-Reihen).

h-Schritt-Vorhersage: Die Wold-Zerlegung (7.15) liefert

$$P_{\mathcal{M}_n} X_{n+h} \;=\; \sum_{j=h}^{\infty} c_j \, e_{n+h-j} + V_{n+h} \,,$$

da $e_{n+1}, \ldots, e_{n+h} \perp \mathcal{M}_n$, $V_{n+h} \in \mathcal{M}_{-\infty} \subset \mathcal{M}_n$.

Vorhersagefehler: $E|X_{n+h} - P_{\mathcal{M}_n} X_{n+h}|^2$

$$= \quad Var\left(\sum_{j=0}^{h-1} c_j e_{n+h-j}\right) \;=\; \sigma^2 \sum_{j=0}^{h-1} c_j^2$$

$$\underset{(h\to\infty)}{\longrightarrow} \quad \sigma^2 \sum_{j=0}^{\infty} c_j^2 \;=\; Var\left(\sum_{j=0}^{\infty} c_j e_{n-j}\right),$$

d.h., falls $\{X_n\}$ rein nicht-deterministisch ist, so konvergiert der h-Schritt-Vorhersagefehler gegen die Varianz des Prozesses.

Beispiel 7.4. Betrachte die stationäre Zeitreihe

$$X_n = e_n + V \,, \qquad n \in \mathbb{Z} \,,$$

wobei $\{e_n\} \overset{\mathcal{D}}{\sim} WN(0,\sigma^2)$, $V \perp \{e_n\}$ mit $EV = 0$, $EV^2 < \infty$.

Da

$$\frac{1}{N} \sum_{j=0}^{N-1} X_{n-j} = \frac{1}{N} \sum_{j=0}^{N-1} e_{n-j} + V \;\underset{(N\to\infty)}{\overset{\mathcal{L}^2(P)}{\longrightarrow}}\; V \,,$$

folgt: $V \in \mathcal{M}_n \;\; \forall\, n \in \mathbb{Z}$, also $V \in \mathcal{M}_{-\infty}$.

Ferner: $e_n \perp \mathcal{M}_j \;\; \forall\, j < n$, d.h. $e_n \perp \mathcal{M}_{-\infty}$.

Fazit: $V = P_{\mathcal{M}_{-\infty}} X_n$ ist die deterministische Komponente

und $e_n = X_n - V$ die rein nicht-deterministische Komponente

in der Wold-Zerlegung von X_n.

Abschließende Bemerkung: Die Spektraldarstellung hat sich als sehr nützlich erwiesen für die probabilistsche Analyse von stationären Zeitreihen. Sie bildet darüber hinaus ein wichtiges Hilfsmittel bei der statistischen Analyse, sowohl im Zeitbereich wie auch im Frequenzbereich der Reihe.