univariate/multivariate
Zeitreihenanalyse

lineare Verfahren

- statistische Momente
- Fourier Transformation
- Hilbert Transformation
- Wavelet Transformation
- Auto- / Kreuzkorrelations-
  funktion
- ARMA-Modelle
- ...

nichtlineare Verfahren

- Phasenraumrekonstruktion
  - Dimensionen
  - Lyapunov-Exponenten
  - Entropien
  - Determinismus
  - instabile periodische Orbits
  - Interdependenzen
  - ...
- nichtlineare ARMA-Modelle

lineare Verfahren

heißen linear, weil Dynamik eines Systems (und damit Zeitreihe) als linear angenommen wird
Sei H die Dynamik eines Systems und (x_1, x_2) zwei Zustandsvektoren

H ist linear, wenn

$$H(c_1 x_1 + c_2 x_2) = c_1 H(x_1) + c_2 H(x_2)$$

(Superpositionsprinzip)

Zerlegung gilt bei nichtlinearen Systemen nicht !!

Lineare Methoden

statistische Datenanalyse

modell-unabhängig

- Momente von Verteilungen
- Gleichheit/Ungleichheit von Verteilungen
- lineare Korrelation
- ...

$\downarrow$

deskriptive Statistik

modell-abhängig

- Datenanpassung an Modelle
- Parameterschätzung
- least-squares fit
- robuste Schätzung
- ...

Lineare Methoden

statistische Datenanalyse

- basiert meist auf Analyse von Verteilungen / Häufigkeiten

- keine Information über Dynamik

- notwendig für verschiedene Definitionen und Diskriminanz-Tests

Momente einer Verteilung

1. Moment: **Mittelwert**

$$\langle v \rangle = \bar{v} = \frac{1}{N} \sum_{j=1}^{N} v_j$$

Zentralwert einer Verteilung

Lineare Methoden

Momente einer Verteilung

2. Moment: **Varianz / Standardabweichung**

$$Var(v_1,\ldots,v_N) = \frac{1}{N-1}\sum_{j=1}^{N}(v_j - \bar{v})^2$$

$$\sigma(v_1,\ldots,v_N) = \sqrt{Var(v_1,\ldots,v_N)}$$

Breite/Variabilität/Dispersion einer Verteilung um den Zentralwert

Lineare Methoden

Momente einer Verteilung

3. Moment: **Skewness**

$$Skew(v_1,\ldots,v_N) = \frac{1}{N}\sum_{j=1}^{N}\left(\frac{v_j - \bar{v}}{\sigma}\right)^3$$

Grad der Asymmetrie einer Verteilung um den Zentralwert (relativ zur Normalverteilung)

$$Skew = \begin{cases} <0 & \text{asymmetrisch, negativere Verteilungsenden} \\ 0 & \text{symmetrisch, Normalverteilung} \\ >0 & \text{asymmetrisch, positivere Verteilungsenden} \end{cases}$$

Lineare Methoden

Momente einer Verteilung

4. Moment: **Kurtosis**

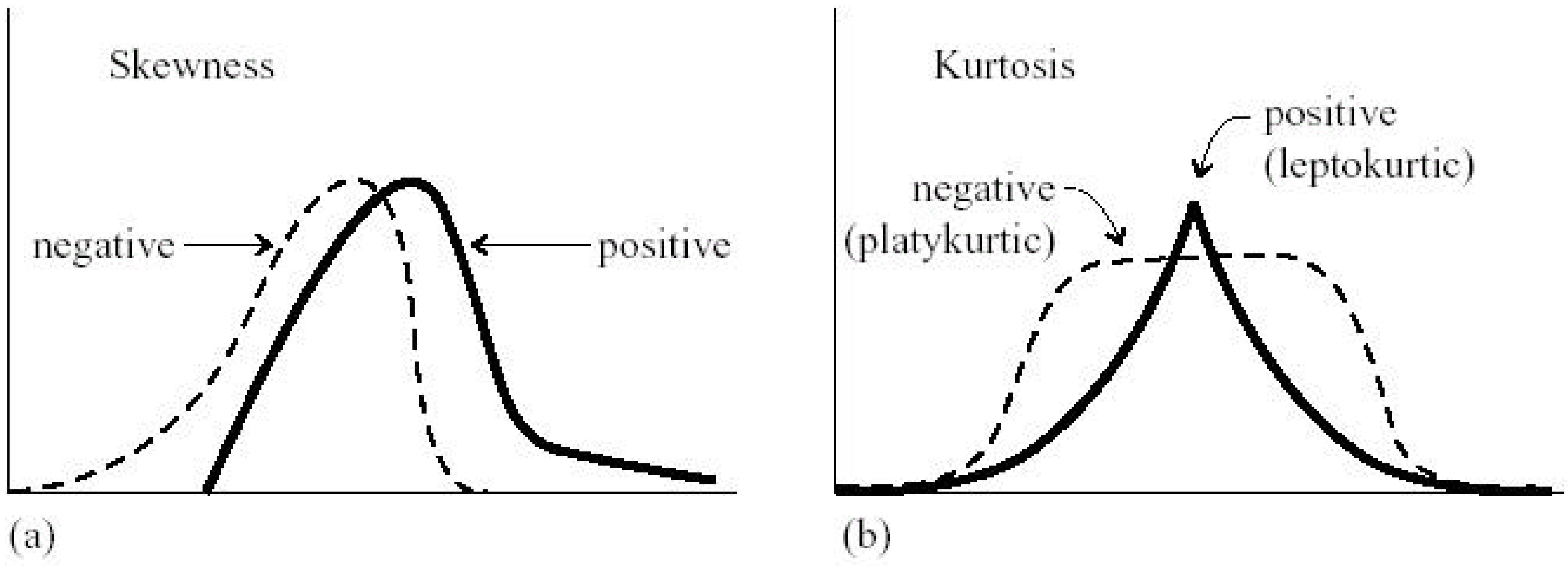
$$Kurt(v_1,\ldots,v_N) = \left\{ \frac{1}{N} \sum_{j=1}^{N} \left(\frac{v_j - \bar{v}}{\sigma} \right)^4 \right\} - 3$$

Grad der Flachheit/Steilheit einer Verteilung (relativ zur Normalverteilung)

$$Kurt = \begin{cases} <0 & \text{platokurtisch, flache Verteilung} \\ \;\; 0 & \text{mesokurtisch (Normalverteilung)} \\ >0 & \text{leptokurtisch, spitze Verteilung} \end{cases}$$

Lineare Methoden

3. und 4. Moment einer Verteilung



Lineare Methoden

Hinweise für Nichtlinearität

betrachte 3. und 4. Moment und deren Standard-abweichungen:

$$\sigma(Skew) = \sqrt{6/N}$$

$$\sigma(Kurt) = \sqrt{24/N}$$

für Normalverteilung

$$Skew(v_i) >> \sigma(Skew)$$

$$Kurt(v_i) >> \sigma(Kurt)$$

Lineare Methoden

Gleichheit/Ungleichheit von Verteilungen

Kolmogorov-Smirnov-Test:
Abstand von kumulativen Verteilungsfunktionen *S*

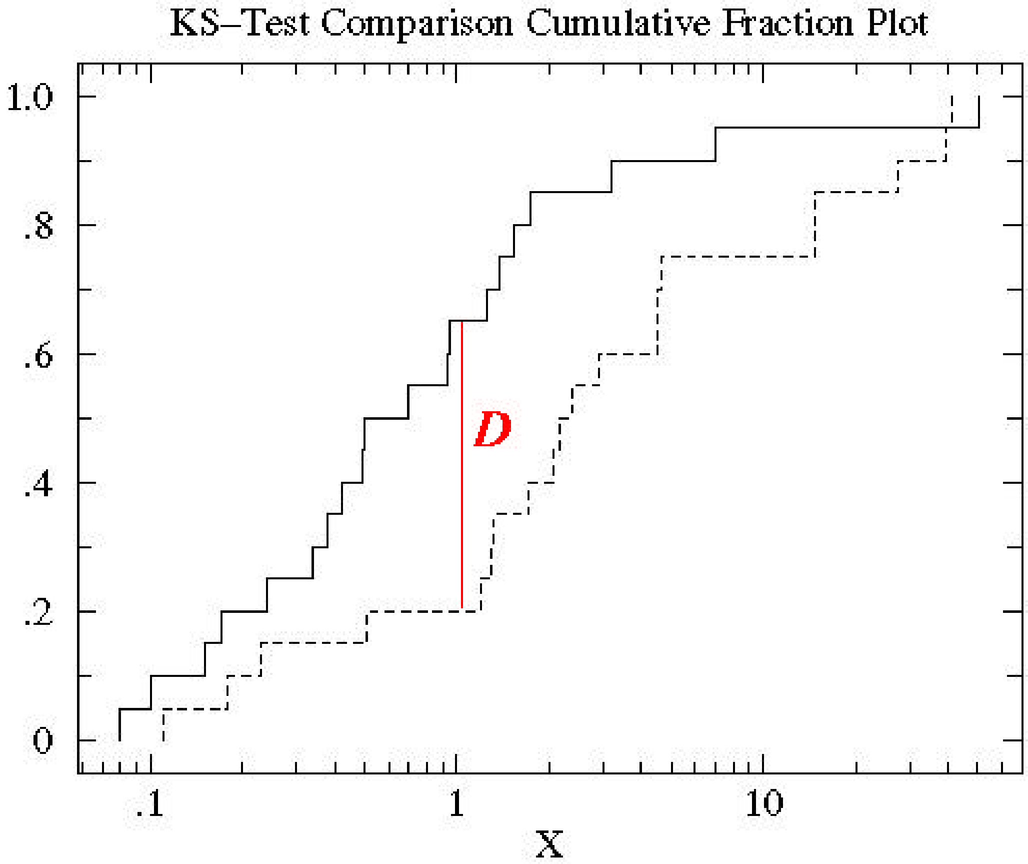
$$D_{KS} = \max_{-\infty<v,w<\infty} \left| S_{N_v}(v) - S_{N_w}(w) \right|$$

Signifikanz für $N_v + N_w$ Freiheitsgrade
- Tabellen

Lineare Methoden

Gleichheit/Ungleichheit von Verteilungen

Kolmogorov-Smirnov-Test:
Abstand von kumulativen Verteilungsfunktionen *S*



Lineare Methoden

Gleichheit/Ungleichheit von Verteilungen

gleiche Mittelwerte ? **Student's t-Test**

2 Zeitreihen

$$v(i) \quad i = 1, \dots, N_v \quad \overline{v}$$

$$w(i) \quad i = 1, \dots, N_w \quad \overline{w}$$

über gepoolte Varianzen:

$$s_D = \sqrt{\frac{\sum_{i=1}^{N_v}(v_i - \overline{v})^2 + \sum_{i=1}^{N_w}(w_i - \overline{w})^2}{N_v + N_w - 2}\left(\frac{1}{N_v} + \frac{1}{N_w}\right)}$$

Standardfehler der Mittelwertdifferenzen

Lineare Methoden

Gleichheit/Ungleichheit von Verteilungen

gleiche Mittelwerte ? **Student's t-Test**

$$t = \frac{\bar{v} - \bar{w}}{s_D}$$

Signifikanz für $N_v + N_w - 2$ Freiheitsgrade

- Tabellen
- unvollständige beta-Funktion

Lineare Methoden

Gleichheit/Ungleichheit von Verteilungen

gleiche Varianzen ? **F-Test**

2 Zeitreihen

$$v(i) \quad i = 1, \ldots, N_v$$

$$w(i) \quad i = 1, \ldots, N_w$$

$$F = \frac{Var(v_i)}{Var(w_i)}$$

Signifikanz für $N_v + N_w$ Freiheitsgrade

- Tabellen
- unvollständige beta-Funktion

Lineare Methoden

lineare Korrelationen:

parametrische Tests

- linearer Korrelationskoeffizient (Pearson's r)

nichtparametrische Tests

- Spearman Rangordnungs Korrelationskoeffizient
- Kendall's τ

Lineare Methoden

lineare Korrelationen: Pearson's *r*

2 Zeitreihen

$$v(i) \quad i = 1, \ldots, N_v \quad \bar{v}$$

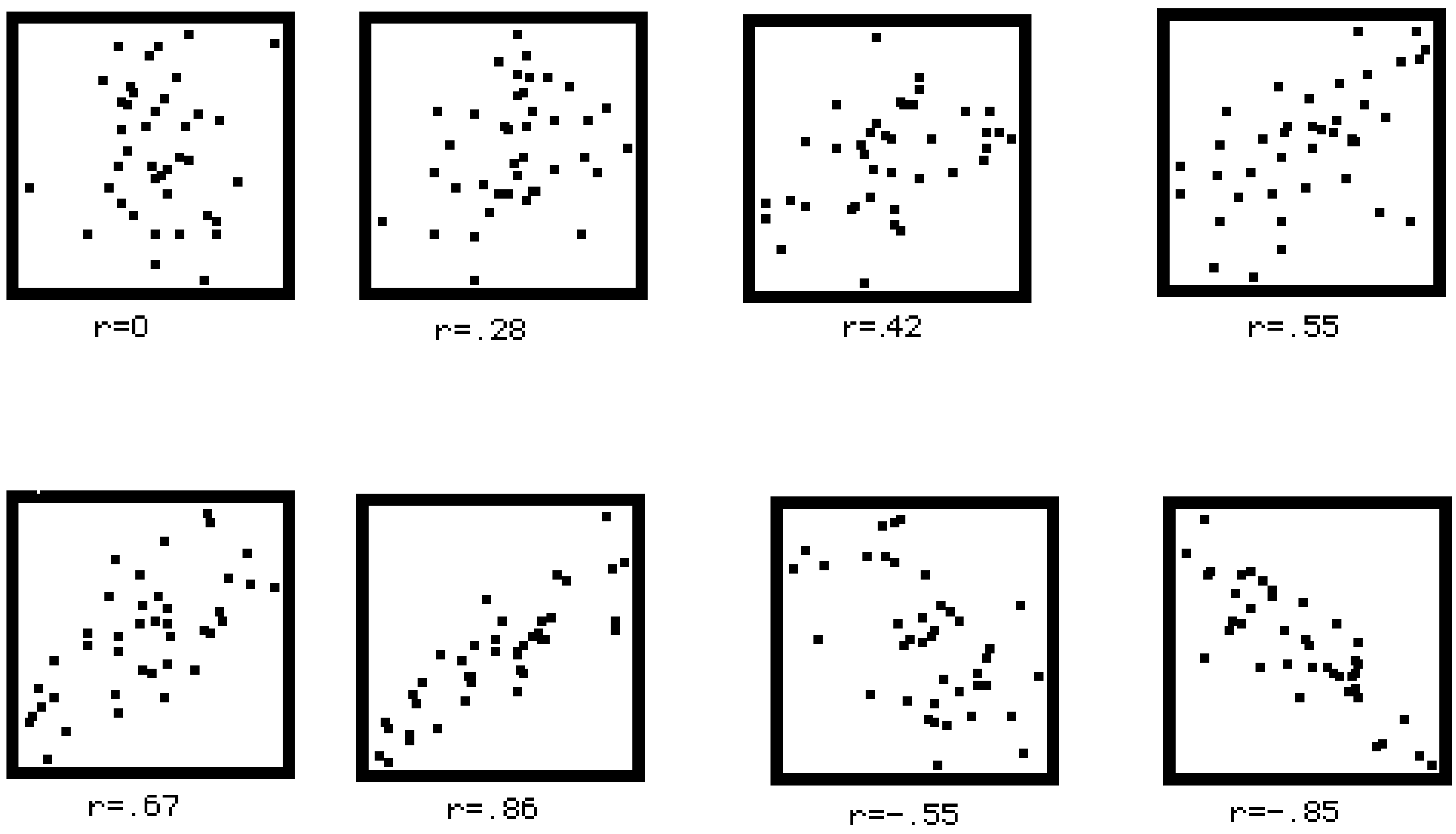
$$w(i) \quad i = 1, \ldots, N_w \quad \bar{w}$$

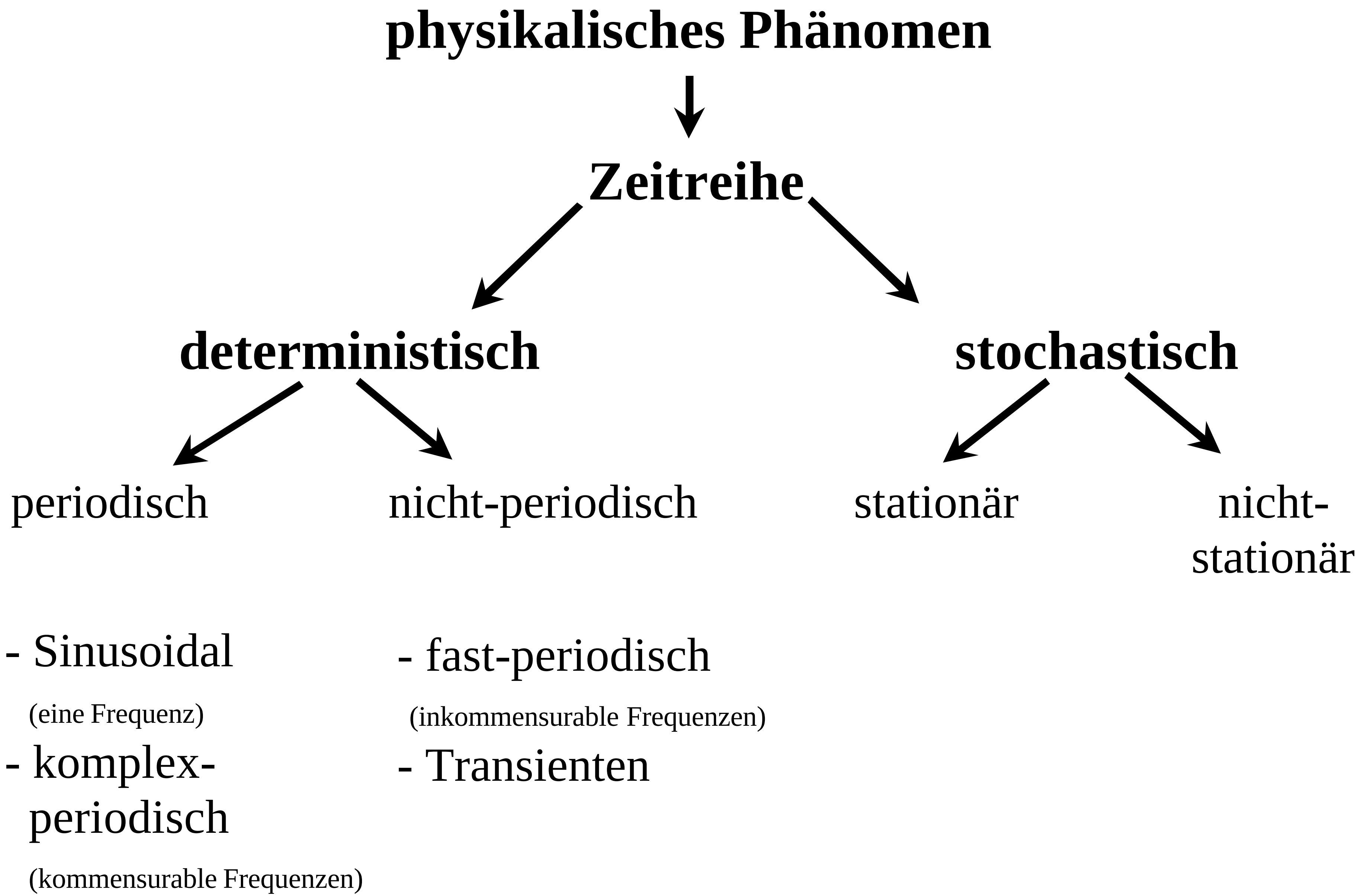
$$r = \frac{\sum_i (v_i - \bar{v})(w_i - \bar{w})}{\sqrt{\sum_i (v_i - \bar{v})^2} \sqrt{\sum_i (w_i - \bar{w})^2}}$$

$$r = \begin{cases} -1 & \text{vollständig antikorreliert} \\ 0 & \text{unkorreliert} \\ +1 & \text{vollständig korreliert} \end{cases}$$

Lineare Methoden

lineare Korrelationen: Pearson's *r*




physikalisches Phänomen
Zeitreihe
deterministisch
stochastisch
periodisch
nicht-periodisch
stationär
nicht-stationär
- Sinusoidal
(eine Frequenz)
- komplex-periodisch
(kommensurable Frequenzen)
- fast-periodisch
(inkommensurable Frequenzen)
- Transienten

Lineare Methoden

Auffinden verborgener periodischer Anteile

physikalischer Prozeß
(Zeitreihe)

Zeitbereich $\longleftrightarrow$ Frequenzbereich

v(t) V(f)

v = Auslenkung, Spannung, etc.

V= Amplitude
komplexe Zahl (Phase)

$-\infty < f < \infty$

Lineare Methoden

Autokorrelationsfunktion

gibt Hinweise darauf, wie Werte **einer** Zeitreihe in bestimmten Zeitabständen τ zusammenhängen

$$C_{vv}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T v(t)v(t+\tau)dt$$

$$C_{vv}(-\tau) = C_{vv}(\tau)$$

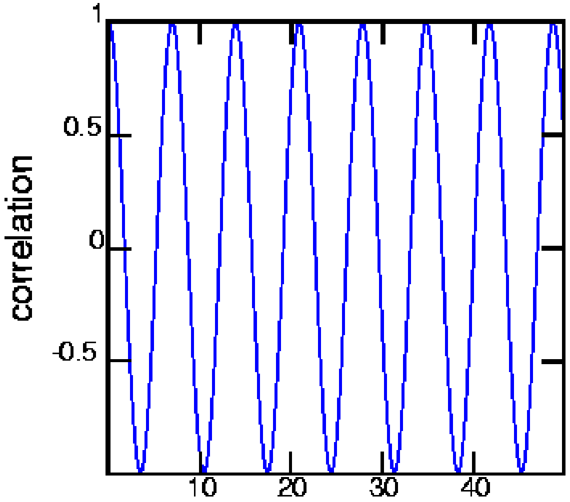
$$C_{vv}(0) \geq \left| C_{vv}(\tau) \right| \quad \forall\tau$$

$\tau = lag$

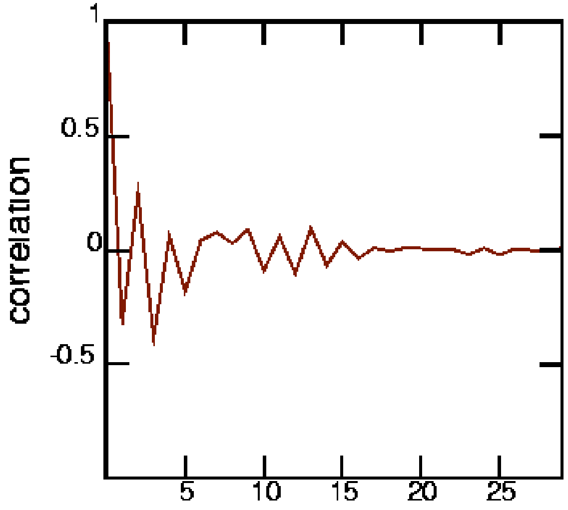
Lineare Methoden

Autokorrelationsfunktion

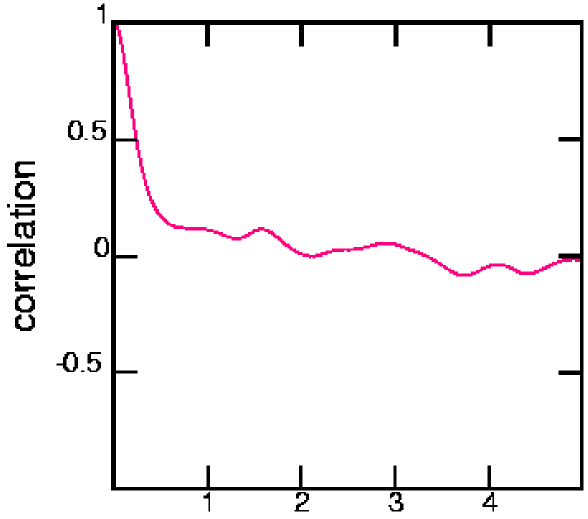
periodisches Signal



stochastisches Signal



Signal mit *Gedächtnis*



Lineare Methoden

Kreuzkorrelationsfunktion

gibt Hinweise darauf, wie Werte einer Zeitreihe mit Werten einer anderen Zeitreihe in bestimmten Zeitabständen τ zusammenhängen

$$C_{vw}(\tau) = \lim_{T\to\infty} \frac{1}{T} \int_0^T v(t)w(t+\tau)dt$$

$$C_{vw}(-\tau) = C_{wv}(\tau)$$

Lineare Methoden

Fourieranalysen

Approximation einer Funktion durch Summe von Sinus- und Kosinustermen (Fourierreihendarstellung). Eindeutig für deterministische Signale.

periodischer Prozeß

$$v(t) = \sum_{k=-\infty}^{\infty} V_k e^{-ik\omega t} \quad \leftrightarrow \quad V_k = \frac{1}{T}\int_0^T v(t) e^{ik\omega t} dt$$

nicht-periodischer Prozeß

$$v(t) = \int_{-\infty}^{\infty} V(\omega) e^{i\omega t} d\omega \quad \leftrightarrow \quad V(\omega) = \frac{1}{2\pi}\int_{-\infty}^{\infty} v(t) e^{i\omega t} dt$$

$$\omega = 2\pi f$$

Lineare Methoden

Fourieranalysen

Approximation einer Funktion durch Summe von Sinus- und Kosinustermen (Fourierreihendarstellung). Eindeutig für deterministische Signale.

$$v(t) = \sum_{k=0}^{\frac{N}{2}} a_k \cos(\omega_k t) + b_k \sin(\omega_k t) \quad \omega_k = \frac{2\pi k}{N}$$

$$a_k = \frac{2}{N} \sum_{t=1}^{N} \cos(\omega_k t) \quad k = 1,2,\ldots,\frac{N-1}{2}$$

$$b_k = \frac{2}{N} \sum_{t=1}^{N} v(t) \sin(\omega_k t)$$

$$v(t) = \sum_{k=-\frac{N-1}{2}}^{\frac{N-1}{2}} c_k e^{i\omega_k t} \quad c_k = \frac{1}{N} \sum_{t=1}^{N} v(t) e^{-i\omega_k t}$$

Bedingung:

$$\sum_{t=-\infty}^{\infty} |v(t)|^2 < \infty$$

Lineare Methoden

Fourieranalysen

Problem:
üblicherweise nur endlich viele Meßwerte (N)
aus endlichem Zeitintervall (Zeitfenster).

Lösung:
diskrete Fouriertransformation

$$V_k := \sum_{n=0}^{N-1} v(n\Delta t) e^{2\pi i n k/N}$$

$$v(n\Delta t) = \frac{1}{N} \sum_{k=0}^{N-1} V_k e^{-2\pi i n k/N}$$

Lineare Methoden

Fourieranalysen

Problem:

FT und DFT sind N^2-Algorithmen

Lösung:

Schnelle Fouriertransformation (FFT)

$N\log_2(N)$-Algorithmus

verschiedene Realisationsmöglichkeiten:

vgl. E.O. Brigham: The Fast Fourier Transform.
Prentice Hall, Englewood Cliffs, New Jersey, 1974

Lineare Methoden

einige nützliche Theoreme:

Faltungs-Theorem

$$v(t) * w(t) \Leftrightarrow V(f)W(f)$$

Korrelations-Theorem

$$C_{vw} \Leftrightarrow V(f)W^*(f)$$

Wiener-Khinchin-Theorem

$$C_{vv} \Leftrightarrow \left|V(f)\right|^2$$

Parseval' sches Theorem

$$Gesamtleistung \equiv \int_{-\infty}^{\infty} \left|v(t)\right|^2 dt = \int_{-\infty}^{\infty} \left|V(f)\right|^2 df$$

Lineare Methoden

Schätzung des Leistungsspektrum *P*(f)

Probleme:

- Varianz der Schätzung ist so groß wie die Schätzung selbst

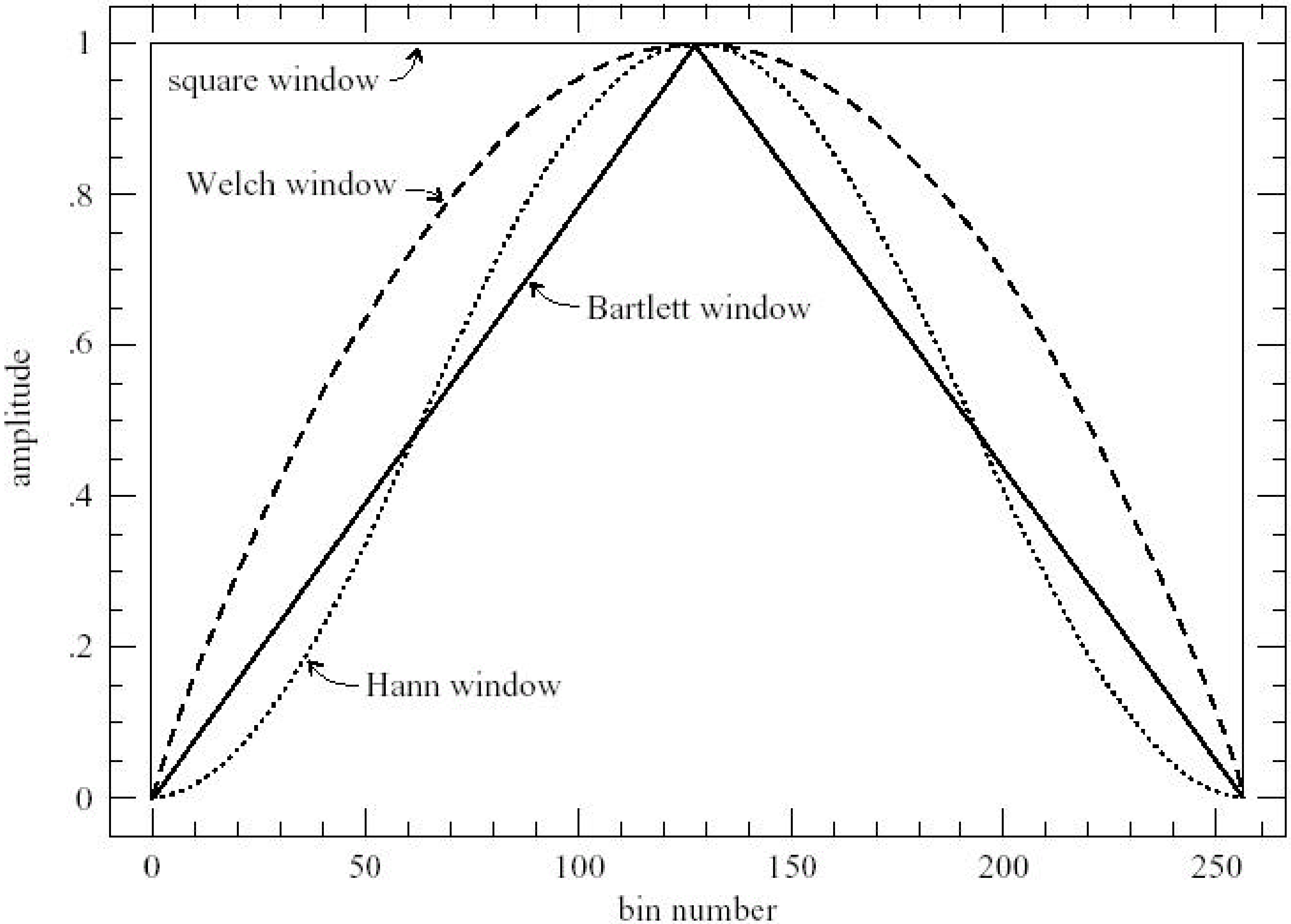
$\longrightarrow$ Zeitbereichs- oder Frequenzbereichsmittelung

- leakage
  diskreter Fall: Schätzung des Leistungsspektrum bei diskreten Frequenzen. Einfluß benachbarter Frequenzen ?

$\longrightarrow$ Windowing

Lineare Methoden

Schätzung des Leistungsspektrum *P*(f)
Fensterfunktionen



Stationarität

- wesentliche Voraussetzung für Beschreibung eines Systems

- gewährleistet Reproduzierbarkeit eines Meßergebnisses

- starke Stationarität / schwache Stationarität

Stationarität

Definition: starke Stationarität

Ein Prozeß wird **stationär** im strengen oder engeren Sinne genannt, wenn

$$\Psi(Z_{t_1},\ldots,Z_{t_n}) = \Psi(Z_{t_1+\tau},\ldots,Z_{t_n+\tau})$$

mit beliebigen $t_1,\ldots,t_n,\tau,n.$

$\Psi(Z_t) = \Psi(Z_{t_1},\ldots,Z_{t_n})$ Verbundverteilungsfunktion (statistischen Momente)

Eine Verschiebung des Zeitbeginns um τ hat keinen Einfluß auf die statistischen Momente und Verbundmomente.

Stationarität

Definition: schwache Stationarität

Ein Prozeß heißt **schwach stationär**, wenn seine Verbundmomente bis zur Ordnung 2 zeitunabhängig, also konstant sind.

$$\langle z(t) \rangle = \bar{z}(t) = \bar{z} = \text{konstant} \quad \text{(mittelwertstationär)}$$

$$\sigma(z(t)) = \sigma = \text{konstant} \quad \text{(varianzstationär)}$$

Lineare Methoden

stochastische Prozesse

- geordnete Menge von Zufallsvariablen

$$\{Z_t, t \in T\}$$

$\{T\}$ Menge der Zeitpunkte, bei denen der Prozeß definiert ist

- Zeitreihe $z_t \in Z_t$ ist Realisation eines stochastischen Prozesses

Lineare Methoden

stochastische Prozesse

Grundmodell: weißes Rauschen

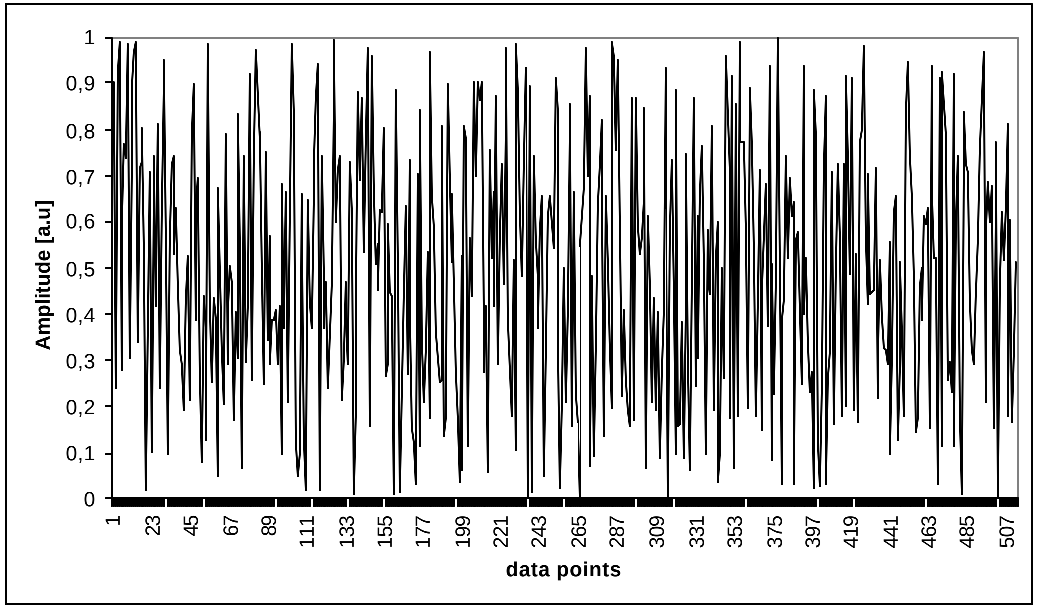
- identisch verteilt
- v(t) unabhängig ---> unkorreliert
- Autokorrelationsfunktion:

$$C_{vv} = \begin{cases} \sigma^2 & t = 0 \\ 0 & t \neq 0 \end{cases}$$

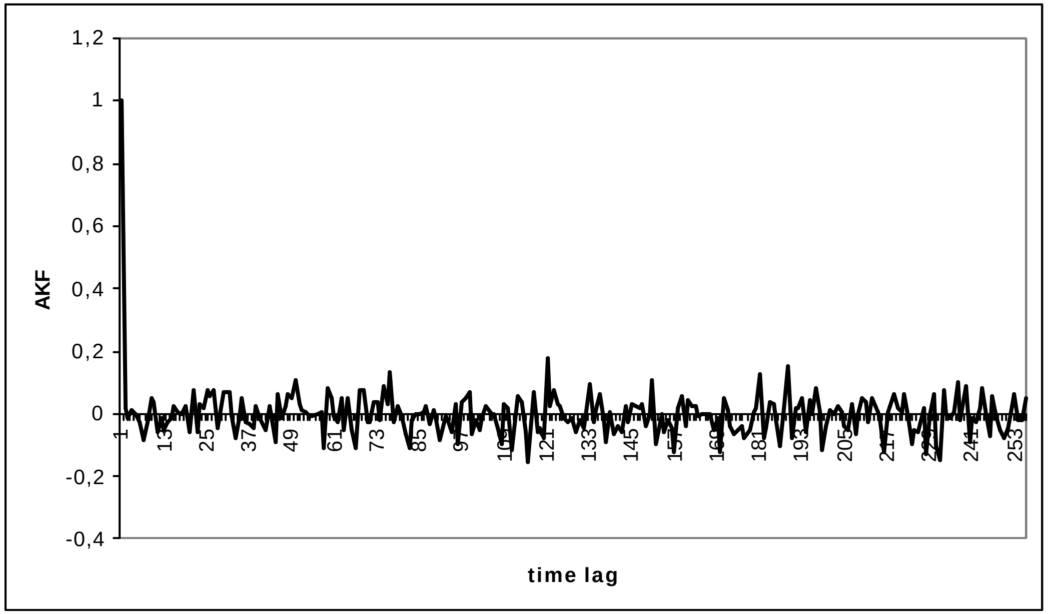
Im Leistungsspektrum sind alle Frequenzen gleichwertig vertreten (Analogie: weißes Licht)

weißes Rauschen

Zeitreihe



Autokorrelations-
funktion



Lineare Methoden

stochastische Prozesse

Autoregressive (AR) Prozesse *p*-ter Ordnung

$$v(t) = \sum_{l=1}^{p} a_l v(t-l) + \varepsilon(t)$$

ε(t) ist reiner Zufallsprozess mit Mittelwert 0 und Varianz σ^2
ist der Antrieb des Systems

v(t) hängt von seiner Vergangenheit bis zur Ordnung *p* ab
(autoregressiv)

Lineare Methoden

stochastische Prozesse

Autoregressive (AR) Prozesse 1-ter Ordnung (Markov-Prozeß)

$$v(t) = av(t-1) + \varepsilon(t)$$

- bei AR(1)-Prozessen nimmt die Autokorrelationsfunktion C_{vv} exponentiell ab

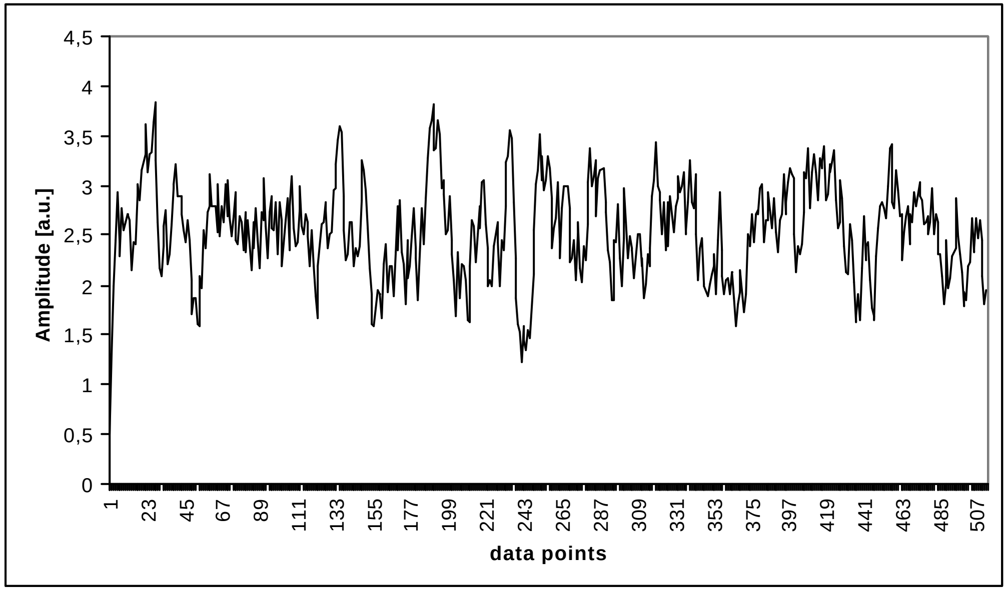
- Korrelationszeit ist die Zeit, nach der C_{vv} auf ihren e-ten Teil abgefallen ist

- für AR(1) Prozesse gilt: $\tau_{corr} = \dfrac{1}{\ln a}$

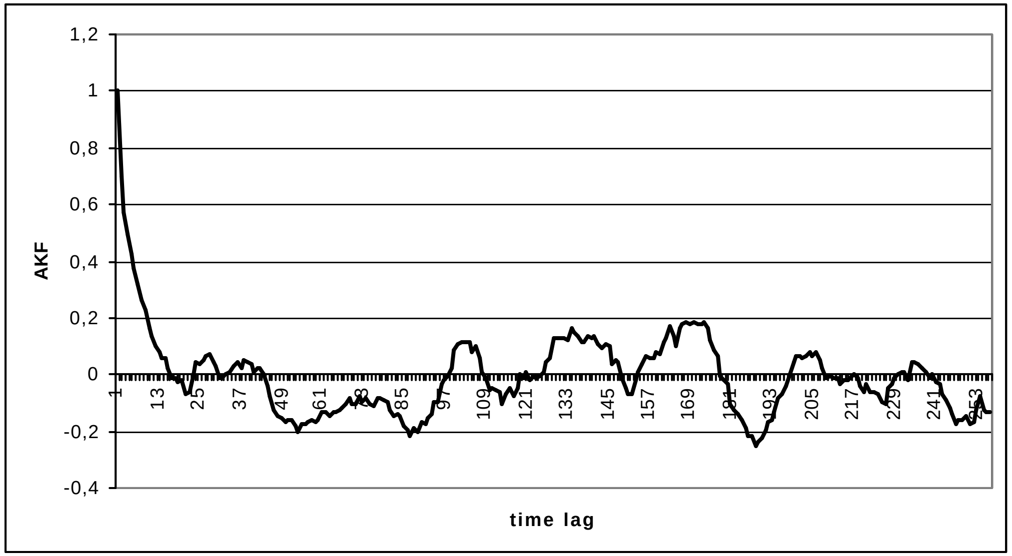
- Prozeß identifizierbar

AR(1)-Prozess: $v(t) = 0{,}8\ v(t-1) + \varepsilon(t)$

Zeitreihe



Autokorrelations-
funktion



Lineare Methoden

stochastische Prozesse

Autoregressive (AR) Prozesse 2-ter Ordnung

$$v(t) = a_1 v(t-1) + a_2 v(t-2) + \varepsilon(t)$$

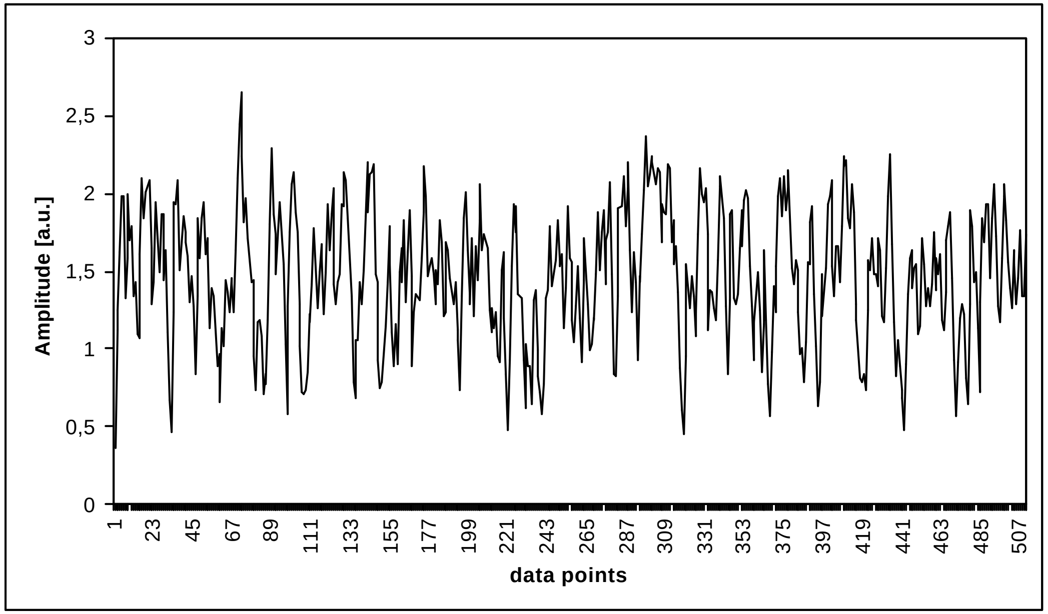
- bei AR-Prozessen höherer Ordnung kann die Autokorrelationsfunktion C_{vv} eine Mischung von gedämpfter Exponential- und sinusförmigen Funktionen sein.

- Prozeß schwer zu identifizieren

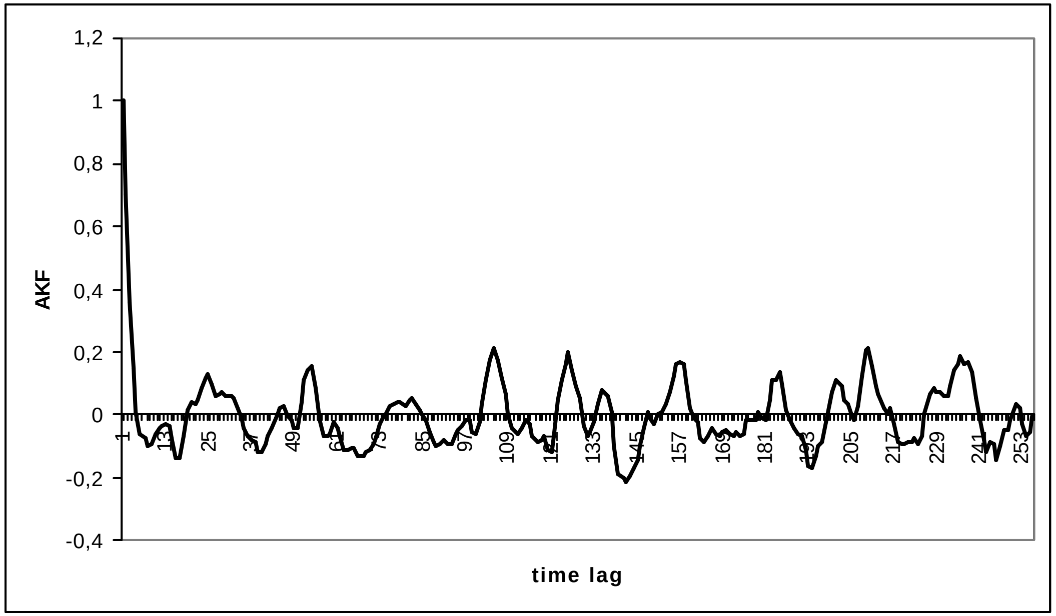
- rekursive Schätzung der Prozeßparameter *a* und *p*

AR(2)-Prozess: $v(t) = 0{,}85\ v(t-1) - 0{,}2\ v(t-2) + \varepsilon(t)$

Zeitreihe



Autokorrelations-funktion



Lineare Methoden

stochastische Prozesse

Erweiterung: Autoregressive (AR) Moving Average (MA) Prozesse *p,q*-ter Ordnung

$$v(t) = \sum_{l=1}^{p} a_l v(t-l) + \sum_{k=1}^{q} b_k \varepsilon(t-k)$$

Leistungsspektrum von ARMA-Prozessen wird durch rationale Funktionen von $e^{-2\pi i f}$ beschrieben:

$$V_{ARMA}(f) = const. \left| \frac{b(e^{-2\pi i f})}{a(e^{-2\pi i f})} \right|^2$$

--> Anwendung als Schätzer