Skriptum

Zeitreihenanalyse I und II

Definitionen und Sätze

Inhaltsverzeichnis

1 Elemente explorativer Zeitreihenanalyse

Statistische Methoden, die Unabhängigkeit oder eine identische Verteilung der Zufallsvariablen erfordern können bei Zeitreihen nicht angewendet werden.

1.1 Das Additive Modell einer Zeitreihe

Es wird angenommen, dass die $y_1, \dots, y_n$ Realisationen einer zufälligen Variablen Y_t sind, die erfüllt:

$$Y_t = T_t + Z_t + S_t + R_t$$

Dabei wird angenommen, dass $E(R_t) = 0$ ist (ggf. Anpassung der Anderen).

1.1.1 Modelle mit einem nichtlinearen Trend

Wir nehmen an es gilt $Y_t = T_t + R_t$ mit $E(Y_T) = T_t =: f(t)$. Man nimmt an, dass diese Funktion von unbekannten Parametern abhängen, die man gewöhnlich durch einen kleinsten Quadrate Ansatz schätzt.

$$\min_\beta \sum_t (y_t - f(t, \beta_1, \dots, \beta_n))^2$$

1.1.2 Die Logistische Funktion

Hier besteht ein linearer Zusammenhang in $1/f_{\log}(t)$.

1.1.3 Die Mitscherlich Funktion

1.1.4 Die Gompertz Kurve

1.1.5 Die Allometrische Funktion

Logtransformation liefert hier einen linearen Zusammenhang.

Zur Beurteilung der Anpassungsgüte eignet sich beispielsweise der R^2 Wert (Achtung nur bei linearer Regression mit kleinsten Quadraten zwischen 0 und 1).

$$R^2 := 1 - \frac{\sum_{t=1}^{n}(y_t - \hat{y}_t)^2}{\sum_{t=1}^{n}(y_t - \bar{y})^2}$$

1.2 Lineare Filterung einer Zeitreihe

Trend und saisonbereinigte Daten. Man betrachte ($E(R_t) = 0$):

$$Y_t = T_t + S_t + R_t$$

Dann ist $y_t - \hat{T}_t$ trendbereinigt und $y_t - \hat{S}_t$ saisonbereinigt.

1.2.1 Lineare Filter

Zusatzdefinition. Definition eines *linearen Filters* mit *Gewichten* $a_{-r}, \ldots, a_s$. (a_u) wird dann auch als linearer Filter bezeichnet.

$$Y_t^* := \sum_{u=-r}^{s} a_u Y_{t-u}, \qquad t = s+1, \ldots, n-r,$$

Zusatzdefinition. Definition eines *moving average*, *simple moving average.* (Trade-Off Situation für die Länge des Filters)

1.2.2 Saisonale Anpassung

Idee 1.2.1. *Bei einer saisonalen Komponente (Periode p) bildet man einen simple moving average Y_t^* der Länge p und erhält dann: $D_t := Y_t - Y_t^* \sim S_t + R_t$. Um nun S_t zu schätzen mittelt man alle diese D_i mit Abstand p: $\bar{D}_t$, $t = 1, \ldots, p$. Da deren Summe nicht zwingend Null ist zieht man von jedem das Mittel ab und erhält $\hat{S}_t$. $Y_t - \hat{S}_t$ sind dann die saisonal angepassten Werte.*

1.2.3 Das Census X-11 Programm

Angenommen wird das additive Modell $Y_t = T_t + S_t + R_t$ mit Periode 12. Vorgehensweise:

- Berechnung eines simple moving average der Länge 12 ($\cong$ Periode) und Berechnung von $D_t := Y_t - Y_t^* \sim S_t + R_t$.
- Anwenden eines moving average auf die D_i (Länge 5)mit lag 12. Anpassung zu $\hat{S}_t$ wie oben.
- Erhalten von $Y_t^{(1)} := Y_t - \hat{S}_t^{(1)} \sim T_t + R_t$ und nochmaliger Durchlauf der Schritte mit Moving Averages anderer Länge (9, 13 oder 23 und dann 7).

1.2.4 Beste lokale Polynomiale Anpassung

Man minimiert hier:

$$\sum_{u=-k}^{k} (y_{t+u} - \beta_0 - \beta_1\, u - \cdots - \beta_p\, u^p)^2 = \min .$$

Die partiellen Ableitungen müssen Null sein und Ergeben in Matrixschreibweise die *Normal Equations* $\boldsymbol{X}^T\boldsymbol{X}\beta = \boldsymbol{X}^T\boldsymbol{y}$. Falls die Matrix invertierbar ist erhält man eine Lösung und kann hieraus die $\hat{y}_{t+u}$ berechnen. Für $u = 0$ ist dies $\beta_0 = \hat{y}_t$. Dies verwendet man dann als beste lokale polynomiale Anpassung.

$$\boldsymbol{X} = \begin{pmatrix} 1 & -k & (-k)^2 & \cdots & (-k)^p \\ \vdots & & \ddots & \vdots & \\ 1 & k & k^2 & \cdots & k^p \end{pmatrix}$$

Zusatzbemerkung. Bei der besten lokalen polynomialen Anpassung handelt es sich um einen Moving Average (1. Zeile mal Spalte liefert Darstellung als Summe der y_{t+u}, Summation zu eins ergibt sich, wenn man Schätzer zu $y_t \equiv 1$ betrachtet).

1.2.5 Differenzfilter

Lemma 1.2.2. *For a polynomial $f(t) := c_0 + c_1 t + \cdots + c_p t^p$ of degree p, the difference*

$$\Delta f(t) := f(t) - f(t-1)$$

is a polynomial of degree at most $p - 1$.

Der lineare Filter $\Delta Y_t = Y_t - Y_{t-1}$ kann also einen linearen Trend entfernen. Analog der Filter $\Delta^2 Y_t$ (auch linearer Filter), der einen quadratischen Trend entfernt. Ebenso kann eine saisonale Komponente durch einen saisonalen Filter der Form $Y_t^* := Y_t - Y_{t-p}$ entfernt werden. (Beachte: Die Reihenfolge der Filter ist egal).

1.2.6 Exponentieller Glätter

$Y_t^* = \alpha Y_t + (1-\alpha)Y_{t-1}^*$, $Y_0^* = Y_0$.

Lemma 1.2.3. *For an exponential smoother with constant $\alpha \in [0,1]$ we have*

$$Y_t^* = \alpha \sum_{j=0}^{t-1} (1-\alpha)^j Y_{t-j} + (1-\alpha)^t Y_0, \qquad t = 1, 2, \ldots, n.$$

BEWEIS: Induktion. □

Korollar 1.2.4.
- *Suppose that the random variables $Y_0, \ldots, Y_n$ have common expectation μ and common variance. Dann gilt:* $\mathrm{E}(Y_t^*) = \mu$.

  If the Y_t are in addition uncorrelated, then $\mathrm{E}((Y_t^* - \mu)^2) \overset{t\to\infty}{\longrightarrow} \frac{\sigma^2\alpha}{2-\alpha} < \sigma^2$.

- *Suppose that the random variables $Y_0, Y_1, \ldots$ satisfy* $\mathrm{E}(Y_t) = \mu$ *for* $0 \le t \le N-1$, *and* $\mathrm{E}(Y_t) = \lambda$ *for* $t \ge N$. *Then we have for* $t \ge N$ $\mathrm{E}(Y_t^*) =\overset{t\to\infty}{\longrightarrow} \lambda$

1.2.7 Autocovarianz und Autocorrelation

Angenommen die $Y_1, \ldots, Y_n$ sind quadratintegrierbare Zufallsvariablen, bei denen die Kovarianz $\mathrm{Cov}(Y_{t+k}, Y_t) = \mathrm{E}((Y_{t+k} - \mathrm{E}(Y_{t+k}))(Y_t - \mathrm{E}(Y_t)))$ nicht von t abhängt, so ist

$$\gamma(k) := \mathrm{Cov}(Y_{k+1}, Y_1) = \mathrm{Cov}(Y_{k+2}, Y_2) = \ldots, \quad c(k) := \frac{1}{n}\sum_{t=1}^{n-k}(y_{t+k} - \bar{y})(y_t - \bar{y})$$

die (empirische) *Autokovarianzfunktion* und

$$\rho(k) := \frac{\gamma(k)}{\gamma(0)}, \qquad k = 0, 1, \ldots, \quad r(k) := \frac{c(k)}{c(0)} = \frac{\sum_{t=1}^{n-k}(y_{t+k} - \bar{y})(y_t - \bar{y})}{\sum_{t=1}^{n}(y_t - \bar{y})^2}.$$

die (empirische) *Autokorrelationsfunktion.* Es ist $\gamma(0) \ge 0$ und $|\gamma(k)| \le \gamma(0) \quad$ für $k \ge 0$ (Cauchy-Schwarz).

1.2.8 Varanzstabilisierende Transformation

Manchmal führen Transformationen (z.B. log) dazu, dass die Variation einer Zufallsvariablen nicht mit der Größe zunimmt ($Y_t = \sigma_t Z_t \Rightarrow \log(Y_t) = \log(\sigma_t) + \log(Z_t)$). Besonders beliebt sind die Box-Cox-Transformationen.

2 Zeitreihenmodelle

Y_t mit $t \in \mathbb{Z}$ wird als *Stoachastischer Prozess* bezeichnet.

2.1 Lineare Filter und Stochastische Prozesse

Zusatzdefinition. Definition einer *komplexwertigen Zufallsgröße* zusammen mit dem *Erwartungswert* und der *Varianz* (*Quadratintegrierbarkeit*) bzw. *Covarianz* (nicht symmetrisch)

$$\mathrm{Cov}(Y,Z) := \mathrm{E}((Y-\mathrm{E}(Y))(\overline{Z-\mathrm{E}(Z)})).$$

Lemma 2.1.1. *For any integrable complex valued random variable $Y = Y_{(1)} + iY_{(2)}$ we have*

$$|\,\mathrm{E}(Y)| \leq \mathrm{E}(|Y|) \leq \mathrm{E}(|Y_{(1)}|) + \mathrm{E}(|Y_{(2)}|).$$

Beweis: $|\,\mathrm{E}(Y)| = r = \mathrm{E}(e^{-i\vartheta}Y) = \mathrm{E}\left(\mathrm{Re}(e^{-i\vartheta}Y)\right) \leq \mathrm{E}(|Y|)$, mit Polarkoordinaten. □

Korollar 2.1.2. *For any square integrable complex valued random variable we have*

$$|\,\mathrm{E}(YZ)| \leq \mathrm{E}(|Y||Z|) \leq \mathrm{E}(|Y|^2)^{1/2}\,\mathrm{E}(|Z|^2)^{1/2}$$

and thus,

$$|\,\mathrm{Cov}(Y,Z)| \leq \mathrm{Var}(Y)^{1/2}\,\mathrm{Var}(Z)^{1/2}.$$

2.1.1 Stationäre Prozesse

Zusatzdefinition. Definition eines *stationären Prozesses* und Darstellung der *Autokovaranzfunktion.* Definition eines *weißen Rauschens*, *absolut summierbaren (linearen) Filters* und eines *allgemeinen linearen Prozess.*

$$Y_t := \sum_{u=-\infty}^{\infty} a_u \varepsilon_{t-u} := \sum_{u\geq 0} a_u\varepsilon_{t-u} + \sum_{u\geq 1} a_{-u}\varepsilon_{t+u}, \quad t\in\mathbb{Z},\ \text{mit}\ ,\ \sum_{t=-\infty}^{\infty}|a_t| < \infty$$

2.1.2 Existenz eines allgemeinen linearen Prozesses

Z.z ist $\sum_{u=-\infty}^{\infty} |a_u\varepsilon_{t-u}| < \infty$ mit Wahrscheinlichkeit Eins. $L_2 := L_2(\Omega, \mathcal{A}, \mathcal{P})$ Raum aller komplexwertigen quadratintegrierbaren Zufallswariablen mit $||Y||_2 := \mathrm{E}(|Y|^2)^{1/2}$.

Satz 2.1.3. *The space $(L_2, ||\cdot||_2)$ is complete i.e., suppose that $X_n \in L_2,\ n\in\mathbb{N}$, has the property that for arbitrary $\varepsilon > 0$ one can find an integer $N(\varepsilon)\in\mathbb{N}$ such that $||X_n - X_m||_2 < \varepsilon$ if $n,\ m \geq N(\varepsilon)$. Then there exists a random variable $X\in L_2$ such that $\lim_{n\to\infty} ||X - X_n||_2 = 0$.*

Beweis: Folgt aus einem vorherigen Lemma, mit Hilfe einer Teleskopsumme. □

Satz 2.1.4. *Suppose that $(Z_t)_{t\in\mathbb{Z}}$ is a stochastic process such that $\sup_t \mathrm{E}(|Z_t|) < \infty$ and let $(a_t)_{t\in\mathbb{Z}}$ be an absolutely summable filter. Then we have $\sum_{u\in\mathbb{Z}} |a_u Z_{t-u}| < \infty$ with probability one for $t\in\mathbb{Z}$ and, thus, $Y_t := \sum_{u\in\mathbb{Z}} a_u Z_{t-u}$ exists almost surely in $\mathbb{C}$. We have moreover $\mathrm{E}(|Y_t|) < \infty$, $t\in\mathbb{Z}$, and* *(i)* $\mathrm{E}(Y_t) = \lim_{n\to\infty} \sum_{u=-n}^{n} a_u \, \mathrm{E}(Z_{t-u})$, $t \in \mathbb{Z}$,

(ii) $\mathrm{E}(|Y_t - \sum_{u=-n}^{n} a_u Z_{t-u}|) \overset{n\to\infty}{\longrightarrow} 0$.

If, in addition, $\sup_t \mathrm{E}(|Z_t|^2) < \infty$, *then we have* $\mathrm{E}(|Y_t|^2) < \infty$, $t \in \mathbb{Z}$, *and*

(iii) $||Y_t - \sum_{u=-n}^{n} a_u Z_{t-u}||_2 \overset{n\to\infty}{\longrightarrow} 0$.

Beweis: Satz von der monotonen Konvergenz erlaubt Hineinziehen des Erwartungswertes $\Rightarrow$ Existenz von Y_t. Aus der dominierten Konvergenz folgt dann 1. und 2. Für 3. zeigen der Cauchy-Eigenschaft und dass der Grenzwert Y_t ist. $\square$

Satz 2.1.5. *Suppose that* $(Z_t)_{t\in\mathbb{Z}}$ *is a stationary process with mean* $\mu_Z := \mathrm{E}(Z_0)$ *and autocovariance function* γ_Z *and let* (a_t) *be an absolutely summable filter. Then* $Y_t = \sum_u a_u Z_{t-u}$, $t \in \mathbb{Z}$, *is also stationary with*

$$\mu_Y = \mathrm{E}(Y_0) = \Big(\sum_u a_u\Big)\mu_Z$$

and autocovariance function

$$\gamma_Y(t) = \sum_u \sum_w a_u \bar{a}_w \gamma_Z(t + w - u).$$

Beweis: Nach vorherigem Satz kann man einfach einsetzen und alles berechnen. $\square$

2.1.3 Die kovarianzerzeugende Funktion

Zusatzdefinition. Definition der *kovarianzerzeugenden Funktion* (Definiert in einem Ringgebiet und Eindeutig)

$$G(z) := \sum_{t\in\mathbb{Z}} \gamma(t) z^t = \sum_{t\geq 0} \gamma(t) z^t + \sum_{t\geq 1} \gamma(-t) z^{-t},$$

Satz 2.1.6. *Suppose that* $Y_t = \sum_u a_u \varepsilon_{t-u}$, $t \in \mathbb{Z}$, *is a general linear process with* $\sum_u |a_u||z^u| < \infty$, *if* $r^{-1} < |z| < r$ *for some* $r > 1$. *Put* $\sigma^2 := \mathrm{Var}(\varepsilon_0)$. *The process* (Y_t) *then has the covariance generating function*

$$G(z) = \sigma^2 \Big(\sum_u a_u z^u\Big)\Big(\sum_u \bar{a}_u z^{-u}\Big), \quad r^{-1} < |z| < r.$$

Beweis: $\mathrm{Cov}(Y_t, Y_0)$ nach vorherigem Satz berechnen mit $\gamma(i) = 0$ für $i \neq 0$ (Weißes Rauschen). Dann $G(z)$ mit Cauchyprodukt. $\square$

2.1.4 Das charakteristische Polynom

Zusatzdefinition. Definition des *charakteristischen Polynoms* zu einem absolut summierbaren Filter.

$$A(z) := \sum_{u \in \mathbb{Z}} a_u z^u, \quad \text{existiert auf einem Kreisgebiet oder gar nicht}$$

2.1.5 Inverse Filter

Lemma 2.1.7. *Let (a_u) and (b_u) be absolutely summable filters with characteristic polynomials $A_1(z)$ and $A_2(z)$, which both exist on some annulus $r < |z| < R$. The product filter $(c_v) = \left(\sum_{u+w=v} b_w a_u\right)$ then has the characteristic polynomial*

$$A(z) = A_1(z) A_2(z).$$

Beweis: Filtern zuerst mit b_u, dann mit a_u, dann Cauchyprodukt. Absolute Summierbarkeit aus der von a_u und b_u. □

Zusatzdefinition. Definition des *inversen Filters*, so dass $A_1(z)A_2(z) = 1$

$$Y_t = \sum_u a_u Z_{t-u} \quad \text{and} \quad \sum_w b_w Y_{t-w} = Z_t, \quad t \in \mathbb{Z}.$$

2.1.6 Kausaler Filter

Zusatzdefinition. Definition eines *kausalen Filters.*

Lemma 2.1.8. *Let $a \in \mathbb{C}$. The filter (a_u) with $a_0 = 1,\ a_1 = -a$ and $a_u = 0$ elsewhere has an absolutely summable and causal inverse filter $(b_u)_{u \geq 0}$ if and only if $|a| < 1$. In this case we have $b_u = a^u,\ u \geq 0$.*

Beweis: Berechnung von $A_2(z) = 1/(1 - az)$ mittels geometrischer Reihe. □

Satz 2.1.9. *Let $a_1, a_2, \ldots, a_p \in \mathbb{C},\ a_p \neq 0$. The filter (a_u) with coefficients $a_0 = 1, a_1, \ldots, a_p$ and $a_u = 0$ elsewhere has an absolutely summable and causal inverse filter if the p roots $z_1, \ldots, z_p \in \mathbb{C}$ of $A(z) = 1 + a_1 z + a_2 z^2 + \cdots + a_p z^p = 0$ are outside of the unit circle i.e., $|z_i| > 1$ for $1 \leq i \leq p$.*

Beweis: Zerlegung von $A(z)$ in Nullstellen. Dann für diese jeweils mit geometrischer Reihe. □

Bemerkung 2.1.10. Falls die a_p und die z_i reell sind, so auch der inverse Filter. Falls keine Nullstelle auf dem Einheitskreis liegt, so existiert der inverse Filter, nur für die Kausalität wird außerhalb des Einheitskreises benötigt.

2.2 Moving Averages and Autoregressive Prozesse

Zusatzdefinition. Definition eines *Moving Average* $MA(q)$.

$$Y_t := \varepsilon_t + a_1\varepsilon_{t-1} + \cdots + a_q\varepsilon_{t-q}, \quad B(z)1 + a_1z + \cdots + a_qz^q$$

Zusatzbemerkung. Gemäß vorherigen Sätzen handelt es sich immer um einen stationären Prozess (Satz 2.1.5).

Lemma 2.2.1. *Suppose that* $Y_t = \sum_{u=0}^{q} a_u\varepsilon_{t-u}$, $t \in \mathbb{Z}$, *is a* $MA(q)$*-process. Put* $\mu := \mathrm{E}(\varepsilon_0)$ *and* $\sigma^2 := \mathrm{Var}(\varepsilon_0)$. *Then we have*

(i) $\mathrm{E}(Y_t) = \mu\sum_{u=0}^{q} a_u$,

(ii) $\gamma(v) = \mathrm{Cov}(Y_v, Y_0) = \begin{cases} 0, & v > q, \\ \sigma^2 \sum\limits_{w=0}^{q-v} a_{v+w}a_w, & 0 \leq v \leq q, \end{cases}$

$\gamma(-v) = \gamma(v)$,

(iii) $\mathrm{Var}(Y_0) = \gamma(0) = \sigma^2\sum_{w=0}^{q} a_w^2$,

Beweis: Berechnung der kovarianzerzeugenden Funktion $G(z)$ eines linearen Filters, dann Cauchyprodukt. □

Beispiel 2.2.2. Beispiel der Autokorrelationsfunktion eines $MA(1)$-Prozesses. Die von a und $1/a$ stimmen überein und es gilt $|\rho(1)| \leq 1/2$.

2.2.1 Invertierbare Prozesse

Zusatzdefinition. Definition eines *invertierbaren* MA(q)-Prozesses.

$$\varepsilon_t = \sum_{u\geq 0} b_uY_{t-u}, \quad t \in \mathbb{Z},$$

2.2.2 Autoregressive Prozesse

Zusatzdefinition. Definition eines *autoregressiven Prozesses der Ordnung p* AR(p).

$$Y_t = a_1Y_{t-1} + \cdots + a_pY_{t-p} + \varepsilon_t, \quad A(z) = 1 - a_1z - \cdots - a_pz^p$$

2.2.3 Stationaritätsbedingung

Satz 2.2.3. *The AR(p)-equation with the given constants* $a_1, \ldots, a_p$ *and white noise* $(\varepsilon_t)_{t\in\mathbb{Z}}$ *has a stationary solution* $(Y_t)_{t\in\mathbb{Z}}$ *if all p roots of the equation* $1 - a_1z - a_2z^2 - \cdots -$

$a_p z^p = 0$ *are outside of the unit circle (*Stationaritätsbedingung eines AR(p) Prozesses*). In this case, the stationary solution is almost surely uniquely determined by*

$$Y_t := \sum_{u \geq 0} b_u \varepsilon_{t-u}, \quad t \in \mathbb{Z},$$

where $(b_u)_{u \geq 0}$ *is the absolutely summable inverse causal filter of* $c_0 = 1,\ c_u = -a_u,\ u = 1, \dots, p$ *and* $c_u = 0$ *elsewhere.*

BEWEIS: Absolut summierbarer inverser kausaler Filter (zu den ϵ) aus Satz, als $MA(\infty)$ dann Stationär. □

Beispiel 2.2.4. Ein $AR(1)$ Prozess. Man erhält die Darstellung von Y_t aus den ϵ aus vorherigem Satz und kann damit aus vorherigem Satz die Autocovarianzfunktion berechnen (generalisierter linearer Prozess). Autokorrelationsfunktion nimmt exponentiell ab.

2.2.4 Die Yule-Walker Gleichungen

Lemma 2.2.5. *Let* $Y_t = \sum_{u=1}^{p} a_u Y_{t-u} + \varepsilon_t$ *be an AR(p)-process, which satisfies the stationarity condition. Its autocorrelation function* ρ *then satisfies for* $s = 1, 2, \dots$ *the recursion*

$$\rho(s) = \sum_{u=1}^{p} a_u \rho(s-u),$$

known as Yule–Walker equations*.*

BEWEIS: Multiplikation von $Y_t - \mu = \sum_{u=1}^{p} a_u (Y_{t-u} - \mu) + \varepsilon_t - \nu$ mit $Y_{t-s} - \mu$ for $s > 0$ liefert die Darstellung, da Y_{t-s} und ϵ_t unkorreliert. □

Zusatzbemerkung. Mit $\rho(-s) = \rho(s)$ kann man die Yule-Walker-Gleichungen in Matrixschreibweise Schreiben und dann falls die Empirische Autokorrelationsmatrix invertierbar ist erhält man einen Schätzer durch $\hat{\boldsymbol{a}} := \boldsymbol{R}^{-1} \boldsymbol{r}$.

2.2.5 Der partielle Autokorrelationskoeffizient

$$\begin{aligned}
\boldsymbol{P}_k &:= \Big(\mathrm{Corr}(Y_i, Y_j) \Big)_{1 \leq i,j \leq k} \\
&= \begin{pmatrix}
1 & \rho(1) & \rho(2) & \dots & \rho(k-1) \\
\rho(1) & 1 & \rho(1) & & \rho(k-2) \\
\rho(2) & \rho(1) & 1 & & \rho(k-3) \\
\vdots & & & \ddots & \vdots \\
\rho(k-1) & \rho(k-2) & \rho(k-3) & \dots & 1
\end{pmatrix}
\end{aligned}$$

ist positiv semidifinit. Wir nehmen nun an positiv definit. Dann gilt

$$\boldsymbol{a}_k := \begin{pmatrix} a_{k1} \\ \vdots \\ a_{kk} \end{pmatrix} = \boldsymbol{P}_k^{-1} \begin{pmatrix} \rho(1) \\ \vdots \\ \rho(k) \end{pmatrix}.$$

Der Wert $\alpha(k) = a_{kk}$ heißt dann *partieller Autokorrelationskoeffizient mit lag k.*

Zusatzbemerkung. Nach den Yule-Walker Gleichungen ist $(a_1, \dots, a_p, 0, \dots, 0)$ eine Lösung für $k \geq p$ und demnach $\alpha(k) = 0$ for $k > p$.

Zusatzdefinition. Definition von *empirischer partieller Autokorrelationskoeffizient*

2.2.6 ARMA-Prozesse

Zusatzdefinition. Definition eines *autoregressiven moving average process der Ordnung p,q* mit den *charakteristischen Polynomen.*

$$Y_t = a_1 Y_{t-1} + a_2 Y_{t-2} + \dots + a_p Y_{t-p} + \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}.$$

$$A(z) := 1 - a_1 z - \dots - a_p z^p, \quad B(z) := 1 + b_1 z + \dots + b_q z^q$$

Zusatzdefinition. Definition von *Stationaritätsbedingung eines ARMA Prozesses*
Aus $Y_t - a_1 Y_{t-1} - \dots - a_p Y_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \dots + b_q \varepsilon_{t-q}$ erhält man, dass der rechte Teil stationär ist. Wenn $A(z)$ nur Nullstellen außerhalb des Einheitskreises hat kann man den linken Teil invertieren und erhält nach Rechnung $Y_t = \sum_{v \geq 0} \alpha_v \varepsilon_{t-v}$, dabei ist α_v der Produktfilter $\frac{1}{A(z)} \cdot B(z)$.

Zusatzdefinition. Definition von *Invertierbarer ARMA Prozesses.*

2.2.7 Die Autocovaranzfunktion eines ARMA-Prozesses

Zusatzbemerkung. Man kann die Koeffizienten α_v aus $A(z) \cdot E(z) = B(z)$ heraus berechnen.

Satz 2.2.6. *Suppose that* $Y_t = \sum_{u=1}^{p} a_u Y_{t-u} + \sum_{v=0}^{q} b_v \varepsilon_{t-v}$, $b_0 := 1$, *is an* $ARMA(p,q)$*-process, which satisfies the stationarity condition. Its autocovariance function* γ *then satisfies the recursion*

$$\gamma(s) - \sum_{u=1}^{p} a_u \gamma(s-u) = \sigma^2 \sum_{v=s}^{q} b_v \alpha_{v-s}, \quad 0 \leq s \leq q,$$
$$\gamma(s) - \sum_{u=1}^{p} a_u \gamma(s-u) = 0, \quad s \geq q+1,$$

where α_v, $v \geq 0$, *are the coefficients in the representation* $Y_t = \sum_{v \geq 0} \alpha_v \varepsilon_{t-v}$ *and* σ^2 *is the variance of* ε_0.

Beweis: Ähnlich wie für Yule-Walker Gleichungen, nur dass nicht alles verschwindet. Erwarungswert bilden, dann gleichmäßig verteilt abziehen und mit $(Y_{t-s-v} - \mu)$ multiplizieren und Erwartungswert bilden (unkorreliert wie in Yule-Walker). □

Beispiel 2.2.7. Ein $ARMA(1,1)$-Prozess.

2.2.8 ARIMA-Prozesse

Zusatzdefinition. Definition eines *autoregressiven integrated moving average der Ordnung p,d,q* und eines *saisonalen ARMA-Prozesses.*

$$\Delta^d Y_t = \sum_{u=1}^{p} a_u \Delta^d Y_{t-u} + \sum_{w=0}^{q} b_w \varepsilon_{t-w}, \quad t \in \mathbb{Z},$$

2.2.9 Cointegration

Zusatzdefinition. Definition von *cointegrierten* $I(1)$ Serien.

$$X_t = \mu + a_1 Y_t + a_2 Z_t, \quad t \in \mathbb{Z}, \quad \mathrm{I}(0)$$

Zusatzbeispiel. $Y_t = aW_t + \varepsilon_t, \quad Z_t = W_t + \delta_t, \quad t \in \mathbb{Z}$ mit wobei W $I(1)$ ist oder Zwei Random Walks (Hund und Betrunkener, *error correction mechanism.*

Zusatzbemerkung. Vorgehensweise:

(a) Determine that the two series are $I(1)$ by standard unit root tests such as Dickey–Fuller or augmented Dickey–Fuller.

(b) Compute $\hat{\varepsilon}_t = Y_t - \hat{\beta}_0 - \hat{\beta}_1 Z_t$ using ordinary least squares. $\sum_{t=1}^{n} \left(Y_t - \hat{\beta}_0 - \hat{\beta}_1 Z_t\right)^2$.

(c) Examine $\hat{\varepsilon}_t$ for stationarity, using for example the Phillips–Ouliaris test.

2.2.10 Dickey-Fuller-Test

Einfacher Fall der drei Modelle. $\Delta Y_t = \gamma Y_{t-1} + \varepsilon_t$. Dies ist Darstellbar als $AR(1)$ Prozess mit $a_1 = \gamma + 1$. Die Nullhypothese ist dann $\gamma = 0$ und der AR Prozess wäre nicht stationär ($a_1 = 1$)

Man schätzt nun $a_1 = \gamma + 1$ by $\hat{a}_1$ durch eine Regression und testet mit einer Dicky-Fuller-Teststatistik ($x := n\hat{\gamma} := n(\hat{a}_1 - 1)$.

2.2.11 Phillips-Ouliaris Test

Nullyhpothese: Keine Cointegration. Berechnung zweier Testgrößen RHO und TAU.

2.2.12 ARCH- und GARCH-Prozesse

Zusatzdefinition. Definition eines *autoregressiven conditional heteroscedastischen Prozesses* ARCH.

Man modelliert die Volatilität einer Zeitreihe Y_t, die Varianz hängt von den vorherigen Beobachtungen ab. Z Standardnormal oder T-Verteilt.

$$Y_t = \sigma_t Z_t, \quad t \in \mathbb{Z}, \quad \sigma_t^2 = a_0 + \sum_{j=1}^{p} a_j Y_{t-j}^2, \quad t \in \mathbb{Z}$$

Zusatzbemerkung. Es gilt: $E(Y_t) = 0$ und $\sigma^2 = a_0 + \sigma^2 \sum_{j=1}^{p} a_j$, d.h. $\sum_{j=1}^{p} a_j < 1$ ist eine notwendige Bedingung für die Stationarität. Wegen $\mathrm{E}(Y_s Y_t) = \mathrm{E}(\sigma_s Z_s \sigma_t Z_t) = \mathrm{E}(\sigma_s Z_s \sigma_t)\,\mathrm{E}(Z_t) = 0$ $s < t$ sind Y_t und Y_s unkorelliert. Sie sind aber nicht unabhängig, da Y_s die Varianz von Y_t beeinflusst.

Lemma 2.2.8. *Let (Y_t) be a stationary and causal ARCH(p)-process with constants $a_0, a_1, \ldots, a_p$. If the process of squared random variables (Y_t^2) is a stationary one, then it is an AR(p)-process:*

$$Y_t^2 = a_1 Y_{t-1}^2 + \cdots + a_p Y_{t-p}^2 + \varepsilon_t,$$

where (ε_t) is a white noise with $\mathrm{E}(\varepsilon_t) = a_0$, $t \in \mathbb{Z}$.

Beweis: Nach ϵ_t auflösen und Zeigen, dass es ein weißes Rauschen ist (Erwartungswert konstant, Varianz konstant, Unkorreliert). □

Zusatzbemerkung. Die Stationarität und die Ordnung des Prozesses (Y_t^2) findet man wie bei einem AR-Prozess. Daraus ist dann auch a_0 schätzbar. Bedingt den vorherigen Beobachtungen ist Y_t auch normalverteilt, wenn es Z_t ist. Dann könnte man die Koeffizienten über einen Maximum Likelihoodansatz schätzen.

Zusatzdefinition. Definition eines *generalisierten ARCH-Prozesses.*

$$\sigma_t^2 = a_0 + \sum_{j=1}^{p} a_j Y_{t-j}^2 + \sum_{k=1}^{q} b_k \sigma_{t-k}^2$$

2.3 Das Box-Jenkins Programm

Wir arbeiten mit varianz-stabilisierten, Trend und saisonal bereinigten Daten $y_1, \ldots, y_n$. Anpassung an ARMA (endlich) und nicht unendlichen MA Prozess.

(a) Auswahl der Ordnungen (Identification)

(b) Schätzung der Koeffizienten (Estimation)

(c) Diagnostik Check

(d) Vorhersage (Forecasting)

2.3.1 Auswahl der Ordnungen

Für reine *MA* und *AR* Prozesse nutzt das ACF oder PACF Diagramm. Bei ARMA-Prozessen wählt man Paare (p, q) so, dass entweder das *Akaike Informationskriterium* (AIC), das *Bayesian Informationkriterium* (BIC) oder das *Hannan-Quinn-Kriterium* minimiert wird. Diese basieren jeweils auf $\hat{\sigma}_{p,q}^2$ von ε_0.

2.3.2 Schätzung der Koeffizienten – Gaussian Model: Maximum Likelihood Estimator

Wir nehmen an, dass $(Y_1, \ldots, Y_n)$ einer n-dimensionalen Normalverteilung unterliegen.

$$P\{Y_i \leq s_i,\ i = 1, \ldots, n\} = \int_{-\infty}^{s_1} \cdots \int_{-\infty}^{s_n} \boldsymbol{\varphi}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(x_1, \ldots, x_n)\, dx_n \ldots dx_1$$

$$\begin{aligned}
&\boldsymbol{\varphi}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(x_1, \ldots, x_n) \\
&\quad = \frac{1}{(2\pi)^{n/2}(\det \boldsymbol{\Sigma})^{1/2}} \exp\Big(-\frac{1}{2}((x_1, \ldots, x_n) - \boldsymbol{\mu}^{\boldsymbol{T}})\boldsymbol{\Sigma}^{-1}((x_1, \ldots, x_n) - \boldsymbol{\mu}^{\boldsymbol{T}})^T \Big)
\end{aligned}$$

mit $\boldsymbol{\mu} = (\mu, \ldots, \mu)^T \in \mathbb{R}^n$ und $\boldsymbol{\Sigma} = (\gamma(i-j))_{1 \leq i,j \leq n}$.

Zusatzbemerkung. Man maximiert die Dichte $\boldsymbol{\varphi}_{\boldsymbol{\mu},\boldsymbol{\Sigma}}(y_1, \ldots, y_n)$. Erfüllt der Prozess (Y_t) die Stationaritätsbedingung, so ist $Y_t = \sum_{v \geq 0} \alpha_v \varepsilon_{t-v},\ t \in \mathbb{Z}$, wobei die α_v nur von den MA und AR Komponenten Abhängen. Damit ist $\gamma(s) = \mathrm{Cov}(Y_0, Y_s) = \sigma^2 \sum_{v \geq 0} \alpha_v \alpha_{s+v}$. Die Dichte hängt also nur von $\boldsymbol{\vartheta} := (\sigma^2, \mu, a_1, \ldots, a_p, b_1, \ldots, b_q)$ ab. Auf $\boldsymbol{\vartheta}$ kann man die Maximum-Likelihood-Methode anwenden (evtl. Log-Likelihood) bzgl. der Dichte.

2.3.3 Schätzung der Koeffizienten – Nichtparametrischer Ansatz: Kleinste Quadrate

Sei $E(\epsilon_t) = 0$, dann ist $\hat{Y}_t = a_1 Y_{t-1} + \cdots + a_p Y_{t-p} + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$ eine logische Vorhersage von Y_t in einem $ARMA(p,q)$-Prozess mit Residuum $Y_t - \hat{Y}_t = \varepsilon_t$. Sei nun $\hat{\varepsilon}_t$ ein Schätzer abhängig von den ARMA Koeffizienten. Dann ist

$$\begin{aligned}
&S^2(a_1, \ldots, a_p,\, b_1, \ldots, b_q) \\
&\quad = \sum_{t=-\infty}^{n} \hat{\varepsilon}_t^2 = \sum_{t=-\infty}^{n} (y_t - a_1 y_{t-1} - \cdots - a_p y_{t-p} - b_1 \hat{\varepsilon}_{t-1} - \cdots - b_q \hat{\varepsilon}_{t-q})^2
\end{aligned}$$

die *Residuuenquadratsumme.* Im kleinsten Quadrateansatz minimiert man diese.

Zusatzbemerkung. Da man für $t \leq 0$ keine Werte zur Verfügung hat sagt man diese durch Null vorher und bestimmt den Prozess iterativ, bis man alle Werte $\hat{\varepsilon}_{\max\{p+q\}}$ hat.

2.3.4 Diagnostik Check

Man verwendet einen Portmanteau-Test, der überprüft, ob die Residuen $\hat{\varepsilon}_t$ sich wie ein weißes Rauschen verhalten und dazu überprüft, ob die Summe der empirischen Autokorrelationskoeffizienten klein genug ist. Diese sind dann approximativ χ^2 verteilt und man erhält einen p-Wert.

2.3.5 Vorhersage

Zusatzdefinition. Definition der *besten h-Schritt Vorhersage* mit minimalem Mittleren quadratischem Fehler.

$$\mathrm{E}\left(\left(Y_{n+h} - \sum_{u=0}^{n-1} c_u^* Y_{n-u}\right)^2\right) = \min_{c_0,\dots,c_{n-1}\in\mathbb{R}} \mathrm{E}\left(\left(Y_{n+h} - \sum_{u=0}^{n-1} c_u Y_{n-u}\right)^2\right).$$

Lemma 2.3.1. *Let (Y_t) be an arbitrary stochastic process with finite second moments. If the weights $c_0^*, \dots, c_{n-1}^*$ have the property that*

$$\mathrm{E}\left(Y_i\left(Y_{n+h} - \sum_{u=0}^{n-1} c_u^* Y_{n-u}\right)\right) = 0, \quad i = 1, \dots, n,$$

then $\hat{Y}_{n+h} := \sum_{u=0}^{n-1} c_u^* Y_{n-u}$ is a best h-step forecast of Y_{n+h}.

Beweis: $\mathrm{E}((Y_{n+h} - \tilde{Y}_{n+h})^2) = \dots \geq \mathrm{E}((Y_{n+h} - \hat{Y}_{n+h})^2)$. □

Zusatzbemerkung. Die Gleichungen sind für einen stationären Prozess mit Erwartungswert Null vom Yule-Walker Typ und somit entspricht c_{n-1}^* den Koeffizienten $\alpha(n)$ für $h = 1$ und ist somit Teil der Einschrittvorhersage.

Beispiel 2.3.2. Vorhersage eines MA(1)-Prozesses. (Ab 2. Schritt Vorhersage durch Null)

Satz 2.3.3. *Suppose that $Y_t = \sum_{u=1}^{p} a_u Y_{t-u} + \varepsilon_t$, $t \in \mathbb{Z}$, is a stationary AR(p)-process, which satisfies the stationarity condition and has zero mean $\mathrm{E}(Y_0) = 0$. Let $n \geq p$. The best one-step forecast is*

$$\hat{Y}_{n+1} = a_1 Y_n + a_2 Y_{n-1} + \dots + a_p Y_{n+1-p}$$

and the best two-step forecast is

$$\hat{Y}_{n+2} = a_1 \hat{Y}_{n+1} + a_2 Y_n + \dots + a_p Y_{n+2-p}.$$

The best h-step forecast for arbitrary $h \geq 2$ is recursively given by

$$\hat{Y}_{n+h} = a_1 \hat{Y}_{n+h-1} + \dots + a_{h-1} \hat{Y}_{n+1} + a_h Y_n + \dots + a_p Y_{n+h-p}.$$

Beweis: Da der Prozess stationär ist er als $Y_t = \sum_{u\geq 0} b_u \varepsilon_{t-u}$ darstellbar. Damit ist $\mathrm{E}((Y_{n+1} - \hat{Y}_{n+1})Y_i) = \mathrm{E}(\varepsilon_{n+1} Y_i) = 0$ und aus dem Lemma folgt die Beh. für $h = 1$. Rest rekursiv. □

Korollar 2.3.4. *Suppose that $Y_t = \sum_{u=1}^{p} a_u Y_{t-u} + \varepsilon_t + \sum_{v=1}^{q} b_v \varepsilon_{t-v}$, $t \in \mathbb{Z}$, is an ARMA(p, q)-process, which satisfies the stationarity condition and has zero mean, precisely $\mathrm{E}(\varepsilon_0) = 0$. Suppose that $n + q - p \geq 0$. The best h-step forecast of Y_{n+h} for $h > q$ satisfies the recursion*

$$\hat{Y}_{n+h} = \sum_{u=1}^{p} a_u \hat{Y}_{n+h-u}.$$

Beispiel 2.3.5. Vorhersage eines $ARMA(1,1)$ Prozesses. Alle ϵ_n berechnen um $\widehat{Y}_{n+1}$ berechnen zu können aus ARMA Gleichung und dann nach vorherigem Satz vorgehen.

2.4 Zustandsraum Modelle – State-Space Models

Zusatzdefinition. Definition eines *Zustandsraummodells.*
Es gibt einen nichtbeobachtbaren Prozess (X_t) und einen beobachtbaren Prozess (Y_t).

$$\boldsymbol{X}_{t+1} = \boldsymbol{A}_t \boldsymbol{X}_t + \boldsymbol{B}_t \boldsymbol{\varepsilon}_{t+1} \in \mathbb{R}^k, \quad \boldsymbol{Y}_t = \boldsymbol{C}_t \boldsymbol{X}_t + \boldsymbol{\eta}_t \in \mathbb{R}^m.$$

(ε_t) und (η_t) unkorrelierte weiße Rauschen mit $\boldsymbol{\mu} = \boldsymbol{0}$ und bekannten $\mathrm{Cov}(\boldsymbol{\varepsilon}_t) =: \boldsymbol{K}_t$ und $\mathrm{Cov}(\boldsymbol{\eta}_t) =: \boldsymbol{Q}_t$.

Beispiel 2.4.1. Ein Random Walk, ein AR(p), ein MA(q) und ein ARMA(p,q) Prozess lassen sich als Zustandsraummodelle angeben.

2.4.1 Kalman-Filter

Zusatzdefinition. Definition der *besten linearen Vorhersage* von X aus Y.

$$\hat{\boldsymbol{X}}_t := \boldsymbol{D}_1 \boldsymbol{Y}_1 + \cdots + \boldsymbol{D}_t \boldsymbol{Y}_t, \text{ minimiert } \mathrm{E}((\boldsymbol{X}_t - \hat{\boldsymbol{X}}_t)^T (\boldsymbol{X}_t - \hat{\boldsymbol{X}}_t))$$

Lemma 2.4.2. *If the estimate $\hat{\boldsymbol{X}}_t$ satisfies*

$$\mathrm{E}((\boldsymbol{X}_t - \hat{\boldsymbol{X}}_t)\boldsymbol{Y}_s^T) = \boldsymbol{0}, \quad 1 \leq s \leq t,$$

then it minimizes the mean squared error.

Beweis: $\mathrm{E}((\boldsymbol{X}_t - \boldsymbol{X}'_t)^T (\boldsymbol{X}_t - \boldsymbol{X}'_t)) = \cdots \geq \mathrm{E}((\boldsymbol{X}_t - \hat{\boldsymbol{X}}_t)^T (\boldsymbol{X}_t - \hat{\boldsymbol{X}}_t))$ □

$\tilde{\boldsymbol{X}}_t := \boldsymbol{A}_{t-1} \hat{\boldsymbol{X}}_{t-1}$ ist dann die beste Vorhersage von $\boldsymbol{X}_t$ durch $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{t-1}$ und $\tilde{\boldsymbol{Y}}_t := \boldsymbol{C}_t \tilde{\boldsymbol{X}}_t$ die beste Vorhersage von $\boldsymbol{Y}_t$ durch $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_{t-1}$ (nach vorherigem Lemma).

Mit

$$\boldsymbol{\Delta}_t := \mathrm{E}((\boldsymbol{X}_t - \hat{\boldsymbol{X}}_t)(\boldsymbol{X}_t - \hat{\boldsymbol{X}}_t)^T) \quad \text{und} \quad \tilde{\boldsymbol{\Delta}}_t := \mathrm{E}((\boldsymbol{X}_t - \tilde{\boldsymbol{X}}_t)(\boldsymbol{X}_t - \tilde{\boldsymbol{X}}_t)^T)$$

gilt dann

$$\tilde{\boldsymbol{\Delta}}_t = \boldsymbol{A}_{t-1} \boldsymbol{\Delta}_{t-1} \boldsymbol{A}_{t-1}^T + \boldsymbol{B}_{t-1} \boldsymbol{K}_t \boldsymbol{B}_{t-1}^T \quad \text{und} \quad \mathrm{E}((\boldsymbol{Y}_t - \tilde{\boldsymbol{Y}}_t)(\boldsymbol{Y}_t - \tilde{\boldsymbol{Y}}_t)^T) = \boldsymbol{C}_t \tilde{\boldsymbol{\Delta}}_t \boldsymbol{C}_t^T + \boldsymbol{R}_t.$$

Zusatzdefinition. Definition und Arbeitsweise des *Kalman gain.*
Angenommen wir beobachten nun auch $\boldsymbol{Y}_t$, dann daten wir $\boldsymbol{X}$ wie folgt up:

$$\tilde{\boldsymbol{X}}_t + \boldsymbol{K}_t (\boldsymbol{Y}_t - \tilde{\boldsymbol{Y}}_t) = \hat{\boldsymbol{X}}_t$$

Lemma 2.4.3. *The matrix $\boldsymbol{K}_t$ is a solution of the equation*

$$\boldsymbol{K}_t (\boldsymbol{C}_t \tilde{\boldsymbol{\Delta}}_t \boldsymbol{C}_t^T + \boldsymbol{R}_t) = \tilde{\boldsymbol{\Delta}}_t \boldsymbol{C}_t^T.$$

Beweis: Es muss gelten (Lemma) $\boldsymbol{0} = \mathrm{E}((\boldsymbol{X}_t - \hat{\boldsymbol{X}}_t)\boldsymbol{Y}_s^T), s \leq t$. Insb. $s = t$ liefert $\boldsymbol{K}_t$. □

Zusatzbemerkung. Falls die Matrix invertierbar ist erhält man und zusammen mit einer Rechnung:

$$\boldsymbol{K}_t := \tilde{\boldsymbol{\Delta}}_t \boldsymbol{C}_t^T (\boldsymbol{C}_t \tilde{\boldsymbol{\Delta}}_t \boldsymbol{C}_t^T + \boldsymbol{R}_t)^{-1} \quad \text{und} \quad \boldsymbol{\Delta}_t = \tilde{\boldsymbol{\Delta}}_t - \boldsymbol{K}_t \boldsymbol{C}_t \tilde{\boldsymbol{\Delta}}_t$$

Zusatzdefinition. Definition des *Kalman filters*.

prediction step

$$\begin{aligned}
\tilde{\boldsymbol{X}}_t &= \boldsymbol{A}_{t-1} \hat{\boldsymbol{X}}_{t-1}, \\
\tilde{\boldsymbol{Y}}_t &= \boldsymbol{C}_t \tilde{\boldsymbol{X}}_t, \\
\tilde{\boldsymbol{\Delta}}_t &= \boldsymbol{A}_{t-1} \boldsymbol{\Delta}_{t-1} \boldsymbol{A}_{t-1}^T + \boldsymbol{B}_{t-1} \boldsymbol{Q}_t \boldsymbol{B}_{t-1}^T.
\end{aligned}$$

updating step

$$\begin{aligned}
\boldsymbol{K}_t &= \tilde{\boldsymbol{\Delta}}_t \boldsymbol{C}_t^T (\boldsymbol{C}_t \tilde{\boldsymbol{\Delta}}_t \boldsymbol{C}_t^T + \boldsymbol{R}_t)^{-1}, \\
\hat{\boldsymbol{X}}_t &= \tilde{\boldsymbol{X}}_t + \boldsymbol{K}_t (\boldsymbol{Y}_t - \tilde{\boldsymbol{Y}}_t), \\
\boldsymbol{\Delta}_t &= \tilde{\boldsymbol{\Delta}}_t - \boldsymbol{K}_t \boldsymbol{C}_t \tilde{\boldsymbol{\Delta}}_t.
\end{aligned}$$

Zusatzbemerkung. Ein Problem ist die Wahl der Startwerte. Die Schätzwerte werden durch diese aber nur gering beeinflusst. Falls das Zustandsraummodell durch eine parametrische Verteilung beschriebene werden kann, dann kann man die Matrizen unter gewissen Bedingungen durch einen Maximum-Likelihood Ansatz schätzen. Durch Iteration der 1-Schritt Vorhersage erhält man h-Schritt Vorhersagen:

$$\tilde{\boldsymbol{X}}_{t+h} := \boldsymbol{A}_{t+h-1} \tilde{\boldsymbol{X}}_{t+h-1} \quad \text{und} \quad \tilde{\boldsymbol{Y}}_{t+h} := \boldsymbol{C}_{t+h} \tilde{\boldsymbol{X}}_{t+h}$$

Beispiel 2.4.4. Kalman-Filter auf ein weißes Rauschen plus einer Konstante und auf die Airlinedaten.

3 The Frequency Domain

Eigentlich müssen die Daten $y_1, \ldots, y_n$ in diesem Kapitel nicht stationär sein, aber die empirische Autokovarianzfunktion hat dann keine Deutung.

3.1 Kleinste Quadrate-Ansatz mit bekannten Frequenzen

Zusatzdefinition. Definition einer *periodischen Funktion*, *Periode*, *Fundamentalperiode*, *Frequenz*, *Fundamentalfrequenz*, *Harmonische Welle der Länge r (harmonic wave)*.

$$m_\lambda(t) := A\cos(2\pi\lambda t) + B\sin(2\pi\lambda t), \qquad A, B \in \mathbb{R},\ \lambda > 0, g(t) = \mu + \sum_{k=1}^{r} m_{\lambda_k}(t)$$

Beispiel 3.1.1. Sterndaten. Anpassung mit bekannten Frequenzen aber unbekannten Koeffizienten.

Zuerst passen wir nur eine harmonische Komponente an die mittelwertangepassten Daten an.

$$m(t) = Am_1(t) + Bm_2(t), \quad m_1(t) := \cos(2\pi\lambda t), \quad m_2(t) = \sin(2\pi\lambda t).$$

A und B schätzen wir dann über kleinste Quadrate.

$$R(A,B) := \sum_{t=1}^{n} (y_t - \bar{y} - m(t))^2.$$

Falls eine Bedingung erfüllt ist können wir die Normal Equations (partiellen Ableitungen) nach A und B auflösen mit den *empirische Kreuz-Kovarianzen* von $(y_t)_{1\le t\le n}$ und $(\sin/\cos(2\pi\lambda t))_{1\le t\le n}$:

$$A = A(\lambda) = n\frac{c_{22}C(\lambda) - c_{12}S(\lambda)}{c_{11}c_{22} - c_{12}c_{21}}, \quad C(\lambda) := \frac{1}{n}\sum_{t=1}^{n}(y_t - \bar{y})\cos(2\pi\lambda t)$$

$$B = B(\lambda) = n\frac{c_{21}C(\lambda) - c_{11}S(\lambda)}{c_{12}c_{21} - c_{11}c_{22}}, \quad S(\lambda) := \frac{1}{n}\sum_{t=1}^{n}(y_t - \bar{y})\sin(2\pi\lambda t)$$

3.1.1 Harmonische Wellen mit Fourier-Frequenzen

Zusatzbemerkung. Die Lösungen von A und B werden besonders einfach, wenn man *Fourier-Frequenzen* wählt $\lambda = k/n,\ k = 0, 1, 2, \ldots, [n/2]$. $(A = 2C(\lambda),\ B = 2S(\lambda)$.

Lemma 3.1.2. *For arbitrary $0 \le k, m \le [n/2]$ we have*

$$\sum_{t=1}^{n} \cos\left(2\pi\frac{k}{n}t\right)\cos\left(2\pi\frac{m}{n}t\right) = \begin{cases} n, & k = m = 0 \text{ or } n/2, \quad \text{if } n \text{ is even} \\ n/2, & k = m \neq 0 \text{ and } \neq n/2, \quad \text{if } n \text{ is even} \\ 0, & k \neq m \end{cases}$$

$$\sum_{t=1}^{n} \sin\left(2\pi\frac{k}{n}t\right)\sin\left(2\pi\frac{m}{n}t\right) = \begin{cases} 0, & k = m = 0 \text{ or } n/2, \quad \text{if } n \text{ is even} \\ n/2, & k = m \neq 0 \text{ and } \neq n/2, \quad \text{if } n \text{ is even} \\ 0, & k \neq m \end{cases}$$

$$\sum_{t=1}^{n} \cos\left(2\pi\frac{k}{n}t\right)\sin\left(2\pi\frac{m}{n}t\right) = 0.$$

Zusatzsatz. *Das Lemma impliziert, dass die $2[n/2] + 1$ Vektoren $(\sin(2\pi(k/n)t))_{1\le t\le n}$ $k = 1, \ldots, [n/2]$ und $(\cos(2\pi(k/n)t))_{1\le t\le n}$, $k = 0, \ldots, [n/2]$ orthogonal stehen und unabhängig sind, also den $\mathbb{R}^n$ aufspannen. Also existieren für jeden Datensatz $y_1, \ldots, y_n$*

$$y_t = \sum_{k=0}^{[n/2]}\left(A_k\cos\left(2\pi\frac{k}{n}t\right) + B_k\sin\left(2\pi\frac{k}{n}t\right)\right), \qquad t = 1, \ldots, n.$$

diese Koeffizienten minimieren die Quadratsumme (sie ist Null). Durch die Normal-Equations erhält man die Lösung:

$$A_k = \begin{cases} \frac{2}{n}\sum_{t=1}^{n} y_t\cos\left(2\pi\frac{k}{n}t\right), & k = 1, \ldots, [(n-1)/2] \\ \frac{1}{n}\sum_{t=1}^{n} y_t\cos\left(2\pi\frac{k}{n}t\right), & k = 0 \text{ and } k = n/2, \text{ if } n \text{ is even} \end{cases}$$

$$B_k = \frac{2}{n}\sum_{t=1}^{n} y_t\sin\left(2\pi\frac{k}{n}t\right), \quad k = 1, \ldots, [(n-1)/2].$$

Zusatzbemerkung. Man arbeitet auch gerne mit der äquivalenten Formulierung

$$y_t = \frac{\tilde{A}_0}{2} + \sum_{k=1}^{[n/2]} \left(A_k \cos\left(2\pi \frac{k}{n} t\right) + B_k \sin\left(2\pi \frac{k}{n} t\right)\right), \qquad t = 1, \ldots, n$$

Bis auf den Faktor 2 stimmen diese Koeffizienten mit den empirischen Kovarianzen $C(k/n)$ und $S(k/n)$ überein, wegen

$$\sum_{t=1}^{n} \cos\left(2\pi \frac{k}{n} t\right) = \sum_{t=1}^{n} \sin\left(2\pi \frac{k}{n} t\right) = 0, \quad k = 1, \ldots, [n/2]$$

3.1.2 Das Periodogramm

Zusatzbemerkung. Zuvor wurde an eine Zeitreihe exakt durch Fourier-Frequenzen beschrieben, oft genügen aber jedoch sehr viel wenigere (die wieder in der Nähe von Fourierfrequenzen liegen).

Zusatzbemerkung. Durch Ausrechnen erhält man $\sum_{t=1}^{n}(y_t - \bar{y})^2 = \sum_{k=1}^{[n/2]} \frac{n}{2}\big(A_k^2 + B_k^2\big)$ also entspricht der rechte Teil $(n/2)(A_k^2 + B_k^2) = 2n(C^2(k/n) + S^2(k/n))$ dem Anteil der Gesamtvarianz dieser Fourierfrequenz *Intensiät der Frequenz* k/n.

Zusatzdefinition. Definition des *Periodogramms.*

$$\begin{aligned} I(\lambda) &= n\big(C(\lambda)^2 + S(\lambda)^2\big) \qquad I(k/n) = \frac{n}{4}\big(A_k^2 + B_k^2\big), \qquad k = 1, \ldots, [(n-1)/2]. \\ &= \frac{1}{n}\left(\Big(\sum_{t=1}^{n}(y_t - \bar{y})\cos(2\pi\lambda t)\Big)^2 + \Big(\sum_{t=1}^{n}(y_t - \bar{y})\sin(2\pi\lambda t)\Big)^2\right). \end{aligned}$$

Satz 3.1.3. *We have for the periodogram*

(a) $I(0) = 0$,

(b) I *is an even function, i.e.,* $I(\lambda) = I(-\lambda)$ *for any* $\lambda \in \mathbb{R}$,

(c) I *has the period 1.*

3.1.3 Fourier Transformation

Zusatzbemerkung. Das Periodogramm ist eine Funktion in Abhängigkeit von $D(\lambda)$, da gilt $I(\lambda) = n|D(\lambda)|^2$. Dabei enthält $D(\lambda)$ aber auch alle Information von C und S.

$$D(\lambda) := C(\lambda) - iS(\lambda) = \frac{1}{n}\sum_{t=1}^{n}(y_t - \bar{y})e^{-i2\pi\lambda t}.$$

Zusatzdefinition. Definition einer *Fouriertransformation.*

$$f_a(\lambda) = \sum_{t\in\mathbb{Z}} a_t e^{-i2\pi\lambda t}, \qquad \lambda \in \mathbb{R}, \quad (a_t) \text{ abs. summierbar}$$

Satz 3.1.4. *We have*

(a) $f_a(0) = \sum_{t\in\mathbb{Z}} a_t$,

(b) $f_a(-\lambda)$ *and* $f_a(\lambda)$ *are conjugate complex numbers i.e.,* $f_a(-\lambda) = \overline{f_a(\lambda)}$,

(c) f_a *has the period 1.*

3.1.4 Autocovarianzfunktion und Periodogramm

Satz 3.1.5 (Fouriertransformation der emp. Autocovarianzfunktion)**.** *Denote by c the empirical autocovariance function of $y_1, \dots, y_n$, i.e., $c(k) = n^{-1}\sum_{j=1}^{n-k}(y_j - \bar{y})(y_{j+k} - \bar{y})$, $k = 0, \dots, n-1$, where $\bar{y} := n^{-1}\sum_{j=1}^{n} y_j$. Then we have with $c(-k) := c(k)$*

$$I(\lambda) = c(0) + 2\sum_{k=1}^{n-1} c(k)\cos(2\pi\lambda k)$$
$$= \sum_{k=-(n-1)}^{n-1} c(k)e^{-i2\pi\lambda k}.$$

Beweis: Ausschreiben von $I(\lambda)$ dann Additionsformeln und Cos in positiv und negative exp-Funktion aufspalten. □

3.1.5 Inverse Fouriertransformation

Satz 3.1.6 (Inverse Fouriertransformation des Periodogramms)**.** *The periodogram*

$$I(\lambda) = \sum_{k=-(n-1)}^{n-1} c(k)e^{-i2\pi\lambda k}, \qquad \lambda \in \mathbb{R},$$

satisfies the inverse formula

$$c(k) = \int_0^1 I(\lambda)e^{i2\pi\lambda k}\,d\lambda, \qquad |k| \le n-1.$$

Zusatzbemerkung. Für $k = 0$ erhält man eine alternative Darstellung der Strichprobenvarianz. Das Periodogramm zeigt also wie die Gesamtvarianz sich über die Frequenzen verteilt.

Satz 3.1.7. *For an absolutely summable sequence $a := (a_t)_{t\in\mathbb{Z}}$ with Fourier transform $f_a(\lambda) = \sum_{t\in\mathbb{Z}} a_t e^{-i2\pi\lambda t}$, $\lambda \in \mathbb{R}$, we have*

$$a_t = \int_0^1 f_a(\lambda)e^{i2\pi\lambda t}\,d\lambda, \qquad t \in \mathbb{Z}.$$

Beweis: Da $\exp(i)$ den Betrag eins hat liefert das Prinzip der Dominanten Konvergenz, dass man Summe und Integral vertauschen kann $\Rightarrow$ Kronecker-Funktion. □

Zusatzbemerkung. Eine Zeitreihe wird also auch eindeutig durch seine Fouriertransformierte bestimmt. Folglich ist die Analyse im Frequenzbereich gleichwertig zum Zeitbereich.

3.1.6 Aliasing

Zusatzdefinition. Definition von *Aliasing* bei kontinuierlichen Prozessen $(Z_t)_{t\in\mathbb{R}}$ und der *Nyquist Frequenz* $\lambda^* \leq 1/(2\Delta)$. Nur diese können bei Intervallbreite Δ unterschieden werden.

4 Spektrum eines stationären Prozesses

Zusatzdefinition. Definition der *Spektraldiche* oder des *Spektrums*.

$$f(\lambda) := \sum_{t\in\mathbb{Z}} \gamma(t)e^{-i2\pi\lambda t} = \gamma(0) + 2\sum_{t\in\mathbb{N}} \gamma(t)\cos(2\pi\lambda t), \quad \lambda \in \mathbb{R},$$

Zusatzbemerkung. Aus der allgemeinen inversen Fouriertransformation folgt sofort $\gamma(t) = \int_0^1 f(\lambda)e^{i2\pi\lambda t}\,d\lambda = \int_0^1 f(\lambda)\cos(2\pi\lambda t)\,d\lambda$ und für $t = 0$ die Varianzdarstellung.

4.1 Charakterisierung der Autokovarianzfunktion

Satz 4.1.1. *A symmetric function $K : \mathbb{Z} \to \mathbb{R}$ is the autocovariance function of a stationary process $(Y_t)_{t\in\mathbb{Z}}$ iff K is a positive semidefinite function, i.e., $K(-n) = K(n)$ and*

$$\sum_{1\leq r,s\leq n} x_r K(r-s)x_s \geq 0$$

for arbitrary $n \in \mathbb{N}$ and $x_1, \ldots, x_n \in \mathbb{R}$.

Beweis: Notwendigkeit folgt aus $0 \leq \mathrm{E}\left((\sum_{t=1}^n x_t Y_t)^2\right)$. Suffizienz lang. □

4.1.1 Sprektralverteilungsfunktion und Spektraldichte

Satz 4.1.2 (Herglotz'sche Satz)**.** *A symmetric function $\gamma : \mathbb{Z} \to \mathbb{R}$ is positive semidefinite iff it can be represented as an integral*

$$\gamma(h) = \int_0^1 e^{i2\pi\lambda h}\,dF(\lambda) = \int_0^1 \cos(2\pi\lambda h)\,dF(\lambda), \quad h \in \mathbb{Z},$$

where F is a real valued measure generating function on $[0, 1]$ with $F(0) = 0$. The function F is uniquely determined.

Beweis: (lang) □

Zusatzdefinition. Definition der *Spektralverteilungsfunktion von γ* und der *Spektraldichte* (existiert falls $\sum_{h\geq 0} |\gamma(h)| < \infty$).

Beispiel 4.1.3. Für ein weißes Rauschen gilt und ebenso

$$\gamma(h) = \begin{cases} \sigma^2, & h = 0 \\ 0, & h \in \mathbb{Z} \setminus \{0\}. \end{cases}, \quad \int_0^1 \sigma^2 e^{i2\pi\lambda h}\, d\lambda = \begin{cases} \sigma^2, & h = 0 \\ 0, & h \in \mathbb{Z} \setminus \{0\}, \end{cases}$$

Es hat also die konstante Spektraldichte $f(\lambda) = \sigma^2$, $0 \le \lambda \le 1$.

Korollar 4.1.4. *A symmetric function $\gamma : \mathbb{Z} \to \mathbb{R}$ is the autocovariance function of a stationary process $(Y_t)_{t\in\mathbb{Z}}$, iff it satisfies one of the following two (equivalent) conditions:*

(a) $\gamma(h) = \int_0^1 e^{i2\pi\lambda h}\, dF(\lambda)$, $h \in \mathbb{Z}$, *where F is a measure generating function on* $[0, 1]$ *with* $F(0) = 0$.

(b) $\sum_{1\le r,s\le n} x_r \gamma(r - s) x_s \ge 0$ *for each $n \in \mathbb{N}$ and $x_1, \dots, x_n \in \mathbb{R}$.*

Korollar 4.1.5. *A symmetric function $\gamma : \mathbb{Z} \to \mathbb{R}$ with $\sum_{t\in\mathbb{Z}} |\gamma(t)| < \infty$ is the autocovariance function of a stationary process iff*

$$f(\lambda) := \sum_{t\in\mathbb{Z}} \gamma(t) e^{-i2\pi\lambda t} \ge 0, \quad \lambda \in [0, 1].$$

The function f is in this case the spectral density of γ.

Beweis: Definition einer Hilfsfunktion $f_N(\lambda)$, die in f übergeht. Spektraldichte folgt aus inverser Fouriertransformation. Andere Richtung durch inverse Fouriertrans $\gamma(t) = \int_0^1 f(\lambda) e^{i2\pi\lambda t}\ d\lambda = \int_0^1 e^{i2\pi\lambda t}\, dF(\lambda)$, mit $F(\lambda) = \int_0^\lambda f(x)\, dx$, $0 \le \lambda \le 1$, also der Satz anwendbar. □

Beispiel 4.1.6. Die Funktion $\gamma(h) = \begin{cases} 1, & \text{if } h = 0 \\ \rho, & \text{if } h \in \{-1, 1\} \\ 0, & \text{elsewhere} \end{cases}$ ist die Autokovarianzfunktion eines stationären Prozesses falls $|\rho| \le 0.5$. Folgt aus $f(\lambda) = \dots \ge 0$.

Bemerkung 4.1.7. The preceding discussion shows that a function $f : [0, 1] \to \mathbb{R}$ is the spectral density of a stationary process iff f satisfies the following three conditions

(i) $f(\lambda) \ge 0$,

(ii) $f(\lambda) = f(1 - \lambda)$,

(iii) $\int_0^1 f(\lambda)\, d\lambda < \infty$.

Beweis: Es ist f reell, also Summe über $\mathbb{Z}$ durch Kosinus darstellbar. □

4.2 Lineare Filter und Frequenzen

Satz 4.2.1. *Let $(Z_t)_{t\in\mathbb{Z}}$ be a stationary process with spectral distribution function F_Z and let $(a_t)_{t\in\mathbb{Z}}$ be an absolutely summable filter with Fourier transform f_a. The linear* *filtered process* $Y_t := \sum_{u\in\mathbb{Z}} a_u Z_{t-u}$*,* $t \in \mathbb{Z}$*, then has the spectral distribution function*

$$F_Y(\lambda) := \int_0^\lambda |f_a(x)|^2 \, dF_Z(x), \quad 0 \le \lambda \le 1, \quad f_a(\lambda) = \sum_{w\in\mathbb{Z}} a_w e^{-i2\pi\lambda w}$$

If in addition $(Z_t)_{t\in\mathbb{Z}}$ *has a spectral density* f_Z*, then*

$$f_Y(\lambda) := |f_a(\lambda)|^2 f_Z(\lambda), \qquad 0 \le \lambda \le 1,$$

is the spectral density of $(Y_t)_{t\in\mathbb{Z}}$*.*

Beweis: Autokovfunktion von gefiltertem Prozess bekannt von vorher, dann Spektraldarstellung des vorherigen und Summen und Integral vertauschen. □

4.2.1 Transferfunktion und Powertransferfunktion

Zusatzdefinition. Definition der *Transferfunktion* und *Powertransferfunktion* (gain) eines Filters $(a_t)_{t\in\mathbb{Z}}$.

$$f_a(\lambda) \quad \text{bzw.} \quad g_a(\lambda) := |f_a(\lambda)|^2$$

Zusatzbemerkung. Da die Spektraldiche die Intensität einer Frequenz λ angibt, hat das Anwenden eines linearen Filters (a_t) direkt und deutbaren Einfluss darauf verringern falls $|f_a(\lambda)| < 1$ und vergrößern falls $|f_a(\lambda)| > 1$.

Beispiel 4.2.2. Berechnung und Ausgabe der Powertransferfunktion eines Simple Moving Average, eines Differenzfilters erster Ordnung und eines saisonalen Differenzfilters.

4.2.2 Kleinste Quadrate basierende Filter Designs

Bemerkung 4.2.3. Aus dem vorherigen Abschnitt stellt sich die Frage, ob man Filter mit vorgelegten Eigenschaften finden kann, wie beispielsweise einen *low pass* Filter, *high pass* Filter und einen *band pass* Filter.

Man betrachtet $f_a(\lambda) = \sum_{u=r}^{s} a_u e^{-i2\pi\lambda u}$ und $\int_0^{0.5} |f(\lambda) - f_a(\lambda)|^2 \, d\lambda$ mit dem gesuchten f.

Beispiel 4.2.4. Low pass Filter aus kleinsten Quadraten mit Filter der Länge 40.

4.3 Spektraldichten von ARMA-Prozessen

Satz 4.3.1. *Suppose that*

$$Y_t = a_1 Y_{t-1} + \cdots + a_p Y_{t-p} + \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}, \quad t \in \mathbb{Z},$$

is a stationary ARMA(p, q)*-process, where* (ε_t) *is a white noise with variance* σ^2*. Put*

$$A(z) := 1 - a_1 z - a_2 z^2 - \cdots - a_p z^p,$$
$$B(z) := 1 + b_1 z + b_2 z^2 + \cdots + b_q z^q$$

and suppose that the process (Y_t) satisfies the stationarity condition (2.3), i.e., the roots of the equation $A(z) = 0$ are outside of the unit circle. The process (Y_t) then has the spectral density

$$f_Y(\lambda) = \sigma^2 \frac{|B(e^{-i2\pi\lambda})|^2}{|A(e^{-i2\pi\lambda})|^2} = \sigma^2 \frac{|1 + \sum_{v=1}^{q} b_v e^{-i2\pi\lambda v}|^2}{|1 - \sum_{u=1}^{p} a_u e^{-i2\pi\lambda u}|^2}.$$

Beweis: Existenz folgt, da als $Y_t = \sum_{v \geq 0} \alpha_v \varepsilon_{t-v}$ darstellbar. Dann zweifache Bestimmung der Spektraldichte von $X_t := Y_t - a_1 Y_{t-1} - \cdots - a_p Y_{t-p} = \varepsilon_t + b_1 \varepsilon_{t-1} + \cdots + b_q \varepsilon_{t-q}$ und auflösen (keine Nullstellen von A auf Einheitskreis). □

Beispiel 4.3.2. Aus Satz ergeben sich Darstellungen für MA und AR. Angabe der Spektraldichten von $ARMA(1,1)$, $MA(1)$ und $AR(1)$ Prozessen.

5 Statistische Analyse im Frequenzbereich

5.1 Test auf weißes Rauschen

$$Y_t = \mu + A\cos(2\pi\lambda t) + B\sin(2\pi\lambda t) + \varepsilon_t, \quad H_0 : A = B = 0, \varepsilon_t \sim \mathcal{N}(0, \sigma^2)$$

Zusatzbemerkung. Das Periodogramm wird verwendet um Frequenzen festzustellen, also auch hier möglich. $(Y_t)_{t\in\mathbb{Z}}$ ist nur unter der Nullhypothese ein stationärer Prozess.

5.1.1 Verteilung des Periodogramms

Lemma 5.1.1. *Let $\varepsilon_1, \ldots, \varepsilon_n$ be independent and identically normal distributed random variables with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$. Denote by*

$$C_\varepsilon\left(\frac{k}{n}\right) = \frac{1}{n}\sum_{t=1}^{n}(\varepsilon_t - \bar{\varepsilon})\cos\left(2\pi\frac{k}{n}t\right),$$
$$S_\varepsilon\left(\frac{k}{n}\right) = \frac{1}{n}\sum_{t=1}^{n}(\varepsilon_t - \bar{\varepsilon})\sin\left(2\pi\frac{k}{n}t\right)$$

the cross covariances with Fourier frequencies k/n, $1 \leq k \leq [(n-1)/2]$. Then the $2[(n-1)/2]$ random variables

$$C_\varepsilon(k/n),\ S_\varepsilon(k/n), \quad 1 \leq k \leq [(n-1)/2],$$

are independent and identically $N(0, \sigma^2/(2n))$-distributed.

Beweis: Zusammensetzten von C_ε und S_ε als Vektor $\boldsymbol{v} = \boldsymbol{A}(\boldsymbol{I}_n - n^{-1}\boldsymbol{E}_n)(\varepsilon_t)_{1\leq t\leq n}$ also normalverteilt. Varianz berechenbar. □

Korollar 5.1.2. *Let $\varepsilon_1, \dots, \varepsilon_n$ be as in the preceding lemma and let*

$$I_\varepsilon(k/n) = n\Big\{C_\varepsilon^2(k/n) + S_\varepsilon^2(k/n)\Big\}$$

be the pertaining periodogram, evaluated at the Fourier frequencies k/n, $1 \le k \le [(n-1)/2]$. The random variables $I_\varepsilon(k/n)/\sigma^2$ are independent and identically standard exponential distributed i.e.,

$$P\{I_\varepsilon(k/n)/\sigma^2 \le x\} = \begin{cases} 1 - \exp(-x), & x > 0 \\ 0, & x \le 0. \end{cases}$$

Beweis: Quadrate Normalverteilter sind Chi-Quadrat verteilt. Mit 2 Freiheitsgraden ist dies genau standard exponential von $x/2$. □

Satz 5.1.3. *Let $\varepsilon_1, \dots, \varepsilon_n$ be independent $N(\mu, \sigma^2)$-distributed random variables and denote by*

$$S_j := \frac{\sum_{k=1}^{j} I_\varepsilon(k/n)}{\sum_{k=1}^{m} I_\varepsilon(k/n)}, \quad j = 1, \dots, m := [(n-1)/2],$$

the cumulated periodogram. Note that $S_m = 1$. Then we have

$$\big(S_1, \dots, S_{m-1}\big) =_D \big(U_{1:m-1}, \dots, U_{m-1:m-1}\big).$$

Beweis: Es ist bekannt, dass $(U_{j:m})_{1\le j\le m}$ die gleiche Verteilung hat wie $((Z_1 + \dots + Z_j)/(Z_1 + \dots + Z_{m+1}))_{1\le j\le m}$ (Z iid exponential) □

Korollar 5.1.4. *The empirical distribution function of $S_1, \dots, S_{m-1}$ is distributed like that of $U_1, \dots, U_{m-1}$, i.e.,*

$$\hat{F}_{m-1}(x) := \frac{1}{m-1}\sum_{j=1}^{m-1} 1_{(0,x]}(S_j) =_D \frac{1}{m-1}\sum_{j=1}^{m-1} 1_{(0,x]}(U_j), \quad x \in [0,1].$$

Korollar 5.1.5. *Put $S_0 := 0$ and*

$$M_m := \max_{1\le j\le m}(S_j - S_{j-1}) = \frac{\max_{1\le j\le m} I_\varepsilon(j/n)}{\sum_{k=1}^{m} I_\varepsilon(k/n)}.$$

The maximum spacing M_m has a distribution function.

5.1.2 Fisher's Test

Zusatzdefinition. Definition des Tests nach *Fisher* (Test auf versteckte Periodizitäten).

$$\kappa_m := \frac{\max_{1\le j\le m} I(j/n)}{(1/m)\sum_{k=1}^{m} I(k/n)} = mM_m$$

5.1.3 Der Bartlett–Kolmogrov–Smirnov Test

Zusatzdefinition. Definition des Tests nach *Bartlett–Kolmogrov–Smirnov.*
Falls $Y_t = \varepsilon_t$ mit ε_t $\mathcal{N}(\mu, \sigma^2)$-Verteilt, so ist $\hat{F}_{m-1}$ of $S_1, \ldots, S_{m-1}$ wie die einer Gleichverteilung. Also Überprüfung, ob

$$\Delta_{m-1} := \sup_{x \in [0,1]} |\hat{F}_{m-1}(x) - x|$$

besonders groß ist (Vergleich mit der Theoretischen Verteilung $F(x) = x$ der Gleichverteilung).

Beispiel 5.1.6. Test der varianzstabilisierten, trend- und saisonalbereinigten Airline Daten. Der Bartlett-Test lässt sich auch als Konfidenzband angeben.

5.2 Schätzung von Spektraldichten

Es sei nun $(Y_t)_{t\in\mathbb{Z}}$ ein stationärer reellwertiger Prozess mit Mittelwert μ und absolut summierbarer Autokovarianzfunktion γ.
Die Spektraldiche ist dann $f(\lambda) = \sum_{h\in\mathbb{Z}} \gamma(h) e^{-i2\pi\lambda h}$.

5.2.1 Asymptotische Eigenschaften des Periodogramms

$$\begin{aligned}
I_n(k/n) = n|D(\lambda)|^2 &= \frac{1}{n}\Big|\sum_{t=1}^{n} Y_t e^{-i2\pi(k/n)t}\Big|^2 \\
&= \frac{1}{n}\left\{\Big(\sum_{t=1}^{n} Y_t \cos\Big(2\pi\frac{k}{n}t\Big)\Big)^2 + \Big(\sum_{t=1}^{n} Y_t \sin\Big(2\pi\frac{k}{n}t\Big)\Big)^2\right\}.
\end{aligned}$$

Zusatzbemerkung. Bis auf $k = 0$ ist dies die Definition. Durch vorherigen Satz erhält man $I_n(k/n) = \begin{cases} n\bar{Y}_n^2, & k = 0 \\ \sum_{|h|<n} c(h) e^{-i2\pi(k/n)h}, & k = 1, \ldots, [n/2] \end{cases}$, mit $c(h) = \frac{1}{n}\sum_{t=1}^{n-|h|}\big(Y_t - \bar{Y}_n\big)\big(Y_{t+|h|} - \bar{Y}_n\big)$.
$I_n(k/n)$ ändert sich auch nicht, wenn man in $c(h)$ den Wert $\bar{Y}_n$ durch μ ersetzt. Für $k \neq 0$:

$$\begin{aligned}
I_n(k/n) &= \sum_{|h|<n} \frac{1}{n}\Big(\sum_{t=1}^{n-|h|} (Y_t - \mu)(Y_{t+|h|} - \mu)\Big) e^{-i2\pi(k/n)h} \\
&= \frac{1}{n}\sum_{t=1}^{n}(Y_t - \mu)^2 + 2\sum_{h=1}^{n-1}\frac{1}{n}\Big(\sum_{t=1}^{n-|h|}(Y_t - \mu)(Y_{t+|h|} - \mu)\Big)\cos\Big(2\pi\frac{k}{n}h\Big).
\end{aligned}$$

Zusatzdefinition. Definition von $I_n(\lambda)$ für $\lambda \in [0, 0.5]$. (Stückweise Treppenförmig)

Satz 5.2.1 (Erwartungswert konvergiert gegen $f(\lambda)$). *Let $(Y_t)_{t\in\mathbb{Z}}$ be a stationary process with absolutely summable autocovariance function γ. Then we have with $\mu = \mathrm{E}(Y_t)$*

$$\mathrm{E}(I_n(0)) - n\mu^2 \overset{n\to\infty}{\longrightarrow} f(0),$$
$$\mathrm{E}(I_n(\lambda)) \overset{n\to\infty}{\longrightarrow} f(\lambda), \qquad \lambda \neq 0.$$

If $\mu = 0$, then the convergence $\mathrm{E}(I_n(\lambda)) \overset{n\to\infty}{\longrightarrow} f(\lambda)$ holds uniformly on $[0, 0.5]$.

Beweis: Für 0 einsetzten mit Césaro Konvergenz, im andern Fall Hilfsfunktion und dann lang. □

Satz 5.2.2. *Let $Z_1, \dots, Z_n$ be independent and identically distributed random variables with mean $\mathrm{E}(Z_t) = 0$ and variance $\mathrm{E}(Z_t^2) = \sigma^2 < \infty$.*

(a) The random vector $(I_n(\lambda_1), \dots, I_n(\lambda_r))$ with $0 < \lambda_1 < \dots < \lambda_r < 0.5$ converges in distribution for $n \to \infty$ to the distribution of r independent and identically exponential distributed random variables with mean σ^2.

(b) If $\mathrm{E}(Z_t^4) = \eta\sigma^4 < \infty$, then we have for $k = 0, \dots, [n/2]$

$$\mathrm{Var}(I_n(k/n)) = \begin{cases} 2\sigma^4 + n^{-1}(\eta - 3)\sigma^4, & k = 0 \text{ or } k = n/2, \text{ if } n \text{ even} \\ \sigma^4 + n^{-1}(\eta - 3)\sigma^4 & \text{elsewhere} \end{cases}$$

and

$$\mathrm{Cov}(I_n(j/n), I_n(k/n)) = n^{-1}(\eta - 3)\sigma^4, \quad j \neq k.$$

Beweis: Erster Teil ähnlich wie Korrolar (Normalität eines Vektors zeigen), verschiedene Literaturquellen. 2. Fall $I_n(k/n) = \frac{1}{n}\sum_{s=1}^n \sum_{t=1}^n Z_s Z_t e^{-i2\pi(k/n)(s-t)}$ schreiben und $\mathrm{E}(I_n(j/n) I_n(k/n))$ berechnen. □

Zusatzbemerkung. Für $\mathcal{N}(0, \sigma^2)$-verteilte Zufallsvariablen gilt $\eta = 3$ und sie sind unkorreliert. Wir haben in einem Korollar sogar gezeigt, dass sie unabhängig sind.

Bemerkung 5.2.3. This Theorem can be generalized to filtered processes

$$Y_t = \sum_{u\in\mathbb{Z}} a_u Z_{t-u}, \text{ with } (Z_t)_{t\in\mathbb{Z}}$$

. In this case one has to replace σ^2, which equals the constant spectral density $f_Z(\lambda)$, in (a) by the spectral density $f_Y(\lambda_i)$, $1 \le i \le r$. If in addition $\sum_{u\in\mathbb{Z}} |a_u||u|^{1/2} < \infty$, then we have in (b) the expansions

$$\mathrm{Var}(I_n(k/n)) = \begin{cases} 2f_Y^2(k/n) + O(n^{-1/2}), & k = 0 \text{ or } k = n/2, \text{ if } n \text{ is even} \\ f_Y^2(k/n) + O(n^{-1/2}) & \text{elsewhere,} \end{cases}$$

and

$$\mathrm{Cov}(I_n(j/n), I_n(k/n)) = O(n^{-1}), \quad j \neq k,$$

where I_n is the periodogram pertaining to $Y_1, \dots, Y_n$. The above terms $O(n^{-1/2})$ and $O(n^{-1})$ are uniformly bounded in k and j by a constant C.

Zusatzbemerkung. Die Familie der Prozesse $Y_t = \sum_{u\in\mathbb{Z}} a_u Z_{t-u}$ enthält insbesondere stationäre $ARMA$-Prozesse.

5.2.2 Discrete Spectral Average Estimator

Zusatzbemerkung. Das Periodogramm ist kein konsistenter Schätzer der Spektraldichte (sie konvergiert nicht gegen sie, nur der Erwartungswert). Der Zentrale Grenzwertsatz und die Bemerkung eben motivieren aber lineare Glätter (z.B. Moving Averages).

Zusatzdefinition. Definition eines *diskreten Spektral-Mittelwerts-Schätzers.*

$$\hat{f}_n(\lambda) := \hat{f}_n(g_n(\lambda)), \quad \hat{f}_n\Big(\frac{k}{n}\Big) := \sum_{|j|\le m} a_{jn} I_n\Big(\frac{k+j}{n}\Big),$$

dabei erfüllt die Sequenz $m = m(n)$ $m \overset{n\to\infty}{\longrightarrow} \infty$ und $m/n \overset{n\to\infty}{\longrightarrow} 0$ und die Gewichte a_{jn} erfüllen

(a) $a_{jn} \ge 0$,

(b) $a_{jn} = a_{-jn}$,

(c) $\sum_{|j|\le m} a_{jn} = 1$,

(d) $\sum_{|j|\le m} a_{jn}^2 \overset{n\to\infty}{\longrightarrow} 0$.

Zusatzbeispiel. Gilt beispielsweise für einfache gleitende Durchschnitte.

Satz 5.2.4 (Konsistenz)**.** *Let* $Y_t = \sum_{u\in\mathbb{Z}} b_u Z_{t-u}$, $t \in \mathbb{Z}$, *where* Z_t *are iid with* $\mathrm{E}(Z_t) = 0$, $\mathrm{E}(Z_t^4) < \infty$ *and* $\sum_{u\in\mathbb{Z}} |b_u||u|^{1/2} < \infty$. *Then we have for* $0 \le \mu, \lambda \le 0.5$

(i) $\lim_{n\to\infty} \mathrm{E}\left(\hat{f}_n(\lambda)\right) = f(\lambda)$,

(ii) $\lim_{n\to\infty} \dfrac{\mathrm{Cov}\left(\hat{f}_n(\lambda), \hat{f}_n(\mu)\right)}{\left(\sum_{|j|\le m} a_{jn}^2\right)} = \begin{cases} 2f^2(\lambda), & \lambda = \mu = 0 \text{ or } 0.5 \\ f^2(\lambda), & 0 < \lambda = \mu < 0.5 \\ 0, & \lambda \ne \mu. \end{cases}$

Beweis: 1. Erweiterung auf Dreiecksungleichung, zweiter verschwindet aufgrund gleichmäßiger Konvergenz, erster da $I_n(g_n(\lambda) + j/n)$ gegen $f(g_n(\lambda) + j/n)$ konvergiert (in der Summe enthalten). 2. Kovarianz ausrechnen und abschätzen. □

Zusatzbemerkung.

- Die Eigenschaft (d) der Gewichte sorgt zusammen mit der zweiten Eigenschaft dabei dafür dass gilt $\mathrm{Var}(\hat{f}_n(\lambda)) \overset{n\to\infty}{\longrightarrow} 0$ und somit verschwindet nach der ersten Eigenschaft auch der MSE.
- Die Voraussetzung $\mathrm{E}(Y_t) = 0$ ist in der Praxis meist sehr einschränkend, aber die Periodogramme von $(Y_t)_{1\le t\le n}$, $(Y_t - \mu)_{1\le t\le n}$ und $(Y_t - \bar{Y})_{1\le t\le n}$ unterscheiden sich nicht für Fourierfrequenzen ungleich Null. Im Fall $\mu \ne 0$ schätzt man $f(0)$ durch benachbarte Werte.

Beispiel 5.2.5. Schätzung der Spektraldiche der Sunspot-Daten (*leakage phenomenon*).

Zusatzbemerkung. Ein bequemer Weg Gewichte a_{jn} zu konstruieren, die (a) bis (d) erfüllen ist über symmetrische Kernel-Funktionen $[-1, 1] \to \infty$ mit $\int_{-1}^{1} K^2(x)\,dx < \infty$. Sei dann $m = m(n) \overset{n\to\infty}{\longrightarrow} \infty$ beliebig mit $m/n \overset{n\to\infty}{\longrightarrow} 0$

$$a_{jn} := \frac{K(j/m)}{\sum_{i=-m}^{m} K(i/m)}, \quad -m \le j \le m.$$

Zusatzbeispiel. Der *truncated Kernel*, *Dreieckskernel*, *Tukey-Hanning Kernel* und *Parzen Kernel.*

5.2.3 Konfidenzintervalle für die Spektraldichte

Die Zufallsvariablen $I_n\big((k+j)/n\big)/f\big((k+j)/n\big)$ sind nach vorheriger Bemerkung 5.2.4 für große n approximativ unab. standardexponential. Also kann die Verteilung des diskreten Spektral-Mittel-Schätzers durch eine gewichtete Summe approximiert werden. Tukey zeigte, dass diese durch die *Gammaverteilung* mit Parameter ν angenähert werden kann. Dabei wählt man die Koeffizienten so, dass

$$\begin{aligned} \mathrm{E}(cY) &= c\nu = f(k/n), \\ \mathrm{Var}(cY) &= 2c^2\nu = f^2(k/n) \sum_{|j|\leq m} a_{jn}^2. \end{aligned}$$

Falls ν ein Integer ist, so ist die Gammaverteilung die χ^2_ν-Verteilung. Die Variable

$$\frac{\nu \hat{f}(k/n)}{f(k/n)} = \frac{\hat{f}(k/n)}{c}$$

ist dann χ^2_ν verteilt.

$$\left(\frac{\nu \hat{f}(k/n)}{\chi^2_{1-\alpha/2}(\nu)}, \frac{\nu \hat{f}(k/n)}{\chi^2_{\alpha/2}(\nu)} \right)$$

ist dann ein $1-\alpha$-Konfidenzbereich für $f(k/n)$.

Zusatzbemerkung. Nimmt man den Logarithmus und bestimmt das Konfidenzintervall von $\log(f(k/n))$, so hat es eine konstante Länge. Es ist aber nur ein $1-\alpha$ Konfidenzbereich für eine Fourierfrequenz und nicht gleichzeitig für alle.