Time Series Analysis

Henrik Madsen

hm@imm.dtu.dk

Informatics and Mathematical Modelling
Technical University of Denmark
DK-2800 Kgs. Lyngby

Outline of the lecture

Stochastic processes, 1st part:

- Stochastic processes in general: Sec 5.1, 5.2, 5.3 [except 5.3.2], 5.4.
- MA, AR, and ARMA-processes, Sec. 5.5

Stochastic Processes – in general

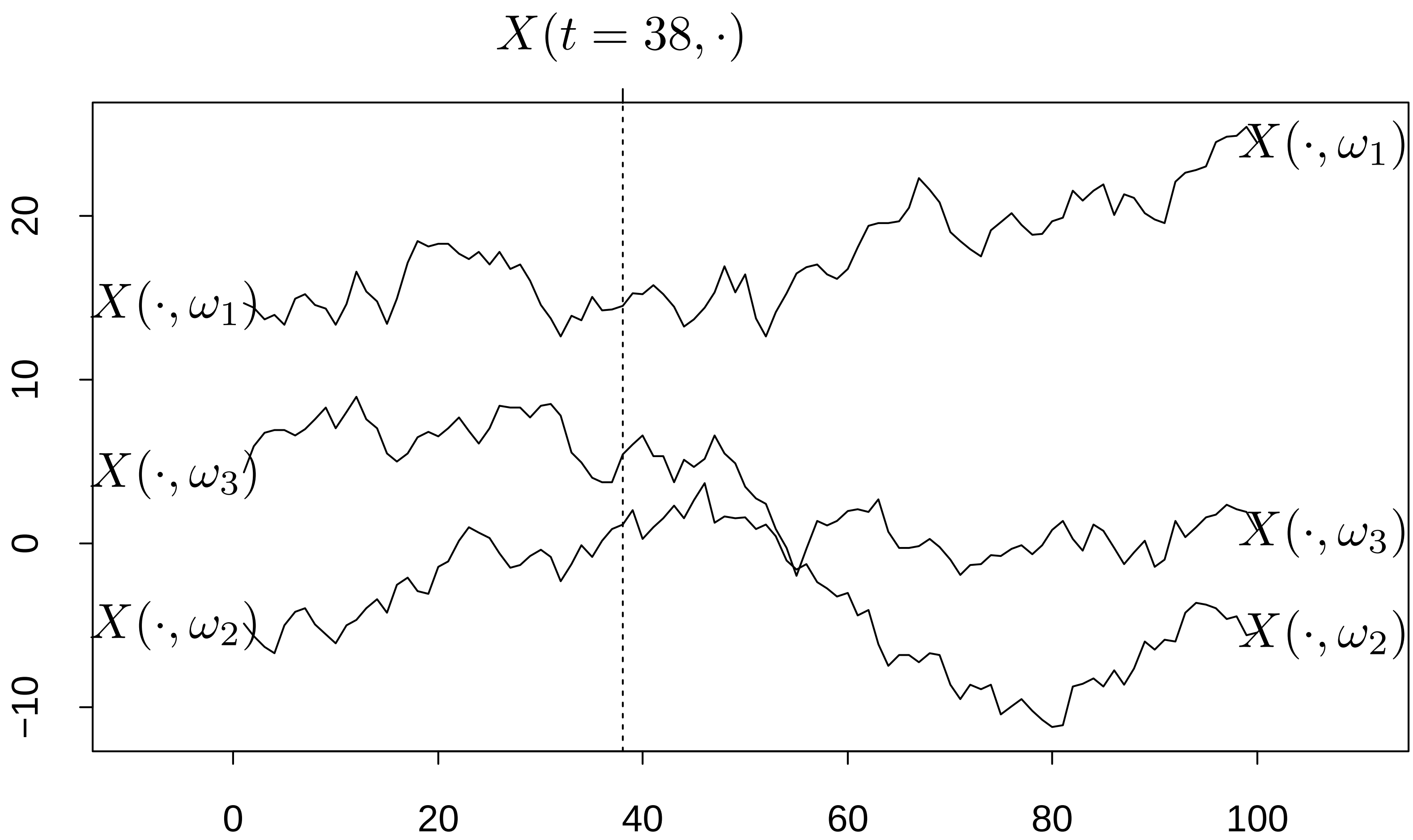
- Function: $X(t, \omega)$
- Time: $t \in T$
- Realization: $\omega \in \Omega$

- Index set: T
- Sample Space: Ω (sometimes called *ensemble*)

- $X(t = t_0, \cdot)$ is a random variable
- $X(\cdot, \omega)$ is a time series (i.e. *one* realization)

- In this course we consider the case where time is discrete and and measurements are continuous

Stochastic Processes – illustration



Complete Characterization

n-dimensional probability distribution:

$$f_{X(t_1),\ldots,X(t_n)}(x_1,\ldots,x_n)$$

Family of probability distribution functions, i.e.:

- For all $n = 1, 2, 3, \ldots$
- and all t

is called the *family of finite-dimensional probability distribution functions for the process*. This family completely characterize the stochastic process.

2'nd order moment representation

Mean function:

$$\mu(t) = E[X(t)] = \int_{-\infty}^{\infty} x f_{X(t)}(x)\, dx,$$

Autocovariance function:

$$\begin{aligned}\gamma_{XX}(t_1, t_2) &= \gamma(t_1, t_2) = Cov\big[X(t_1), X(t_2)\big] \\ &= E\big[(X(t_1) - \mu(t_1))(X(t_2) - \mu(t_2))\big]\end{aligned}$$

The variance function is obtained from $\gamma(t_1, t_2)$ when $t_1 = t_2 = t$:

$$\sigma^2(t) = V[X(t)] = E\big[(X(t) - \mu(t))^2\big]$$

Stationarity

- A process $\{X(t)\}$ is said to be *strongly stationary* if all finite-dimensional distributions are invariant for changes in time, i.e. for every n, and for any set $(t_1, t_2, \ldots, t_n)$ and for any h it holds

$$f_{X(t_1),\cdots,X(t_n)}(x_1, \cdots, x_n) = f_{X(t_1+h),\cdots,X(t_n+h)}(x_1, \cdots, x_n)$$

- A process $\{X(t)\}$ is said to be *weakly stationary of order k* if all the first k moments are invariant to changes in time
- A weakly stationary process of order 2 is simply called *weakly stationary* or just *stationary*:

$$\mu(t) = \mu \quad \sigma^2(t) = \sigma^2 \quad \gamma(t_1, t_2) = \gamma(t_1 - t_2)$$

Ergodicity

- In time series analysis we normally assume that we have access to one realization only
- We therefore need to be able to determine characteristics of the random variable X_t from one realization x_t
- It is often enough to require the process to be mean-ergodic:

$$E[X(t)] = \int_\Omega x(t,\omega) f(\omega)\, d\omega = \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} x(t,\omega)\, dt$$

i.e. if the *mean of the ensemble* equals the *mean over time*

Some intuitive examples, not directly related to time series:
`http://news.softpedia.com/news/What-is-ergodicity-15686.shtml`

Special processes

- *Normal processes* (also called *Gaussian processes*): All finite dimensional distribution functions are (multivariate) normal distributions
- *Markov processes*: The conditional distribution depends only of the latest state of the process:

$$\mathrm{P}\{X(t_n) \leq x | X(t_{n-1}), \cdots, X(t_1)\} = \mathrm{P}\{X(t_n) \leq x | X(t_{n-1})\}$$

- *Deterministic processes*: Can be predicted without uncertainty from past observations
- *Pure stochastic processes*: Can be written as a (infinite) linear combination of uncorrelated random variables
- Decomposition: $X_t = S_t + D_t$

Autocovariance and autocorrelation

- For stationary processes: Only dependent of the time difference $\tau = t_2 - t_1$
- Autocovariance:

$$\gamma(\tau) = \gamma_{XX}(\tau) = Cov[X(t), X(t+\tau)] = E[X(t)X(t+\tau)]$$

- Autocorrelation:

$$\rho(\tau) = \rho_{XX}(\tau) = \gamma_{XX}(\tau)/\gamma_{XX}(0) = \gamma_{XX}(\tau)/\sigma_X^2$$

- Some properties of the autocovariance function:
  - $\gamma(\tau) = \gamma(-\tau)$
  - $|\gamma(\tau)| = \gamma(0)$

Linear processes

- A linear process $\{Y_t\}$ is a process that can be written on the form

$$Y_t - \mu = \sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}$$

  where μ is the mean value of the process and
- $\{\varepsilon_t\}$ is white noise, i.e. a sequence of i.i.d. random variables.
- $\{\varepsilon_t\}$ can be scaled so that $\psi_0 = 1$
- Without loss of generality we assume $\mu = 0$

ψ- and π-weights

- Transfer function and linear process:

$$\psi(\mathsf{B}) = 1 + \sum_{i=1}^{\infty} \psi_i \mathsf{B}^i \qquad Y_t = \psi(\mathsf{B})\varepsilon_t$$

- Inverse operator (if it exists) and the linear process:

$$\pi(\mathsf{B}) = 1 + \sum_{i=1}^{\infty} \pi_i \mathsf{B}^i \qquad \pi(\mathsf{B})Y_t = \varepsilon_t,$$

- Autocovariance using ψ-weights:

$$\gamma(k) = Cov\left[\sum_{i=0}^{\infty} \psi_i \varepsilon_{t-i}, \sum_{i=0}^{\infty} \psi_i \varepsilon_{t+k-i}\right] = \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \psi_i \psi_{i+k}$$

Autocovariance Generating Function

- Let us define *autocovariance generating function*:

$$\Gamma(z) = \sum_{k=-\infty}^{\infty} \gamma(k) z^{-k}, \qquad (1)$$

which is the z–transformation of the autocovariance function.

Autocovariance Generating Function

- We obtain (since $\psi_i = 0$ for $i < 0$)

$$\begin{aligned} \Gamma(z) &= \sigma_\varepsilon^2 \sum_{k=-\infty}^{\infty} \sum_{i=0}^{\infty} \psi_i \psi_{i+k} z^{-k} \\ &= \sigma_\varepsilon^2 \sum_{i=0}^{\infty} \psi_i z^i \sum_{j=0}^{\infty} \psi_j z^{-j} \\ &= \sigma_\varepsilon^2 \psi(z^{-1}) \psi(z). \end{aligned}$$

- $$\Gamma(z) = \sigma_\varepsilon^2 \psi(z^{-1}) \psi(z) = \sigma_\varepsilon^2 \pi^{-1}(z^{-1}) \pi^{-1}(z). \tag{2}$$

Stationarity and invertibility

- The linear process $Y_t = \psi(\mathrm{B})\varepsilon_t$ is *stationary* if

$$\psi(z) = \sum_{i=0}^{\infty} \psi_i z^{-i}$$

converges for $|z| \geq 1$ (i.e. old values of ε_t is weighted down)

- The linear process $\pi(\mathrm{B})Y_t = \varepsilon_t$ is said to be *invertible* if

$$\pi(z) = \sum_{i=0}^{\infty} \pi_i z^{-i}$$

converges for $|z| \geq 1$ (i.e. ε_t can be calculated from recent values of Y_t)

Stationary processes in the frequency domain

- It has been shown that the autocovariance function is non-negative definite.
- Following a theorem of Bochner such a non-negative definite function can be written as a Stieltjes integral

$$\gamma(\tau) = \int_{-\infty}^{\infty} e^{i\omega\tau} \, dF(\omega) \tag{3}$$

for a process in continuous time, or

$$\gamma(\tau) = \int_{-\pi}^{\pi} e^{i\omega\tau} \, dF(\omega) \tag{4}$$

for a process in discrete time.

Processes in the frequency domain

For a purely stochastic process we have the following relations between the spectrum and the autocovariance function

$$\begin{aligned} f(\omega) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} e^{-i\omega\tau}\gamma(\tau)\,d\tau \\ \gamma(\tau) &= \int_{-\infty}^{\infty} e^{i\omega\tau} f(\omega)\,d\omega \end{aligned} \quad \text{(continuous time)} \quad (5)$$

$$\begin{aligned} f(\omega) &= \frac{1}{2\pi}\sum_{k=-\infty}^{\infty} \gamma(k) e^{-i\omega k} \\ \gamma(k) &= \int_{-\pi}^{\pi} e^{ik\omega} f(\omega)\,d\omega \end{aligned} \quad \text{(discrete time)} \quad (6)$$

Processes in the frequency domain

- We have seen that any stationary process can be formulated as a sum of a purely stochastic process and a purely deterministic process.
- Similar, the spectral density can be written

$$F(\omega) = F_S(\omega) + F_D(\omega), \tag{7}$$

where $F_S(\omega)$ is an even continuous function and $F_D(\omega)$ is a step function.

Processes in the frequency domain

- For a pure deterministic process

$$Y_t = \sum_{i=1}^{k} A_i \cos(\omega_i t + \phi_i), \tag{8}$$

  F_S will become 0, and thus $F(\omega)$ will become a step function with steps at the frequencies $\pm\omega_i$, $i = 1, \ldots, k$.

- In this case F can be written as

$$F(\omega) = F_D(\omega) = \sum_{\omega_i \leq \omega} f(\omega_i) \tag{9}$$

  and $\{f(\omega_i); i = 1, \ldots, k\}$ is often called the *line spectrum*.

Linear process as a statistical model?

$$Y_t = \varepsilon_t + \psi_1 \varepsilon_{t-1} + \psi_2 \varepsilon_{t-2} + + \psi_3 \varepsilon_{t-3} + \dots$$

- Observations: $Y_1, Y_2, Y_3, \dots, Y_N$
- Task: Find an infinite number of parameters from N observations!
- Solution: Restrict the sequence $1, \psi_1, \psi_2, \psi_3, \dots$

$MA(q)$, $AR(p)$, and $ARMA(p,q)$ processes

$$
\begin{aligned}
Y_t &= \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} \\
Y_t + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} &= \varepsilon_t \\
Y_t + \phi_1 Y_{t-1} + \cdots + \phi_p Y_{t-p} &= \varepsilon_t + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q}
\end{aligned}
$$

$\{\varepsilon_t\}$ is white noise

$$
\begin{aligned}
Y_t &= \theta(\mathsf{B})\varepsilon_t \\
\phi(\mathsf{B}) Y_t &= \varepsilon_t \\
\phi(\mathsf{B}) Y_t &= \theta(\mathsf{B})\varepsilon_t
\end{aligned}
$$

$\phi(\mathsf{B})$ and $\theta(\mathsf{B})$ are polynomials in the backward shift operator B, ($\mathsf{B} X_t = X_{t-1}$, $\mathsf{B}^2 X_t = X_{t-2}$)

Stationarity and invertibility

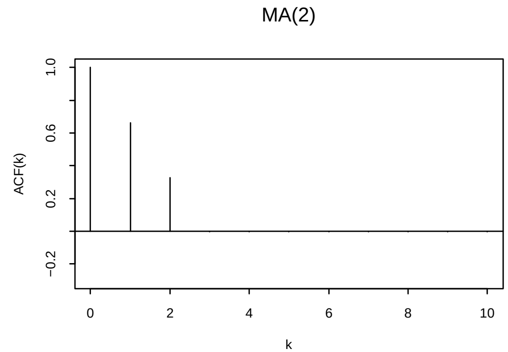
- $MA(q)$
  - Always stationary
  - Invertible if the roots in $\theta(z^{-1}) = 0$ with respect to z all are within the unit circle
- $AR(p)$
  - Always invertible
  - Stationary if the roots of $\phi(z^{-1}) = 0$ with respect to z all lie within the unit circle
- $ARMA(p, q)$
  - Stationary if the roots of $\phi(z^{-1}) = 0$ with respect to z all lie within the unit circle
  - Invertible if the roots in $\theta(z^{-1}) = 0$ with respect to z all are within the unit circle

Autocorrelations

MA(2):

$Y_t = (1 + 0.9\text{B} + 0.8\text{B}^2)\varepsilon_t$

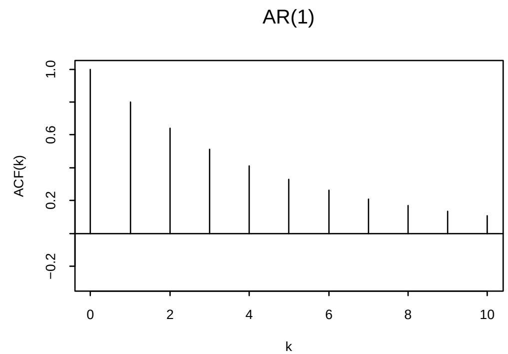
zero after lag 2



AR(1):

$(1 - 0.8\text{B})Y_t = \varepsilon_t$

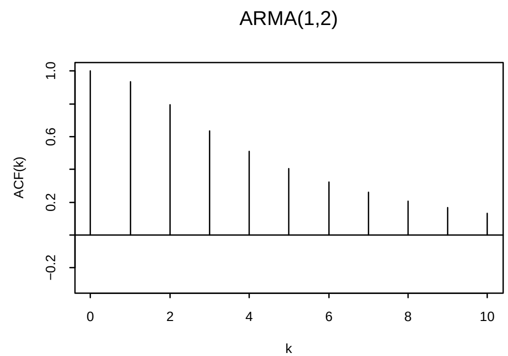
exponential decay (damped sine in case of complex roots)



ARMA(1,2):

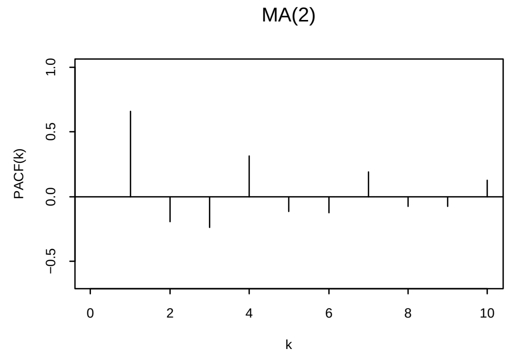
$(1 - 0.8\text{B})Y_t = (1 + 0.9\text{B} + 0.8\text{B}^2)\varepsilon_t$

exponential decay from lag $q+1-p = 2+1-1 = 2$ (damped sine in case of complex roots)

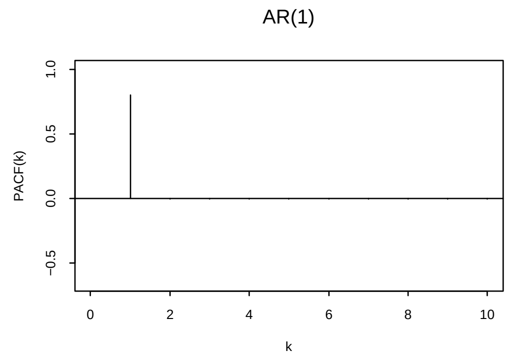


Partial autocorrelations

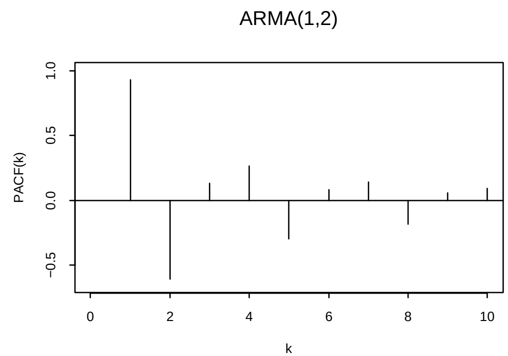
MA(2):
$Y_t = (1 + 0.9\mathsf{B} + 0.8\mathsf{B}^2)\varepsilon_t$



AR(1):
$(1 - 0.8\mathsf{B})Y_t = \varepsilon_t$
zero after lag 1



ARMA(1,2):
$(1 - 0.8\mathsf{B})Y_t = (1 + 0.9\mathsf{B} + 0.8\mathsf{B}^2)\varepsilon_t$



Inverse autocorrelation

- The process: $\phi(\mathsf{B})Y_t = \theta(\mathsf{B})\varepsilon_t$
- The dual process: $\theta(\mathsf{B})Z_t = \phi(\mathsf{B})\varepsilon_t$
- The inverse autocorrelation is the autocorrelation for the dual process
- Thus, the IACF can be used i a similar way as the PACF