Time Series Analysis

Week 7 – Linear systems

Peder Bacher
Department of Applied Mathematics and Computer Science
Technical University of Denmark

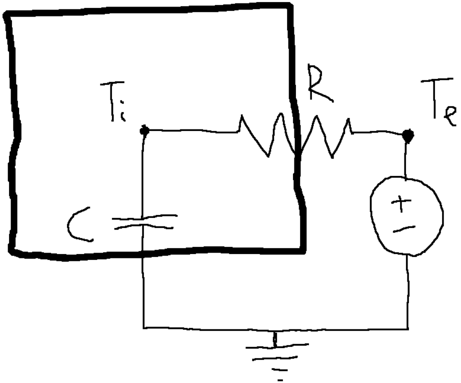
March 15, 2024

Week 7: Outline of the lecture

- Input-Output systems, sec. 4 introduction and 4.1
- Linear system notation
- The z-transform, section 4.4
- Cross Correlation Functions – from Sec. 6.2.2
- Transfer function models; identification, estimation, validation, prediction, Chap. 8

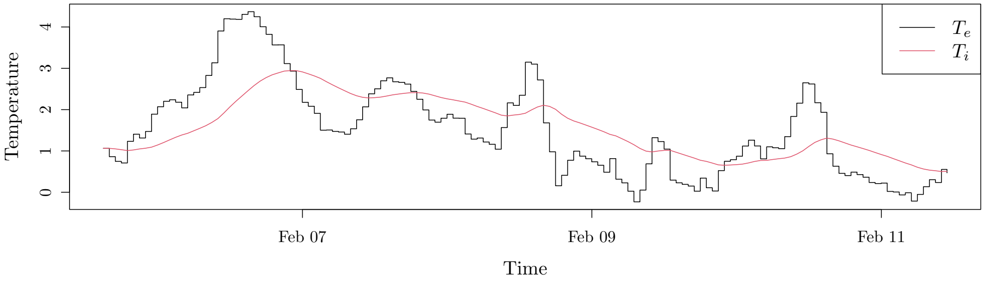
Simplest first order RC-system

Single state model of the temperature in a box:



Simplest RC-system

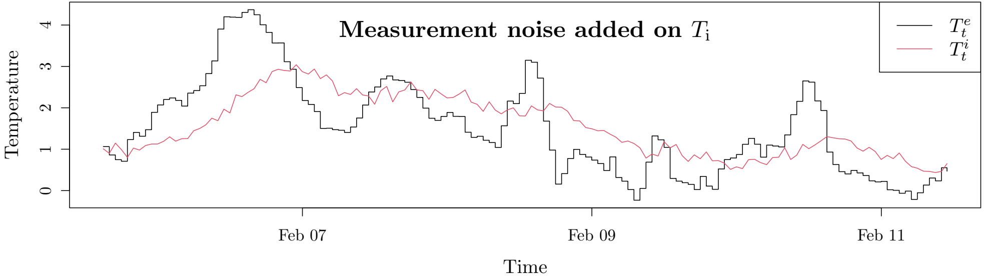
- T_t^{e} external and T_t^{i} internal temperature at time $t = [1, 2, \ldots, n]$
- ODE model
$$\frac{dT_{\mathsf{i}}}{dt} = \frac{1}{RC}(Te - T_{\mathsf{i}})$$



Simplest RC-system

- T_t^{e} external and T_t^{i} internal temperature at time $t = [1, 2, \ldots, n]$
- ODE model

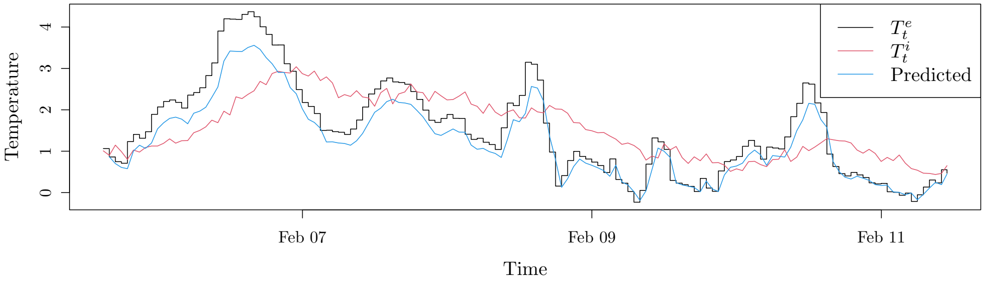
$$\frac{dT_{\mathrm{i}}}{dt} = \frac{1}{RC}(Te - T_{\mathrm{i}})$$



Try a static model

- A simple linear regression model (ε_t is the error)
- Not describing dynamics

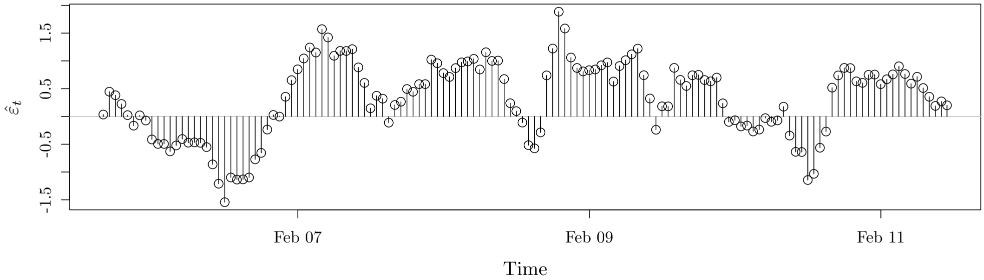
$$T_t^{\mathrm{i}} = \omega_e T_t^{\mathrm{e}} + \varepsilon_t$$



Model validation: check i.i.d. of residuals

Are residuals like white noise?

- Check if they are *independent and identically distributed*
- Is $\hat{\varepsilon}_t$ independent of $\hat{\varepsilon}_{t-k}$ for all t and k?

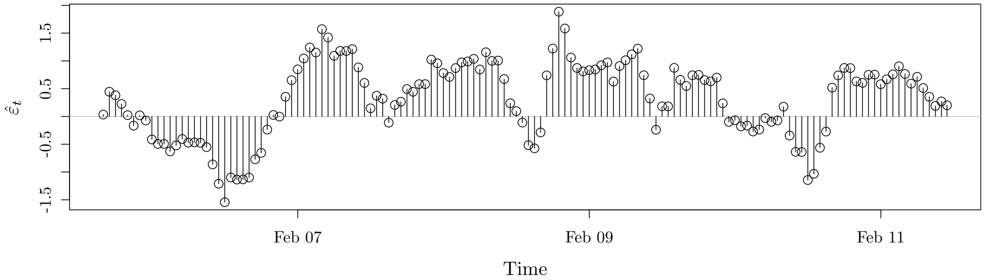


Model validation: check i.i.d. of residuals

Are residuals like white noise?

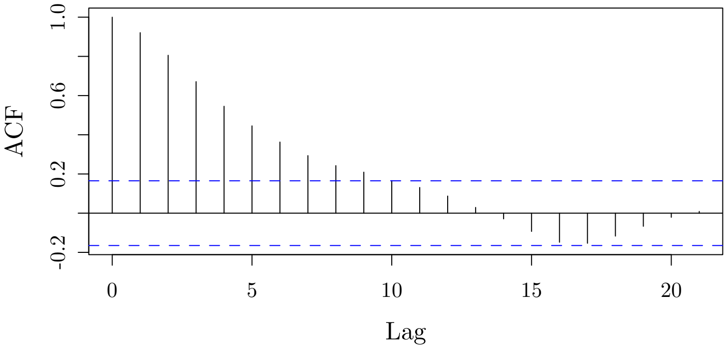
- Check if they are *independent and identically distributed*
- Is $\hat{\varepsilon}_t$ independent of $\hat{\varepsilon}_{t-k}$ for all t and k?

Nope! There is a pattern left...



Model validation: Test for i.i.d. with ACF

TEST if residuals are white noise?



It's not white nose! How do we find a better model?

Discretize the ODE

$$\frac{dT_{\mathrm{i}}}{dt} = \frac{1}{RC}(T_{\mathrm{e}} - T_{\mathrm{i}})$$

It has the solution

$$T_{\mathrm{i}}(t + \Delta t) = T_{\mathrm{e}}(t) + e^{-\frac{\Delta t}{RC}}\big(T_{\mathrm{i}}(t) - T_{\mathrm{e}}(t)\big)$$

Discretize the ODE

$$\frac{dT_{\mathrm{i}}}{dt} = \frac{1}{RC}(T_{\mathrm{e}} - T_{\mathrm{i}})$$

It has the solution

$$T_{\mathrm{i}}(t + \Delta t) = T_{\mathrm{e}}(t) + e^{-\frac{\Delta t}{RC}}\big(T_{\mathrm{i}}(t) - T_{\mathrm{e}}(t)\big)$$

if $\Delta t = 1$ and T_{e} is constant between the sample points then

$$T^{\mathrm{i}}_{t+1} = e^{-\frac{1}{RC}} T^{\mathrm{i}}_t + (1 - e^{-\frac{1}{RC}}) T^{\mathrm{e}}_t$$

Discretize the ODE

$$\frac{dT_{\mathrm{i}}}{dt} = \frac{1}{RC}(T_{\mathrm{e}} - T_{\mathrm{i}})$$

It has the solution

$$T_{\mathrm{i}}(t + \Delta t) = T_{\mathrm{e}}(t) + e^{-\frac{\Delta t}{RC}}\big(T_{\mathrm{i}}(t) - T_{\mathrm{e}}(t)\big)$$

if $\Delta t = 1$ and T_{e} is constant between the sample points then

$$T^{\mathrm{i}}_{t+1} = e^{-\frac{1}{RC}}\, T^{\mathrm{i}}_t + (1 - e^{-\frac{1}{RC}})\, T^{\mathrm{e}}_t$$

since $e^{-\frac{1}{RC}}$ is between 0 and 1, then write it as

$$T^{\mathrm{i}}_{t+1} = \phi_1\, T^{\mathrm{i}}_t + \omega_1\, T^{\mathrm{e}}_t$$

where ϕ_1 and ω_1 are between 0 and 1.

Discretize the ODE

$$\frac{dT_{\mathrm{i}}}{dt} = \frac{1}{RC}(T_{\mathrm{e}} - T_{\mathrm{i}})$$

It has the solution

$$T_{\mathrm{i}}(t + \Delta t) = T_{\mathrm{e}}(t) + e^{-\frac{\Delta t}{RC}}\left(T_{\mathrm{i}}(t) - T_{\mathrm{e}}(t)\right)$$

if $\Delta t = 1$ and T_{e} is constant between the sample points then

$$T^{\mathrm{i}}_{t+1} = e^{-\frac{1}{RC}}\, T^{\mathrm{i}}_t + (1 - e^{-\frac{1}{RC}})\, T^{\mathrm{e}}_t$$

since $e^{-\frac{1}{RC}}$ is between 0 and 1, then write it as

$$T^{\mathrm{i}}_{t+1} = \phi_1\, T^{\mathrm{i}}_t + \omega_1\, T^{\mathrm{e}}_t$$

where ϕ_1 and ω_1 are between 0 and 1.

Add a noise term and we have the ARX model

$$T^{\mathrm{i}}_{t+1} = \phi_1\, T^{\mathrm{i}}_t \quad + \omega_1\, T^{\mathrm{e}}_t \quad + \varepsilon_{t+1}$$

Discretize the ODE

$$\frac{dT_{\rm i}}{dt} = \frac{1}{RC}(T_{\rm e} - T_{\rm i})$$

It has the solution

$$T_{\rm i}(t+\Delta t) = T_{\rm e}(t) + e^{-\frac{\Delta t}{RC}}\big(T_{\rm i}(t) - T_{\rm e}(t)\big)$$

if $\Delta t = 1$ and $T_{\rm e}$ is constant between the sample points then

$$T^{\rm i}_{t+1} = e^{-\frac{1}{RC}}\, T^{\rm i}_t + (1 - e^{-\frac{1}{RC}})\, T^{\rm e}_t$$

since $e^{-\frac{1}{RC}}$ is between 0 and 1, then write it as

$$T^{\rm i}_{t+1} = \phi_1\, T^{\rm i}_t + \omega_1\, T^{\rm e}_t$$

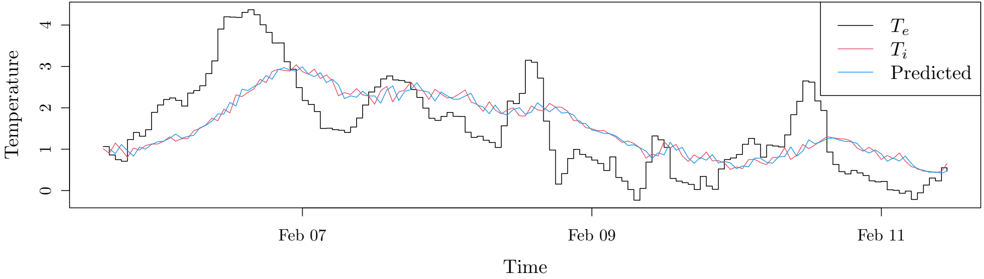
where ϕ_1 and ω_1 are between 0 and 1.

Add a noise term and we have the ARX model

$$T^{\rm i}_t = \phi_1\, T^{\rm i}_{t-1} + \omega_1\, T^{\rm e}_{t-1} + \varepsilon_t$$

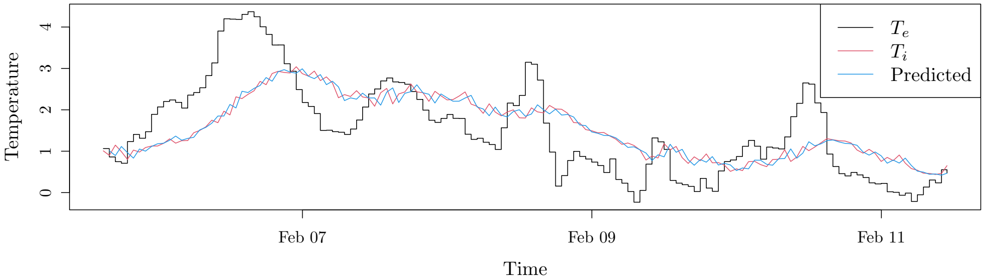
An ARX model

$$T_t^{\mathrm{i}} = \phi_1 T_{t-1}^{\mathrm{i}} + \omega_1 T_{t-1}^{\mathrm{e}} + \varepsilon_t$$



An ARX model

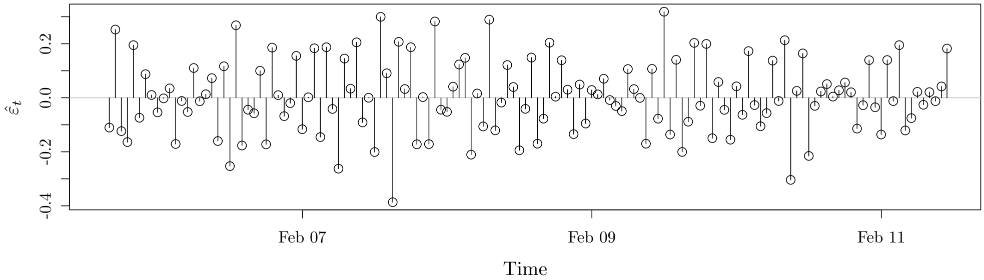
$$T_t^{\mathrm{i}} = \phi_1 \, T_{t-1}^{\mathrm{i}} + \omega_1 \, T_{t-1}^{\mathrm{e}} + \varepsilon_t$$



ARX model

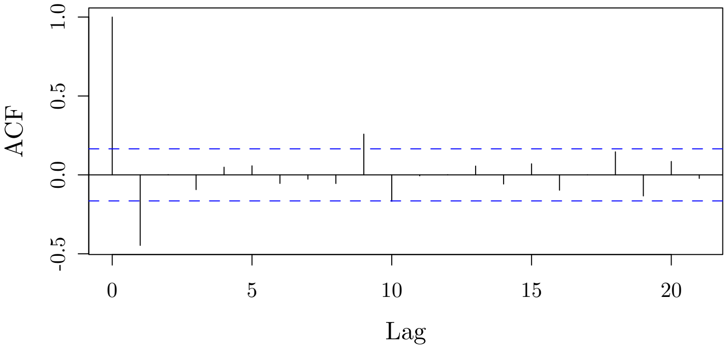
The residuals

$$\hat{\varepsilon}_t = T_t^{\mathrm{i}} - \frac{\hat{\omega}_1 \mathrm{B}}{1 - \hat{\phi}_1 \mathrm{B}} T_t^{\mathrm{e}}$$



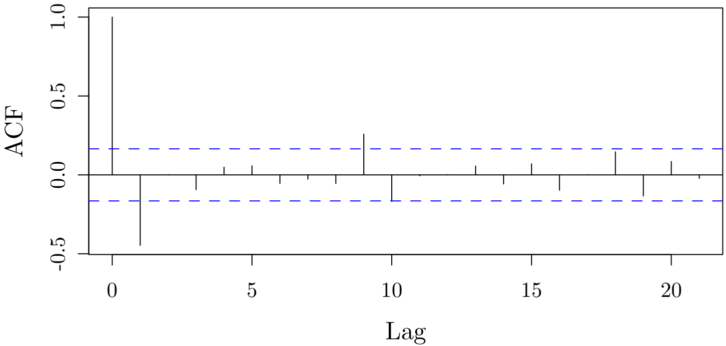
Check for i.i.d. of residuals

Is it likely that this is white noise?



Check for i.i.d. of residuals

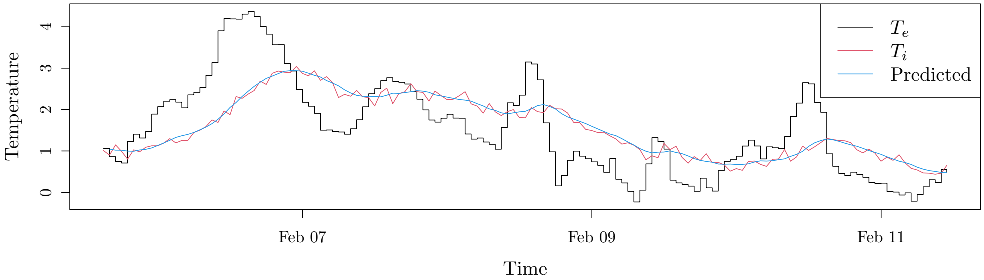
Is it likely that this is white noise? Almost!



Actually we miss an MA part!

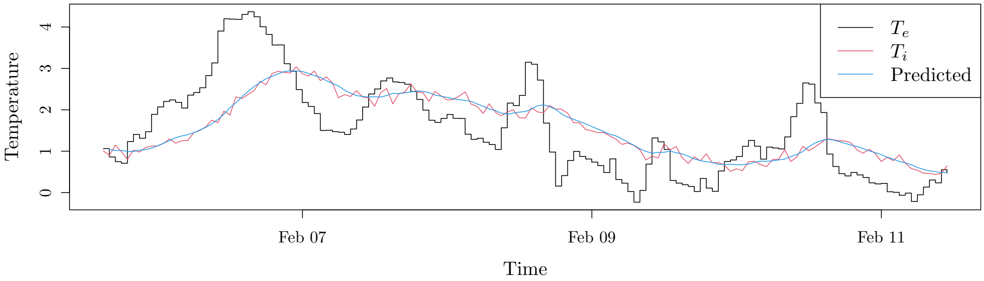
An ARMAX model

$$T_t^{\mathrm{i}} = \phi_1 T_{t-1}^{\mathrm{i}} + \omega_1 T_t^{\mathrm{e}} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

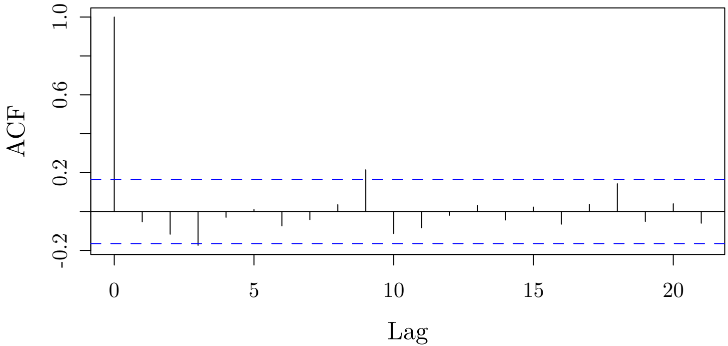


An ARMAX model

$$T_t^{\mathrm{i}} = \phi_1 T_{t-1}^{\mathrm{i}} + \omega_1 T_t^{\mathrm{e}} + \varepsilon_t + \theta_1 \varepsilon_{t-1}$$

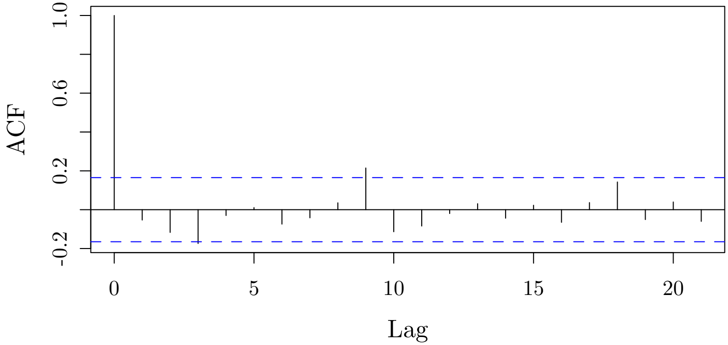


Validate the model with the residuals ACF



Now we have *white noise residuals*, that is want to have after applying the model!

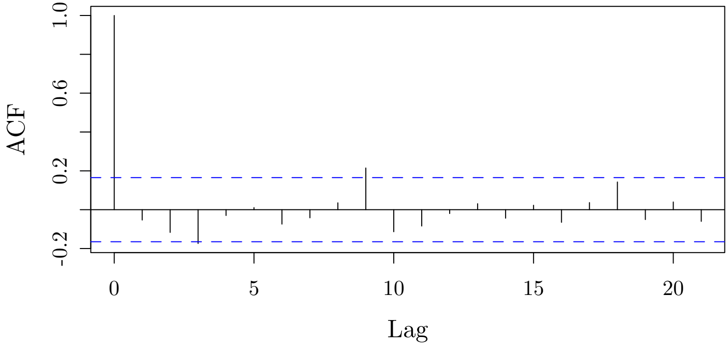
Validate the model with the residuals ACF



Now we have *white noise residuals*, that is want to have after applying the model!

Remember, we are validating the *one-step prediction* residuals: $\hat{\varepsilon}_{t+1} = y_{t+1} - \hat{y}_{t+1|t}$

Validate the model with the residuals ACF



Now we have *white noise residuals*, that is want to have after applying the model!

Remember, we are validating the *one-step prediction* residuals: $\hat{\varepsilon}_t = y_t - \hat{y}_{t|t-1}$

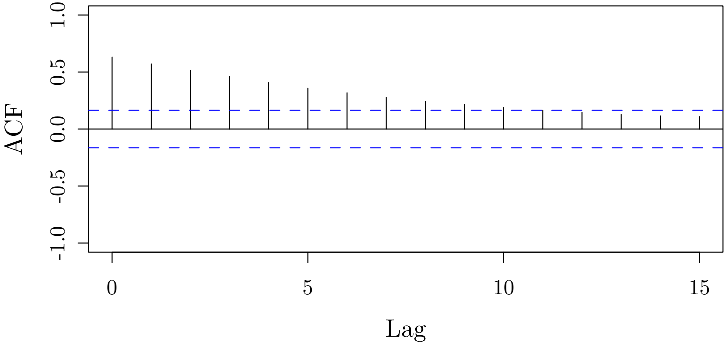
Dependence between variables: Cross-correlation function

Simply shift the index to lag *another* series:

t	T^{e}_t	T^{i}_t	T^{i}_{t-1}
1	4	2	
2	5	3	2
3	2	8	3
4	3	3	8
5	4	1	3
6	5	7	1
7	5	8	7
8			8

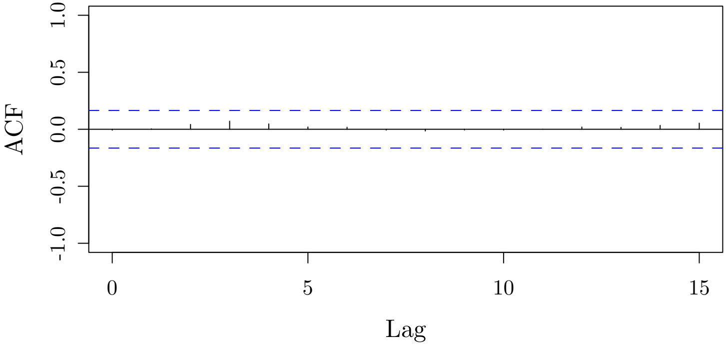
Cross-correlation function

Cross-Correlation Function (CCF) between T_t^{e} and T_t^{i}

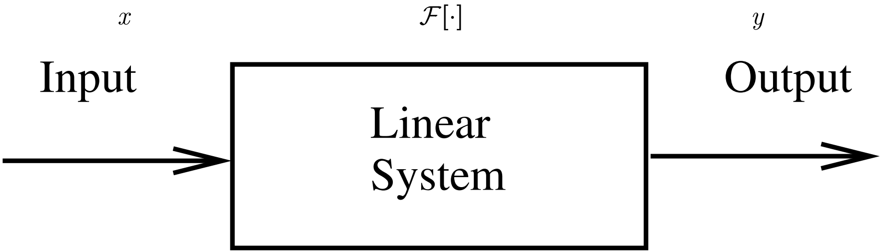


Cross-correlation function

Cross-Correlation Function (CCF) between T_t^{e} and ε_t



Linear Dynamic Systems – notation



$x(t)$	Differential eq., $h(u)$	$y(t)$
x_t	Difference eq., h_k, $h(B)$	y_t
$X(\omega)$	$\mathcal{H}(\omega)$	$Y(\omega)$
$X(z)$	$H(z)$	$Y(z)$

Description in the time domain (Convolution)

For *linear time invariant systems*:

- Discrete time:

$$y_t = \sum_{k=-\infty}^{\infty} h(k) x_{t-k} \tag{1}$$

  Think about

$$y_t = \sum_{k=0}^{t} h(k) x_{t-k} \tag{2}$$

- $h(k)$ is called the *impulse response*

Description in the time domain (Convolution)

For *linear time invariant systems*:

- Discrete time:

$$y_t = \sum_{k=-\infty}^{\infty} h(k) x_{t-k} \quad (1)$$

  Think about

$$y_t = \sum_{k=0}^{t} h(k) x_{t-k} \quad (2)$$

- $h(k)$ is called the *impulse response*, why?

Description in the time domain (Convolution)

For *linear time invariant systems*:

- Discrete time:

$$y_t = \sum_{k=-\infty}^{\infty} h(k) x_{t-k} \tag{1}$$

  Think about

$$y_t = \sum_{k=0}^{t} h(k) x_{t-k} \tag{2}$$

- $h(k)$ is called the *impulse response*, why? What happens if $x_0 = 1$ and $x_k = 0$ for $k \neq 0$?

Description in the time domain (Convolution)

For *linear time invariant systems*:

- Discrete time:
$$y_t = \sum_{k=-\infty}^{\infty} h(k) x_{t-k} \tag{1}$$
  Think about
$$y_t = \sum_{k=0}^{t} h(k) x_{t-k} \tag{2}$$
- $h(k)$ is called the *impulse response*, why? What happens if $x_0 = 1$ and $x_k = 0$ for $k \neq 0$?
- $S_k = \sum_{j=-\infty}^{k} h_j$ is called the *step response*

Description in the time domain (Convolution)

For *linear time invariant systems*:

- Discrete time:

$$y_t = \sum_{k=-\infty}^{\infty} h(k)x_{t-k} \tag{1}$$

  Think about

$$y_t = \sum_{k=0}^{t} h(k)x_{t-k} \tag{2}$$

- $h(k)$ is called the *impulse response*, why? What happens if $x_0 = 1$ and $x_k = 0$ for $k \neq 0$?
- $S_k = \sum_{j=-\infty}^{k} h_j$ is called the *step response*, why?

Description in the time domain (Convolution)

For *linear time invariant systems*:

- Discrete time:

$$y_t = \sum_{k=-\infty}^{\infty} h(k)x_{t-k} \tag{1}$$

  Think about

$$y_t = \sum_{k=0}^{t} h(k)x_{t-k} \tag{2}$$

- $h(k)$ is called the *impulse response*, why? What happens if $x_0 = 1$ and $x_k = 0$ for $k \neq 0$?
- $S_k = \sum_{j=-\infty}^{k} h_j$ is called the *step response*, why? What happens if $x_k = 1$ for all k?

Description in the time domain (Convolution)

For *linear time invariant systems*:

- Discrete time:

$$y_t = \sum_{k=-\infty}^{\infty} h(k) x_{t-k} \quad (1)$$

  Think about

$$y_t = \sum_{k=0}^{t} h(k) x_{t-k} \quad (2)$$

- $h(k)$ is called the *impulse response*, why? What happens if $x_0 = 1$ and $x_k = 0$ for $k \neq 0$?
- $S_k = \sum_{j=-\infty}^{k} h_j$ is called the *step response*, why? What happens if $x_k = 1$ for all k?
- Sometimes people try to observe the impulse or step response directly. How could one do so?

Description in the time domain (Convolution)

For *linear time invariant systems*:

- Discrete time:

$$y_t = \sum_{k=-\infty}^{\infty} h(k) x_{t-k} \tag{1}$$

  Think about

$$y_t = \sum_{k=0}^{t} h(k) x_{t-k} \tag{2}$$

- $h(k)$ is called the *impulse response*, why? What happens if $x_0 = 1$ and $x_k = 0$ for $k \neq 0$?
- $S_k = \sum_{j=-\infty}^{k} h_j$ is called the *step response*, why? What happens if $x_k = 1$ for all k?
- Sometimes people try to observe the impulse or step response directly. How could one do so? gun shot

Description in the time domain (Convolution)

For *linear time invariant systems*:

- Discrete time:
$$y_t = \sum_{k=-\infty}^{\infty} h(k) x_{t-k} \tag{1}$$
Think about
$$y_t = \sum_{k=0}^{t} h(k) x_{t-k} \tag{2}$$
- $h(k)$ is called the *impulse response*, why? What happens if $x_0 = 1$ and $x_k = 0$ for $k \neq 0$?
- $S_k = \sum_{j=-\infty}^{k} h_j$ is called the *step response*, why? What happens if $x_k = 1$ for all k?
- Sometimes people try to observe the impulse or step response directly. How could one do so? gun shot, step increase in temperature set-point.

Dynamic response characteristics from data

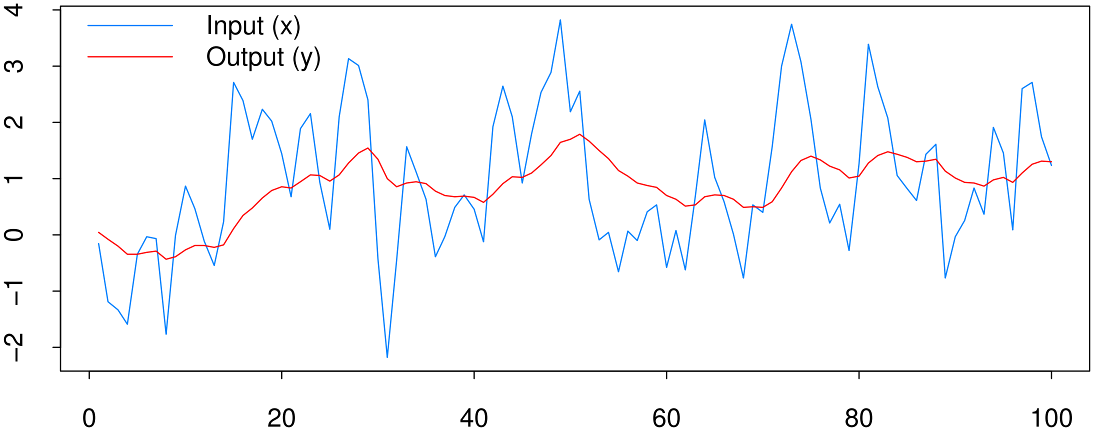
- While easy, direct observations of the impulse or step responses do not yield a lot of statistical information.

Dynamic response characteristics from data

- While easy, direct observations of the impulse or step responses do not yield a lot of statistical information.
- Instead, we use parameter estimation from data with varying inputs.

Dynamic response characteristics from data

- While easy, direct observations of the impulse or step responses do not yield a lot of statistical information.
- Instead, we use parameter estimation from data with varying inputs.



Stability based on the impulse response function

If the impulse response function is absolutely convergent, the system is stable (Theorem 4.3).

- Continuous time:

$$\int_{-\infty}^{\infty} |h(u)| \mathrm{d}u < \infty$$

- Discrete time:

$$\sum_{k=-\infty}^{\infty} |h_k| < \infty$$

The z-transform

- A way to describe dynamical systems in discrete time in the frequency domain:

$$Z(\{x_t\}) = \sum_{t=-\infty}^{\infty} x_t z^{-t}$$

The z-transform

- A way to describe dynamical systems in discrete time in the frequency domain:

$$Z(\{x_t\}) = \sum_{t=-\infty}^{\infty} x_t z^{-t} = X(z) \qquad (z \in \mathbb{C})$$

The z-transform

- A way to describe dynamical systems in discrete time in the frequency domain:

$$Z(\{x_t\}) = \sum_{t=-\infty}^{\infty} x_t z^{-t} = X(z) \qquad (z \in \mathbb{C})$$

- The z-transform of a time delay: $Z(\{x_{t-\tau}\}) = z^{-\tau} X(z)$

The z-transform

- A way to describe dynamical systems in discrete time in the frequency domain:

$$Z(\{x_t\}) = \sum_{t=-\infty}^{\infty} x_t z^{-t} = X(z) \qquad (z \in \mathbb{C})$$

- The z-transform of a time delay: $Z(\{x_{t-\tau}\}) = z^{-\tau} X(z)$
- The *transfer function* of the system is called $H(z) = \sum_{t=-\infty}^{\infty} h_t z^{-t}$

$$y_t = \sum_{k=-\infty}^{\infty} h_k x_{t-k} \quad \Leftrightarrow \quad Y(z) = H(z)X(z)$$

Linear Difference Equation

$$
\begin{aligned}
y_t + a_1 y_{t-1} + \cdots + a_p y_{t-p} &= b_0 x_{t-\tau} + b_1 x_{t-\tau-1} + \cdots + b_q x_{t-\tau-q} \\
(1 + a_1 z^{-1} + \cdots + a_p z^{-p}) Y(z) &= z^{-\tau}(b_0 + b_1 z^{-1} + \cdots + b_q z^{-q}) X(z)
\end{aligned}
$$

Linear Difference Equation

$$y_t + a_1 y_{t-1} + \cdots + a_p y_{t-p} = b_0 x_{t-\tau} + b_1 x_{t-\tau-1} + \cdots + b_q x_{t-\tau-q}$$
$$(1 + a_1 z^{-1} + \cdots + a_p z^{-p}) Y(z) = z^{-\tau}(b_0 + b_1 z^{-1} + \cdots + b_q z^{-q}) X(z)$$

Transfer function:

$$H(z) = \frac{z^{-\tau}(b_0 + b_1 z^{-1} + \cdots + b_q z^{-q})}{(1 + a_1 z^{-1} + \cdots + a_p z^{-p})}$$

Linear Difference Equation

$$y_t + a_1 y_{t-1} + \cdots + a_p y_{t-p} = b_0 x_{t-\tau} + b_1 x_{t-\tau-1} + \cdots + b_q x_{t-\tau-q}$$
$$(1 + a_1 z^{-1} + \cdots + a_p z^{-p}) Y(z) = z^{-\tau}(b_0 + b_1 z^{-1} + \cdots + b_q z^{-q}) X(z)$$

Transfer function:

$$\begin{aligned} H(z) &= \frac{z^{-\tau}(b_0 + b_1 z^{-1} + \cdots + b_q z^{-q})}{(1 + a_1 z^{-1} + \cdots + a_p z^{-p})} \\ &= \frac{z^{-\tau}(1 - n_1 z^{-1})(1 - n_2 z^{-1}) \cdots (1 - n_q z^{-1}) b_0}{(1 - \lambda_1 z^{-1})(1 - \lambda_2 z^{-1}) \cdots (1 - \lambda_p z^{-1})} \end{aligned}$$

Where the roots $n_1, n_2, \ldots, n_q$ are called the *zeros of the system* and $\lambda_1, \lambda_2, \ldots, \lambda_p$ are called the *poles of the system*.

Linear Difference Equation

$$y_t + a_1 y_{t-1} + \cdots + a_p y_{t-p} = b_0 x_{t-\tau} + b_1 x_{t-\tau-1} + \cdots + b_q x_{t-\tau-q}$$
$$(1 + a_1 z^{-1} + \cdots + a_p z^{-p}) Y(z) = z^{-\tau}(b_0 + b_1 z^{-1} + \cdots + b_q z^{-q}) X(z)$$

Transfer function:

$$\begin{aligned} H(z) &= \frac{z^{-\tau}(b_0 + b_1 z^{-1} + \cdots + b_q z^{-q})}{(1 + a_1 z^{-1} + \cdots + a_p z^{-p})} \\ &= \frac{z^{-\tau}(1 - n_1 z^{-1})(1 - n_2 z^{-1}) \cdots (1 - n_q z^{-1}) b_0}{(1 - \lambda_1 z^{-1})(1 - \lambda_2 z^{-1}) \cdots (1 - \lambda_p z^{-1})} \end{aligned}$$

Where the roots $n_1, n_2, \ldots, n_q$ are called the *zeros of the system* and $\lambda_1, \lambda_2, \ldots, \lambda_p$ are called the *poles of the system*. What does these roots say about stability and invertibility of the system?

Linear Difference Equation

$$y_t + a_1 y_{t-1} + \cdots + a_p y_{t-p} = b_0 x_{t-\tau} + b_1 x_{t-\tau-1} + \cdots + b_q x_{t-\tau-q}$$
$$(1 + a_1 z^{-1} + \cdots + a_p z^{-p}) Y(z) = z^{-\tau}(b_0 + b_1 z^{-1} + \cdots + b_q z^{-q}) X(z)$$

Transfer function:

$$\begin{aligned} H(z) &= \frac{z^{-\tau}(b_0 + b_1 z^{-1} + \cdots + b_q z^{-q})}{(1 + a_1 z^{-1} + \cdots + a_p z^{-p})} \\ &= \frac{z^{-\tau}(1 - n_1 z^{-1})(1 - n_2 z^{-1}) \cdots (1 - n_q z^{-1}) b_0}{(1 - \lambda_1 z^{-1})(1 - \lambda_2 z^{-1}) \cdots (1 - \lambda_p z^{-1})} \end{aligned}$$

Where the roots $n_1, n_2, \ldots, n_q$ are called the *zeros of the system* and $\lambda_1, \lambda_2, \ldots, \lambda_p$ are called the *poles of the system*. What does these roots say about stability and invertibility of the system?
The system is stable if all poles lie within the unit circle

Linear Difference Equation

$$y_t + a_1 y_{t-1} + \cdots + a_p y_{t-p} = b_0 x_{t-\tau} + b_1 x_{t-\tau-1} + \cdots + b_q x_{t-\tau-q}$$
$$(1 + a_1 z^{-1} + \cdots + a_p z^{-p}) Y(z) = z^{-\tau}(b_0 + b_1 z^{-1} + \cdots + b_q z^{-q}) X(z)$$

Transfer function:

$$\begin{aligned} H(z) &= \frac{z^{-\tau}(b_0 + b_1 z^{-1} + \cdots + b_q z^{-q})}{(1 + a_1 z^{-1} + \cdots + a_p z^{-p})} \\ &= \frac{z^{-\tau}(1 - n_1 z^{-1})(1 - n_2 z^{-1}) \cdots (1 - n_q z^{-1}) b_0}{(1 - \lambda_1 z^{-1})(1 - \lambda_2 z^{-1}) \cdots (1 - \lambda_p z^{-1})} \end{aligned}$$

Where the roots $n_1, n_2, \ldots, n_q$ are called the *zeros of the system* and $\lambda_1, \lambda_2, \ldots, \lambda_p$ are called the *poles of the system*. What does these roots say about stability and invertibility of the system?
The system is stable if all poles lie within the unit circle
The system is invertible if all zeroes lie within the unit circle

Estimating the impulse response

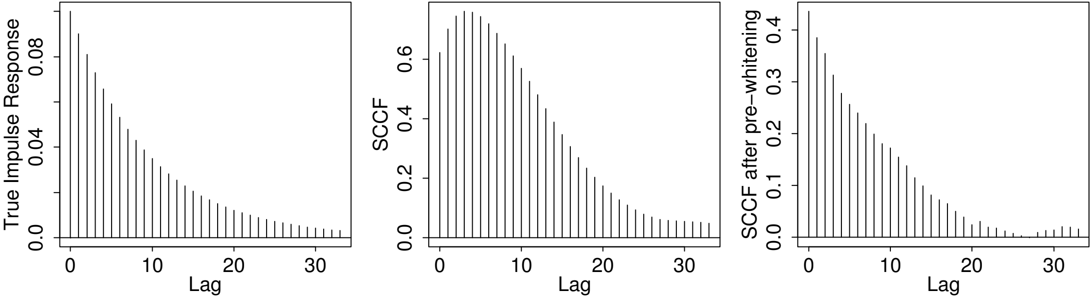
- The shape of the impulse response function is dictated by what kind of relationship there is between the input, X and the output, Y.

Estimating the impulse response

- The shape of the impulse response function is dictated by what kind of relationship there is between the input, X and the output, Y.
- The CCF (cross-correlation function) can be used to reveal this relationship, but requires pre-whitening

Estimating the impulse response

- The shape of the impulse response function is dictated by what kind of relationship there is between the input, X and the output, Y.
- The CCF (cross-correlation function) can be used to reveal this relationship, but requires pre-whitening:



Cross covariance and cross correlation functions

Estimate of the cross covariance function:

$$
\begin{aligned}
C_{XY}(k) &= \frac{1}{N}\sum_{t=1}^{N-k}(X_t - \overline{X})(Y_{t+k} - \overline{Y}) \\
C_{XY}(-k) &= \frac{1}{N}\sum_{t=1}^{N-k}(X_{t+k} - \overline{X})(Y_t - \overline{Y})
\end{aligned}
$$

Cross covariance and cross correlation functions

Estimate of the cross covariance function:

$$
\begin{aligned}
C_{XY}(k) &= \frac{1}{N}\sum_{t=1}^{N-k}(X_t - \overline{X})(Y_{t+k} - \overline{Y}) \\
C_{XY}(-k) &= \frac{1}{N}\sum_{t=1}^{N-k}(X_{t+k} - \overline{X})(Y_t - \overline{Y})
\end{aligned}
$$

Estimate of the cross correlation function:

$$
\widehat{\rho}_{XY}(k) = C_{XY}(k)/\sqrt{C_{XX}(0)\,C_{YY}(0)}
$$

Cross covariance and cross correlation functions

Estimate of the cross covariance function:

$$
\begin{aligned}
C_{XY}(k) &= \frac{1}{N}\sum_{t=1}^{N-k}(X_t - \overline{X})(Y_{t+k} - \overline{Y}) \\
C_{XY}(-k) &= \frac{1}{N}\sum_{t=1}^{N-k}(X_{t+k} - \overline{X})(Y_t - \overline{Y})
\end{aligned}
$$

Estimate of the cross correlation function:

$$
\widehat{\rho}_{XY}(k) = C_{XY}(k)/\sqrt{C_{XX}(0)C_{YY}(0)}
$$

What is a defining property of the CCF for causal systems with no feedback?

Cross covariance and cross correlation functions

Estimate of the cross covariance function:

$$
\begin{aligned}
C_{XY}(k) &= \frac{1}{N}\sum_{t=1}^{N-k}(X_t - \overline{X})(Y_{t+k} - \overline{Y}) \\
C_{XY}(-k) &= \frac{1}{N}\sum_{t=1}^{N-k}(X_{t+k} - \overline{X})(Y_t - \overline{Y})
\end{aligned}
$$

Estimate of the cross correlation function:

$$
\widehat{\rho}_{XY}(k) = C_{XY}(k)/\sqrt{C_{XX}(0)C_{YY}(0)}
$$

What is a defining property of the CCF for causal systems with no feedback? If at least one of the processes is white noise and if the processes are uncorrelated then $\widehat{\rho}_{XY}(k)$ is approximately normally distributed with mean 0 and variance $1/N$

Cross Correlations for systems without measurement noise

$$Y_t = \sum_{i=-\infty}^{\infty} h_i X_{t-i}$$

Given γ_{XX} and the system description we obtain

$$\gamma_{YY}(k) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h_i h_j \gamma_{XX}(k-j+i)$$

$$\gamma_{XY}(k) = \sum_{i=-\infty}^{\infty} h_i \gamma_{XX}(k-i).$$

Cross Correlations for systems without measurement noise

$$Y_t = \sum_{i=-\infty}^{\infty} h_i X_{t-i}$$

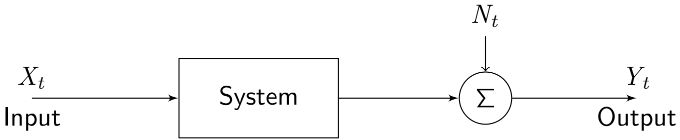
Given γ_{XX} and the system description we obtain

$$\gamma_{YY}(k) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h_i h_j \gamma_{XX}(k - j + i)$$

$$\gamma_{XY}(k) = \sum_{i=-\infty}^{\infty} h_i \gamma_{XX}(k - i).$$

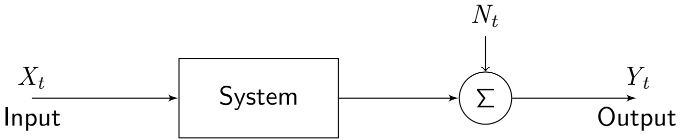
What happens when $\{X_t\}$ is white noise?

Systems with measurement noise



$$Y_t = \sum_{i=-\infty}^{\infty} h_i X_{t-i} + N_t.$$

Systems with measurement noise



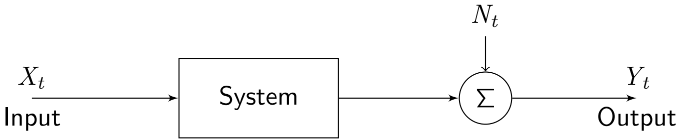
$$Y_t = \sum_{i=-\infty}^{\infty} h_i X_{t-i} + N_t.$$

Given γ_{XX} and γ_{NN} we obtain*:

$$\gamma_{YY}(k) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h_i h_j \gamma_{XX}(k-j+i) + \gamma_{NN}(k)$$

$$\gamma_{XY}(k) = \sum_{i=-\infty}^{\infty} h_i \gamma_{XX}(k-i).$$

Systems with measurement noise



$$Y_t = \sum_{i=-\infty}^{\infty} h_i X_{t-i} + N_t.$$

Given γ_{XX} and γ_{NN} we obtain*:

$$\gamma_{YY}(k) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} h_i h_j \gamma_{XX}(k-j+i) + \gamma_{NN}(k)$$

$$\gamma_{XY}(k) = \sum_{i=-\infty}^{\infty} h_i \gamma_{XX}(k-i).$$

Again, notice what happens for white noise. *Assumption: No feedback in the system.

Estimating the impulse response

- When our input is not white, we try and make it so by using pre-whitening.

Estimating the impulse response

- When our input is not white, we try and make it so by using pre-whitening.
- The reason is that

$$\gamma_{XY}(k) = \sum_{i=-\infty}^{\infty} h_i \gamma_{XX}(k-i)$$

so if and only if $\{X_t\}$ is white noise: $\boxed{\gamma_{XY}(k) = h_k \sigma_X^2}$

Pre-whitening

a) A suitable ARMA-model is applied to the input series:

$$\eta(B)X_t = \nu(B)\alpha_t.$$

Pre-whitening

a) A suitable ARMA-model is applied to the input series:

$$\eta(B)X_t = \nu(B)\alpha_t.$$

b) We perform a *prewhitening* of the input series

$$\alpha_t = \nu(B)^{-1}\eta(B)X_t$$

Pre-whitening

a) A suitable ARMA-model is applied to the input series:

$$\eta(B)X_t = \nu(B)\alpha_t.$$

b) We perform a *prewhitening* of the input series

$$\alpha_t = \nu(B)^{-1}\eta(B)X_t$$

c) The output–series $\{Y_t\}$ is filtered with the same model, i.e.

$$\beta_t = \nu(B)^{-1}\eta(B)Y_t.$$

Pre-whitening

a) A suitable ARMA-model is applied to the input series:

$$\eta(B)X_t = \nu(B)\alpha_t.$$

b) We perform a *prewhitening* of the input series

$$\alpha_t = \nu(B)^{-1}\eta(B)X_t$$

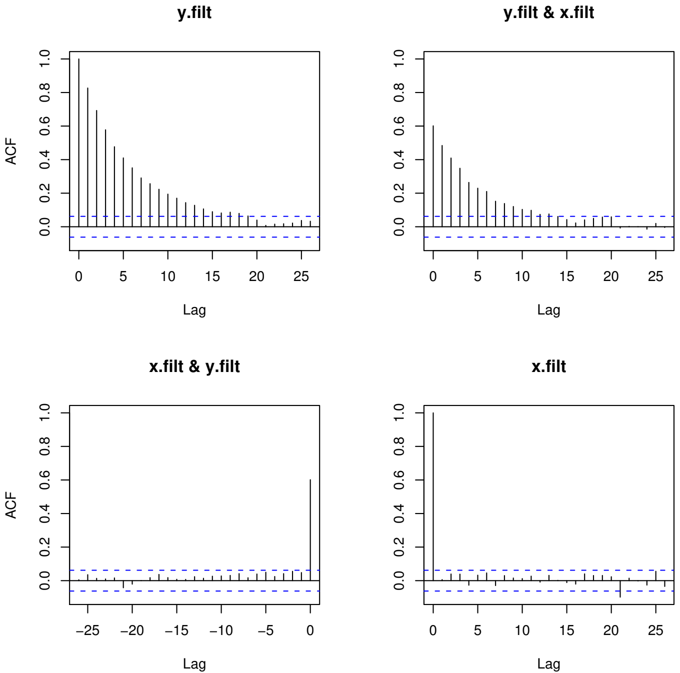
c) The output–series $\{Y_t\}$ is filtered with the same model, i.e.

$$\beta_t = \nu(B)^{-1}\eta(B)Y_t.$$

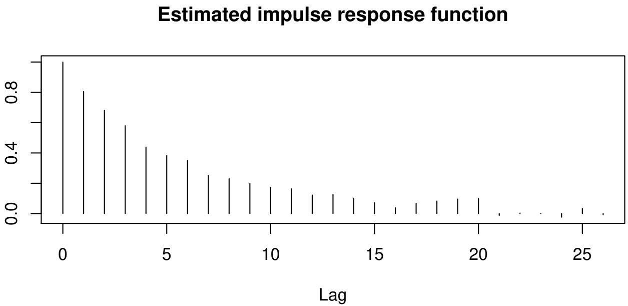
d) Now the *impulse response function is estimated* by

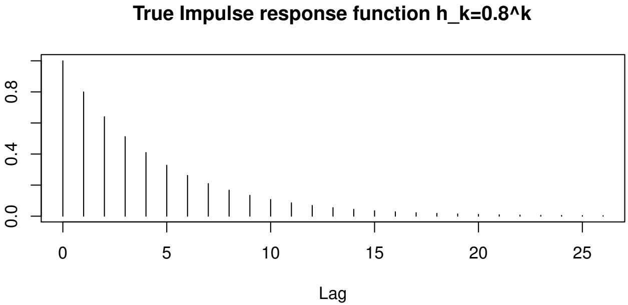
$$\widehat{h}_k = C_{\alpha\beta}(k)/C_{\alpha\alpha}(0) = C_{\alpha\beta}(k)/S_\alpha^2.$$

Graphical output

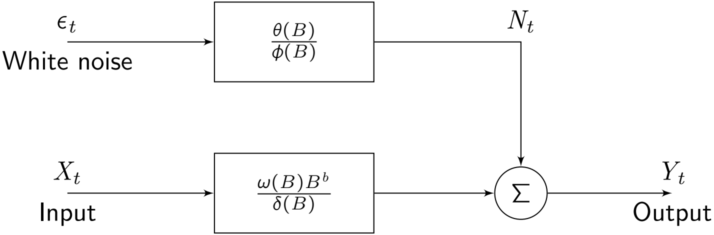


Impulse response functions





Transfer function models



$$Y_t = \frac{\omega(B)}{\delta(B)} B^b X_t + \frac{\theta(B)}{\varphi(B)} \varepsilon_t$$

- Also called Box-Jenkins models

Some names

The following are all sub-models of transfer function models:

Some names

The following are all sub-models of transfer function models:

- FIR: Finite Impulse Response (impulse response function(s) of finite length): $y_t = \sum_{k=-\infty}^{\infty} h(k) x_{t-k}$.

Some names

The following are all sub-models of transfer function models:

- FIR: Finite Impulse Response (impulse response function(s) of finite length): $y_t = \sum_{k=-\infty}^{\infty} h(k) x_{t-k}$.
- ARX: Auto Regressive with eXogenous input: $\varphi(B) Y_t = \omega(B) u_t + \epsilon_t$.

Some names

The following are all sub-models of transfer function models:

- FIR: Finite Impulse Response (impulse response function(s) of finite length): $y_t = \sum_{k=-\infty}^{\infty} h(k) x_{t-k}$.
- ARX: Auto Regressive with eXogenous input: $\varphi(B) Y_t = \omega(B) u_t + \epsilon_t$.
- ARMAX: Auto Regressive Moving Average, eXogenous input: $\varphi(B) Y_t = \omega(B) X_t + \theta(B) \varepsilon_t$.

Some names

The following are all sub-models of transfer function models:

- FIR: Finite Impulse Response (impulse response function(s) of finite length): $y_t = \sum_{k=-\infty}^{\infty} h(k) x_{t-k}$.
- ARX: Auto Regressive with eXogenous input: $\varphi(B) Y_t = \omega(B) u_t + \epsilon_t$.
- ARMAX: Auto Regressive Moving Average, eXogenous input: $\varphi(B) Y_t = \omega(B) X_t + \theta(B) \varepsilon_t$.
- OE: Output Error model: $Y_t = \frac{\omega(B)}{\delta(B)} B^b X_t + \varepsilon_t$.

Some names

The following are all sub-models of transfer function models:

- FIR: Finite Impulse Response (impulse response function(s) of finite length): $y_t = \sum_{k=-\infty}^{\infty} h(k) x_{t-k}$.
- ARX: Auto Regressive with eXogenous input: $\varphi(B) Y_t = \omega(B) u_t + \epsilon_t$.
- ARMAX: Auto Regressive Moving Average, eXogenous input: $\varphi(B) Y_t = \omega(B) X_t + \theta(B) \varepsilon_t$.
- OE: Output Error model: $Y_t = \frac{\omega(B)}{\delta(B)} B^b X_t + \varepsilon_t$.
- Regression models with ARMA noise (the `xreg` option to `arima` in R):

$$Y_t = X_t + \frac{\theta(B)}{\varphi(B)} \varepsilon_t \text{ or } \varphi(B) Y_t = \varphi(B) X_t + \theta(B) \varepsilon_t$$

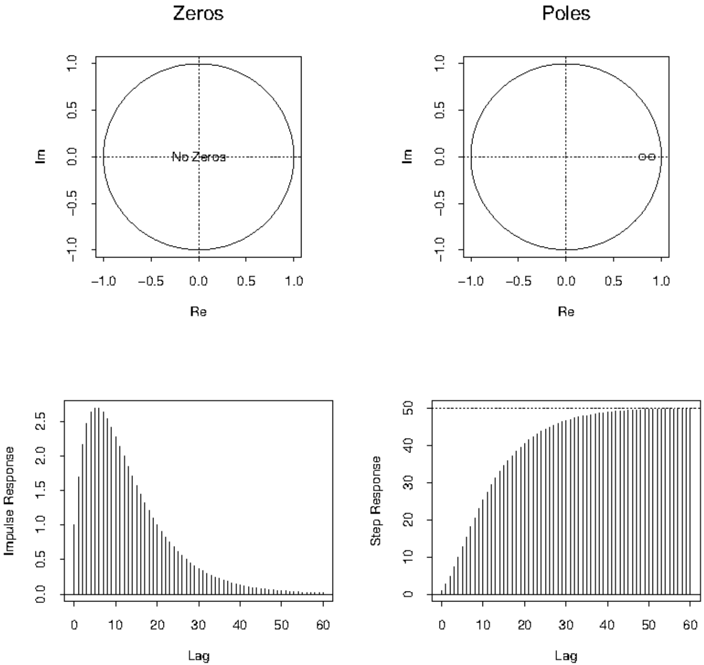
Identification of transfer function models

$$h(B) = \frac{\omega(B)B^b}{\delta(B)} = h_0 + h_1 B + h_2 B^2 + h_3 B^3 + h_4 B^4 + \ldots$$

- Using pre-whitening we estimate the impulse response and “guess” an appropriate structure of $h(B)$ based on this.

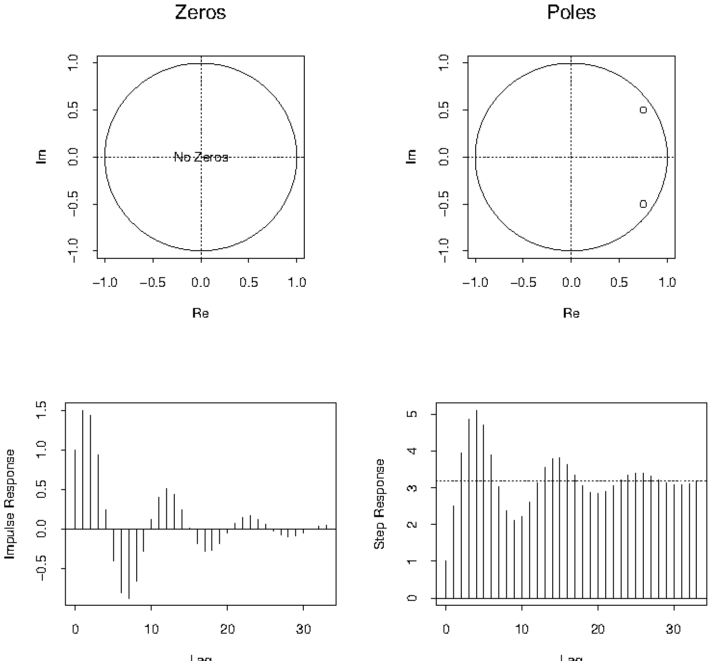
2 real poles

$$h(B) = \frac{1}{1 - 1.7B + 0.72B^2}$$



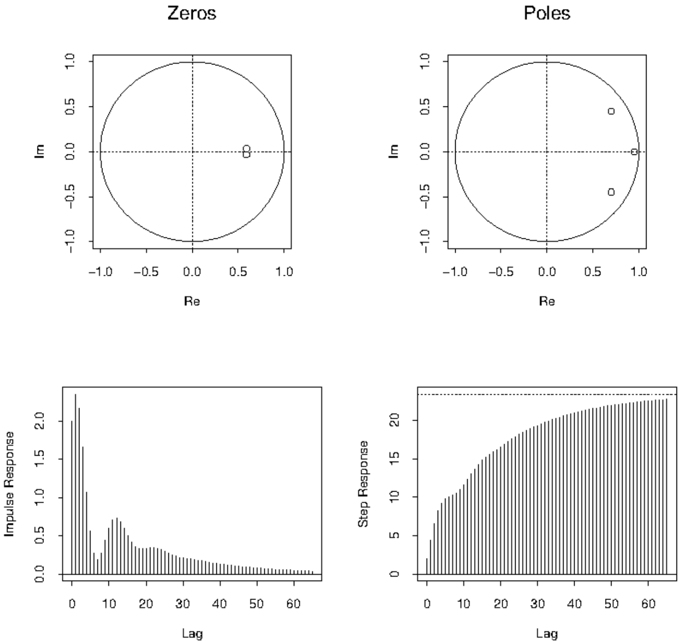
2 complex

$$h(B) = \frac{1}{1 - 1.5B + 0.81B^2}$$



1 real, 2 comp

$$h(B) = \frac{2 - 2.35B + 0.69B^2}{1 - 2.35B + 2.02B^2 - 0.66B^3}$$



Identification of the transfer function for the noise

- After selection of the structure of the transfer function of the input we estimate the parameters of the model (assuming N_t to be white)

$$Y_t = \frac{\omega(B)}{\delta(B)} B^b X_t + N_t$$

Identification of the transfer function for the noise

- After selection of the structure of the transfer function of the input we estimate the parameters of the model (assuming N_t to be white)

$$Y_t = \frac{\omega(B)}{\delta(B)} B^b X_t + N_t$$

- Then, we extract the residuals $\{N_t\}$ and identify a structure for an ARMA model of this series

$$N_t = \frac{\theta(B)}{\varphi(B)} \varepsilon_t \quad \Leftrightarrow \quad \varphi(B) N_t = \theta(B) \varepsilon_t$$

Identification of the transfer function for the noise

- After selection of the structure of the transfer function of the input we estimate the parameters of the model (assuming N_t to be white)

$$Y_t = \frac{\omega(B)}{\delta(B)} B^b X_t + N_t$$

- Then, we extract the residuals $\{N_t\}$ and identify a structure for an ARMA model of this series

$$N_t = \frac{\theta(B)}{\varphi(B)} \varepsilon_t \quad \Leftrightarrow \quad \varphi(B) N_t = \theta(B) \varepsilon_t$$

- Finally, we have the full structure of the model and we estimate all parameters simultaneously

Estimation

- Form 1-step predictions, treating the input $\{X_t\}$ as known (corresponds to conditioning on observed $\{X_t\}$ if it is actually stochastic)

Estimation

- Form 1-step predictions, treating the input $\{X_t\}$ as known (corresponds to conditioning on observed $\{X_t\}$ if it is actually stochastic)
- Select the parameters so that the sum of squares of these errors is as small as possible (implicit assumption of $\{\epsilon_t\}$ being gaussian).

Estimation

- Form 1-step predictions, treating the input $\{X_t\}$ as known (corresponds to conditioning on observed $\{X_t\}$ if it is actually stochastic)
- Select the parameters so that the sum of squares of these errors is as small as possible (implicit assumption of $\{\epsilon_t\}$ being gaussian).
- For FIR and ARX models we can write the model as $\boldsymbol{Y}_t = \boldsymbol{X}_t^T \boldsymbol{\theta} + \varepsilon_t$ and use LS-estimates

Model validation

As for ARMA models with the additions:

- Test for cross correlation between the residuals and the input. If $\{\varepsilon_t\}$ is white noise and when there is no correlation between the input and the residuals then (approximately)

$$\hat{\rho}_{\varepsilon X}(k) \sim \mathcal{N}(0, 1/N)$$

Model validation

As for ARMA models with the additions:

- Test for cross correlation between the residuals and the input. If $\{\varepsilon_t\}$ is white noise and when there is no correlation between the input and the residuals then (approximately)

$$\hat{\rho}_{\varepsilon X}(k) \sim \mathcal{N}(0, 1/N)$$

- A *Portmanteau test* (*Ljung-Box*) can also be performed to test for significent ccf's.