# Stochastik-Praktikum

## Zeitreihenanalyse

Peter Frentrup

Humboldt-Universität zu Berlin

19. Dezember 2017

# Übersicht

1. Zeitreihen
   - Weißes Rauschen (WN)
   - Moving Average (MA)
   - Autoregressive (AR, ARMA)
   - Autoregressiv mit bedingter Heteroskedastizität (ARCH)

2. Vorhersagen
   - Yule-Walker-Schätzer

# Zeitreihen

## Definition

Ein stochastischer Prozess $(X_t)_{t\in T}$ mit diskretem $T = \mathbb{N}_0$ oder $T = \mathbb{Z}$ heißt **Zeitreihe**.

Eine Zeitreihe $(X_t)$ heißt **(strikt) stationär**, falls

$$\forall n \in \mathbb{N}, t_1, \ldots, t_n, t \in T : (X_{1_t}, \ldots, X_{t_n}) \stackrel{d}{=} (X_{t_1+t}, \ldots, X_{t_n+t})\,.$$

Eine Zeitreihe $(X_t)$ mit $X_t \in L^2(\mathbb{P})$ heißt **schwach stationär** falls $\forall r, s, t \in T$:

$$\mathbb{E}[X_s] = \mathbb{E}[X_{s+t}]\,,$$
$$\mathrm{Cov}(X_r, X_s) = \mathrm{Cov}(X_{r+t}, X_{s+t})\,.$$

In diesem Fall heißt $t \mapsto c(t) := \mathrm{Cov}(X_s, X_{s+t})$ **Autokovarianzfunktion** ($s \in T$ beliebig) und $t \mapsto \rho(t) := c(t)/c(0)$ **Autokorrelationsfunktion**.

# Zeitreihen

Stationäre Zeitreihen

1. Jährliche Regenmengen
2. Umtauschkurs Euro–Dollar
3. Anzahl Autounfälle
4. Herzfrequenz einer gesunden Person

Nicht-Stationäre Zeitreihen

1. Jährliche Marienkäferpopulation
2. Tägliche Regenmengen
3. Stündliches Verkehrsaufkommen an der Rudower Chaussee

Nach herausrechnen von **Trends** und **saisonalen Komponenten** können letztere teilweise auch als strationär angenommen werden.

# Zeitreihen

Stationäre Zeitreihen

1. Jährliche Regenmengen
2. Umtauschkurs Euro–Dollar
3. Anzahl Autounfälle
4. Herzfrequenz einer gesunden Person

Nicht-Stationäre Zeitreihen

1. Jährliche Marienkäferpopulation
2. Tägliche Regenmengen
3. Stündliches Verkehrsaufkommen an der Rudower Chaussee

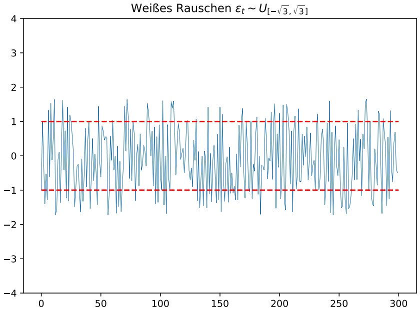
Nach herausrechnen von **Trends** und **saisonalen Komponenten** können letztere teilweise auch als strationär angenommen werden.

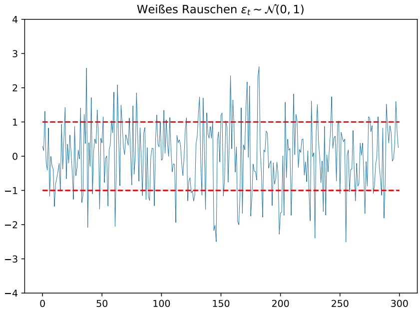
## Definition

Ein *schwach stationärer* Prozess $(\varepsilon_t)_{t\in\mathbb{Z}}$ mit $\mathbb{E}[\varepsilon_t] = 0$ und Autokovarianzfunktion

$$c(t) = \begin{cases} \sigma^2 & t = 0, \\ 0 & t \neq 0 \end{cases}$$

heißt **weißes Rauschen** (engl. **white noise**), $(\varepsilon_t) \sim \mathrm{WN}(0, \sigma^2)$.
Falls zudem die $\varepsilon_t$ i.i.d. sind, so schreiben wir $(\varepsilon_t) \sim \mathrm{IID}(0, \sigma^2)$.



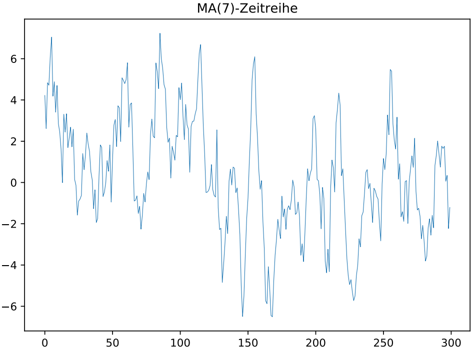



## Definition (MA)

Für $(\varepsilon_t) \sim \mathrm{WN}(0, \sigma^2)$ und $b_1, \ldots, b_q \in \mathbb{R}$, $q \in \mathbb{N}$, definiert

$$X_t := \varepsilon_t + b_1 \varepsilon_{t-1} + \ldots + b_q \varepsilon_{t-q}\,, \qquad t \in \mathbb{Z}\,,$$

eine stationäre Zeitreihe, genannt **Moving Average** Zeitreihe $(X_t) \sim \mathrm{MA}(q)$.



### Definition (AR und ARMA)

Eine Zeitreihe $(X_t)_{t\in\mathbb{Z}}$ heißt **Autoregressiv** mit Ordnung $p \in \mathbb{N}$, $(X_t) \sim \mathrm{AR}(p)$, falls

$$X_t := a_1X_{t-1} + a_2X_{t-2} + \ldots + a_pX_{t-p} + \varepsilon_t\,, \qquad t \in \mathbb{Z},$$

für $(\varepsilon_t) \sim \mathrm{WN}(0, \sigma^2)$ und Parameter $a_1, \ldots, a_p \in \mathbb{R}$, $p \in \mathbb{N}$.

Mit $b_1, \ldots, b_q \in \mathbb{R}$, $q \in \mathbb{N}$, heißt eine Lösung von

$$X_t := a_1X_{t-1} + \ldots + a_pX_{t-p} + \varepsilon_t + b_1\varepsilon_{t-1} + \ldots + b_q\varepsilon_{t-q}\,, \quad t \in \mathbb{Z},$$

**Autoregressive Moving Average** Zeitreihe, $(X_t) \sim \mathrm{ARMA}(p, q)$

$X \sim \mathrm{ARMA}(p, q)$:

$$X_t := a_1 X_{t-1} + \ldots + a_p X_{t-p} + \varepsilon_t + b_1 \varepsilon_{t-1} + \ldots + b_q \varepsilon_{t-q}\,, \quad t \in \mathbb{Z}.$$

**Lemma**

*Sei $(\varepsilon_t) \sim \mathrm{WN}(0, \sigma^2)$. Falls $|a| < 1$, so gibt genau eine schwach stationäre AR(1)-Zeitreihe $X_t = aX_{t-1} + \varepsilon_t$, $t \in \mathbb{Z}$. (nämlich $X_t = \sum_{j=0}^{\infty} a^j \varepsilon_{t-j}$)*

Für viele konkrete $(\varepsilon_t)_{t\in\mathbb{Z}} \sim \mathrm{WN}(0, \sigma^2)$ und Parameter $a_j, b_j$ lassen sich die Anfangswerte $X_{-p+1}, \ldots, X_0$ einer (schwach) stationären ARMA$(p, q)$-Zeitreihe $(X_t)_{t\in\mathbb{N}}$ explizit berechnen/simulieren, ohne alle $(\varepsilon_t)_{t\leq 0}$ simulieren zu müssen.

# Autoregressiv mit bedingter Heteroskedastizität (ARCH)

## Definition (ARCH(1)-Zeitreihe)

Für zufälliges $X_0$ mit $\mathbb{E}[X_0^2] < \infty$, unabhängiges $(\varepsilon_t) \sim \mathrm{IID}(0,1)$, und Parameter $a_0 > 0$ und $a_1 \in (0,1)$ heißt

$$X_t = \sqrt{a_0 + a_1 X_{t-1}^2}\,\varepsilon_t\,, \qquad t \in \mathbb{N}, \tag{1}$$

**autoregressiv mit bedingter Heteroskedastizität** (engl. *ARCH*(1)).

*ARCH*-Prozesse wurden 1982 von Robert Engle zur Modellierung finanzieller Zeitreihen (z.B. Renditen eines Finanzgutes) eingeführt.

## Lemma

$X_t$ *aus* (1) *ist ein weißes Rauschen (bei geeignet gewähltem* $X_0$*).*

Beweis: Übung.

# Autoregressiv mit bedingter Heteroskedastizität (ARCH)

## Definition (ARCH(1)-Zeitreihe)

Für zufälliges $X_0$ mit $\mathbb{E}[X_0^2] < \infty$, unabhängiges $(\varepsilon_t) \sim \mathrm{IID}(0, 1)$, und Parameter $a_0 > 0$ und $a_1 \in (0, 1)$ heißt

$$X_t = \sqrt{a_0 + a_1 X_{t-1}^2}\, \varepsilon_t\,, \qquad t \in \mathbb{N}, \tag{1}$$

**autoregressiv mit bedingter Heteroskedastizität** (engl. *ARCH*(1)).

*ARCH*-Prozesse wurden 1982 von Robert Engle zur Modellierung finanzieller Zeitreihen (z.B. Renditen eines Finanzgutes) eingeführt.

## Lemma

*$X_t$ aus (1) ist ein weißes Rauschen (bei geeignet gewähltem $X_0$).*

Beweis: Übung.

# Übersicht

1. Zeitreihen
   - Weißes Rauschen (WN)
   - Moving Average (MA)
   - Autoregressive (AR, ARMA)
   - Autoregressiv mit bedingter Heteroskedastizität (ARCH)

2. Vorhersagen
   - Yule-Walker-Schätzer

# Vorhersagen

AR($p$)-Zeitreihe $X_{t+1} = a_1 X_t + \ldots + a_p X_{t-p+1} + \varepsilon_{t+1}$, $(\varepsilon_t) \sim \mathrm{WN}(0, \sigma^2)$.

Die **beste lineare Vorhersage** für $X_{t+1}$ gegeben der Beobachtungen $X_0, \ldots, X_t$ $(t \geq p)$ ist

$$\hat{X}_{t+1} := a_1 X_t + \ldots + a_p X_{t-p+1} + \underbrace{\mathbb{E}[\varepsilon_{t+1}]}_{=0},$$

d.h. $\mathbb{E}[(\hat{X}_{t+1} - X_{t+1})^2 \mid X_0, \ldots, X_t]$ ist minimal für diesen Schätzer.

Problem: Parameter $a_1, \ldots, a_p$ unbekannt.

# Vorhersagen

AR($p$)-Zeitreihe $X_{t+1} = a_1 X_t + \ldots + a_p X_{t-p+1} + \varepsilon_{t+1}$, $(\varepsilon_t) \sim \mathrm{WN}(0, \sigma^2)$.

Die **beste lineare Vorhersage** für $X_{t+1}$ gegeben der Beobachtungen $X_0, \ldots, X_t$ $(t \geq p)$ ist

$$\hat{X}_{t+1} := a_1 X_t + \ldots + a_p X_{t-p+1} + \underbrace{\mathbb{E}[\varepsilon_{t+1}]}_{=0},$$

d.h. $\mathbb{E}[(\hat{X}_{t+1} - X_{t+1})^2 \mid X_0, \ldots, X_t]$ ist minimal für diesen Schätzer.

Problem: Parameter $a_1, \ldots, a_p$ unbekannt.

**Empirisches Mittel** gegeben $X_1, \ldots, X_n$:

$$\hat{\mu}_n := \frac{1}{n} \sum_{i=1}^{n} X_i \,.$$

**Empirische Autokovarianz** gegeben $X_1, \ldots, X_n$:

$$\hat{c}(h) := \frac{1}{n} \sum_{i=1}^{n-h} (X_i - \hat{\mu}_n)(X_{i+h} - \hat{\mu}_n)\,, \quad c(-h) := c(h)\,, \quad h \in \mathbb{N}_0 \,.$$

**Empirische Autokorrelation** gegeben $X_1, \ldots, X_n$:

$$\hat{\rho}(h) := \frac{\hat{c}(h)}{\hat{c}(0)} \,.$$

# Yule-Walker-Schätzer

$X$ : schwach stationärer, kausaler (d.h. $X_t$ hängt nur von Vergangenheit ab) AR(1)-Prozess mit $(\varepsilon_t) \sim \mathrm{WN}(0, \sigma^2)$.
**Problem:** Parameter $a_1, \ldots, a_p$ Schätzen.

$\mathbb{E}[X_t] = 0$, Autokovarianz:

$$\begin{aligned} c(h) &= \mathrm{Cov}(X_t, X_{t-h}) = \mathrm{Cov}(a_1 X_{t-1} + \ldots + a_p X_{t-p} + \varepsilon_t, X_{t-h}) \\ &= a_1 c(h-1) + \ldots + a_p c(h-p) \quad \text{für } h \geq 1, \\ c(0) &= \mathrm{Cov}(X_t, X_t) = \mathrm{Cov}(a_1 X_{t-1} + \ldots + a_p X_{t-p} + \varepsilon_t, X_t) \\ &= a_1 c(-1) + \ldots + a_p c(-p) + \sigma^2 . \end{aligned}$$

$\rightsquigarrow$

$$\left.\begin{aligned} c(1) &= a_1 c(0) + \cdots + a_p c(p-1) \\ &\vdots \\ c(p) &= a_1 c(p-1) + \cdots + a_p c(0) \end{aligned}\right\} \Rightarrow c_p = C_p \cdot a$$

mit $c_p = (c(1), \ldots, c(p))^\top$, $C_p = \big(c(i-j)\big)_{1 \leq i,j \leq p}$, $a = (a_1, \ldots, a_p)^\top$.

# Yule-Walker-Schätzer

$X$ : schwach stationärer, kausaler (d.h. $X_t$ hängt nur von Vergangenheit ab) AR(1)-Prozess mit $(\varepsilon_t) \sim \mathrm{WN}(0, \sigma^2)$.
**Problem:** Parameter $a_1, \ldots, a_p$ Schätzen.

$\mathbb{E}[X_t] = 0$, Autokovarianz:

$$\begin{aligned} c(h) &= \mathrm{Cov}(X_t, X_{t-h}) = \mathrm{Cov}(a_1 X_{t-1} + \ldots + a_p X_{t-p} + \varepsilon_t, X_{t-h}) \\ &= a_1 c(h-1) + \ldots + a_p c(h-p) \quad \text{für } h \geq 1, \\ c(0) &= \mathrm{Cov}(X_t, X_t) = \mathrm{Cov}(a_1 X_{t-1} + \ldots + a_p X_{t-p} + \varepsilon_t, X_t) \\ &= a_1 c(-1) + \ldots + a_p c(-p) + \sigma^2 \,. \end{aligned}$$

$\rightsquigarrow$

$$\left.\begin{aligned} c(1) &= a_1 c(0) + \cdots + a_p c(p-1) \\ &\vdots \\ c(p) &= a_1 c(p-1) + \cdots + a_p c(0) \end{aligned}\right\} \Rightarrow c_p = C_p \cdot a$$

mit $c_p = (c(1), \ldots, c(p))^\top$, $C_p = \big(c(i-j)\big)_{1 \leq i,j \leq p}$, $a = (a_1, \ldots, a_p)^\top$.

# Yule-Walker-Schätzer

$X$ : schwach stationärer, kausaler (d.h. $X_t$ hängt nur von Vergangenheit ab) AR(1)-Prozess mit $(\varepsilon_t) \sim \mathrm{WN}(0, \sigma^2)$.
**Problem:** Parameter $a_1, \ldots, a_p$ Schätzen.

$\mathbb{E}[X_t] = 0$, Autokovarianz:

$$\begin{aligned} c(h) &= \mathrm{Cov}(X_t, X_{t-h}) = \mathrm{Cov}(a_1 X_{t-1} + \ldots + a_p X_{t-p} + \varepsilon_t, X_{t-h}) \\ &= a_1 c(h-1) + \ldots + a_p c(h-p) \quad \text{für } h \geq 1, \\ c(0) &= \mathrm{Cov}(X_t, X_t) = \mathrm{Cov}(a_1 X_{t-1} + \ldots + a_p X_{t-p} + \varepsilon_t, X_t) \\ &= a_1 c(-1) + \ldots + a_p c(-p) + \sigma^2 . \end{aligned}$$

$\rightsquigarrow$

$$\left.\begin{aligned} c(1) &= a_1 c(0) + \cdots + a_p c(p-1) \\ &\vdots \\ c(p) &= a_1 c(p-1) + \cdots + a_p c(0) \end{aligned}\right\} \Rightarrow c_p = C_p \cdot a$$

mit $c_p = (c(1), \ldots, c(p))^\top$, $C_p = \big(c(i-j)\big)_{1 \leq i,j \leq p}$, $a = (a_1, \ldots, a_p)^\top$.

# Yule-Walker-Schätzer

$X$ : schwach stationärer, kausaler (d.h. $X_t$ hängt nur von Vergangenheit ab) AR(1)-Prozess mit $(\varepsilon_t) \sim \mathrm{WN}(0, \sigma^2)$.
**Problem:** Parameter $a_1, \ldots, a_p$ Schätzen.

$$\left.\begin{aligned} c(1) &= a_1 c(0) + \cdots + a_p c(p-1) \\ &\vdots \\ c(p) &= a_1 c(p-1) + \cdots + a_p c(0) \end{aligned}\right\} \Rightarrow c_p = C_p \cdot a$$

mit $c_p = (c(1), \ldots, c(p))^\top$, $C_p = \big(c(i-j)\big)_{1 \le i,j \le p}$, $a = (a_1, \ldots, a_p)^\top$.

### Definition (Yule-Walker-Schätzer)

Mit empirischer Autokovarianz $\hat{c}(h) = \frac{1}{n} \sum_{i=1}^{n-h} X_i X_{i+h}$ (da $\mathbb{E}[X_i] = 0$) heißt die Lösung $\hat{a}$ von $\hat{C}_p \cdot \hat{a} = \hat{c}_p$ **Yule-Walker-Schätzer.**