HSLU Hochschule Luzern

Time Series Analysis

ARIMA

Dr. Ludovic Amruthalingam
ludovic.amruthalingam@hslu.ch

Informatik

FH Zentralschweiz

Outline

- Autoregressive model
- Partial autocorrelation function
- Moving average model
- Invertibility condition
- ARMA model
- ARIMA model
- SARIMA model

Autoregressive (AR) model

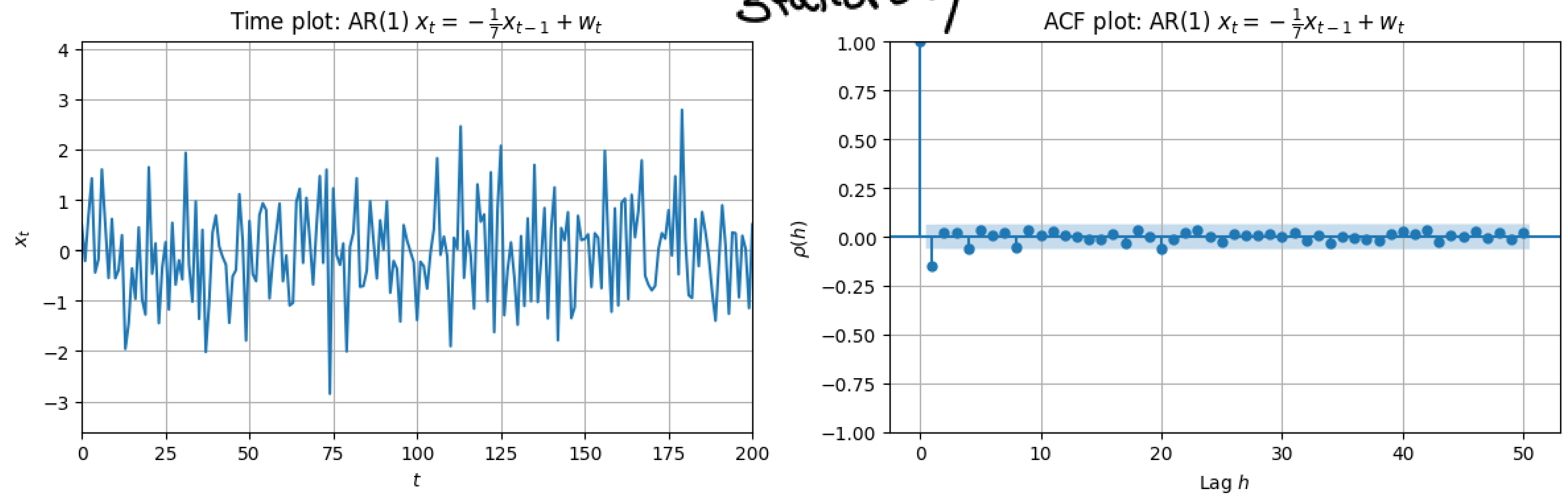
independent for all before timestamp t

innovation term

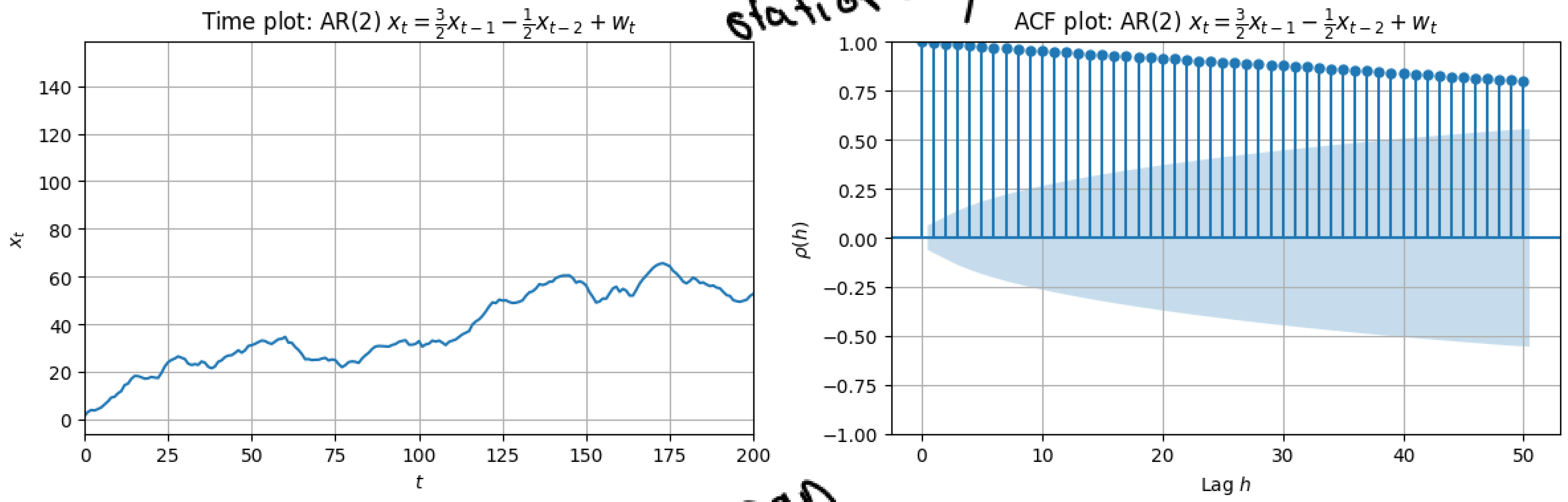
Autoregressive model of order p, AR(p): $X_t = c + \sum_{i=1}^{p} \phi_i X_{t-i} + W_t$ with $W_t \sim WN(0, \sigma^2)$ and c a constant.

We focus on the **zero-mean** AR model ($c = 0$) since we can always choose to model $\mathrm{Y_t} = X_t - \mu_X$.

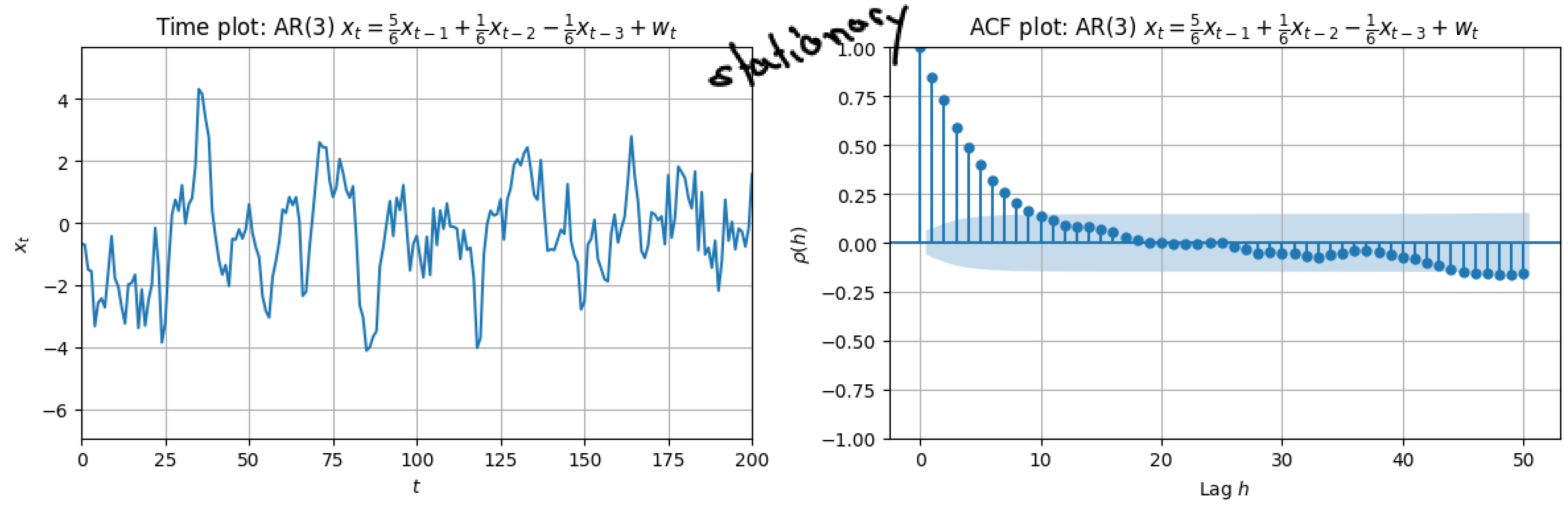
stationary



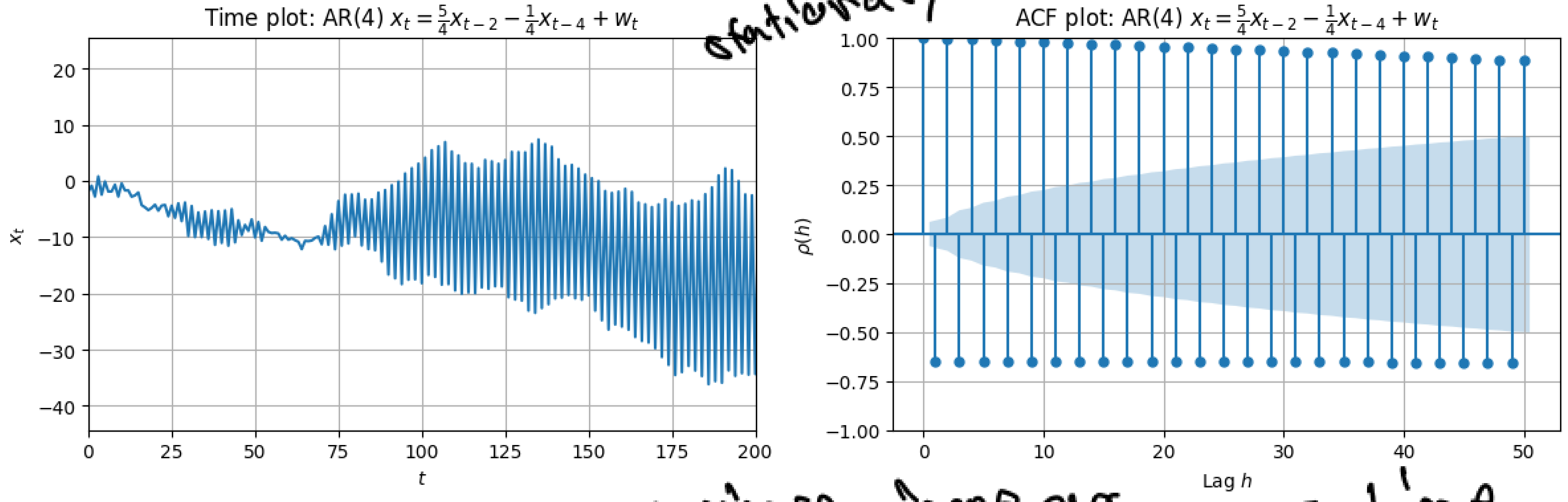
non stationary



stationary



non stationary



variance increases over time

Stationarity of AR(1)

A zero-mean AR(1) can be expanded as $X_t = \phi X_{t-1} + W_t = W_t + \phi(\phi X_{t-2} + W_{t-1}) = \phi^k X_{t-k} + \sum_{i=0}^{k} \phi^i W_{t-i}$

When $|\phi| < 1$ and $k \to \infty$ then $X_t = \sum_{i=0}^{\infty} \phi^i W_{t-i} + 0$ *(under these conditions $\lim_{k\to\infty} \phi^k = 0$ and $\sum_{i=0}^{\infty} \phi^i = \frac{1}{1-\phi}$)*

because white noise with mean 0

Mean: $E[X_t] = E\left[\sum_{i=0}^{\infty} \phi^i W_{t-i}\right] = \sum_{i=0}^{\infty} \phi^i E[W_{t-i}] = 0$

Auto-covariance: $cov(X_{t+h}, X_t) = cov\left(\sum_{j=0}^{\infty} \phi^j W_{t+h-j}, \sum_{i=0}^{\infty} \phi^i W_{t-i}\right) = \sum_{i,j=0}^{\infty} \phi^j \phi^i cov(W_{t+h-j}, W_{t-i})$

$$= \sum_{i=0}^{\infty} \phi^{i+h} \phi^i cov(W_{t+h-(i+h)}, W_{t-i}) = \sum_{i=0}^{\infty} \phi^{2i} \phi^h \sigma^2 = \phi^h \sigma^2 \sum_{i=0}^{\infty} \phi^{2i} = \sigma^2 \frac{\phi^h}{1-\phi^2} = \gamma_h$$

Thus, an AR(1) with $|\phi| < 1$ is <u>asymptotically</u> **stationary**.

Stationarity of AR(p)

The zero-mean AR(p) equation can be rewritten as $W_t = X_t - \sum_{i=1}^{p} \phi_i X_{t-i} = \left(1 - \sum_{i=1}^{p} \phi_i B^i\right) X_t = \Phi(B) X_t$
where $\Phi(B) = \left(1 - \sum_{i=1}^{p} \phi_i B^i\right)$ is called the **characteristic polynomial**.

Theorem: An AR(p) process is stationary $\Leftrightarrow$ all roots of $\boldsymbol{\Phi(B)}$ are strictly greater than 1 in absolute value.

Recall that the roots of a polynomial $P(x)$ are the values r that satisfy $P(r) = 0$ e.g., $r_{1,2} = \pm 2$ for $P(x) = x^2 - 4$.

Examples (cf. figure with AR examples):

- AR(1) with $\Phi(B) = 1 + \frac{1}{7}B = \frac{1}{7}(7 + B)$, the root is $r_1 = -7 \Rightarrow$ stationary
- AR(2) with $\Phi(B) = 1 - \frac{3}{2}B + \frac{1}{2}B^2 = \frac{1}{2}(1 - B)(2 - B)$, the roots are $r_1 = 1$ and $r_2 = 2 \Rightarrow$ not stationary
- AR(3) with $\Phi(B) = 1 - \frac{5}{6}B - \frac{1}{6}B^2 + \frac{1}{6}B^3 = \frac{1}{6}(2 - B)\left(3 - B - B^2\right)$, then $r_1 = 2$ and $r_{2,3} = \frac{-1 \pm \sqrt{13}}{2} \Rightarrow$ stationary
- AR(4) with $\Phi(B) = 1 - \frac{5}{4}B^2 + \frac{1}{4}B^4 = \frac{1}{4}(1 - B)(1 + B)(2 - B)(2 + B)$, then $r_{1,2} = \pm 1$, $r_{3,4} = \pm 2 \Rightarrow$ not stationary

Auto-correlation function of autoregressive models

From here on, we focus on **stationary AR(p)** models $X_t = \sum_{i=1}^{p} \phi_i X_{t-i} + W_t$ with **zero-mean**.

The ACF can be computed as follows: $\rho_h = \frac{\gamma_h}{\gamma_0} = \frac{1}{\gamma_0} \sum_{i=1}^{p} \phi_i cov(X_{t+h-i}, X_t) = \frac{1}{\gamma_0} \sum_{i=1}^{p} \phi_i \gamma_{h-i} = \sum_{i=1}^{p} \phi_i \rho_{h-i}$

Since $\rho_0 = 1$ and $\rho_h = \rho_{-h}$ (γ is symmetric), we get a linear equation system called the **Yule-Walker equations**:

$$\underbrace{\begin{pmatrix} \rho_1 \\ \rho_2 \\ \rho_3 \\ \vdots \\ \rho_p \end{pmatrix}}_{P} = \underbrace{\begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{p-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{p-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{p-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \rho_{p-3} & \cdots & 1 \end{pmatrix}}_{R} \underbrace{\begin{pmatrix} \phi_1 \\ \phi_2 \\ \phi_3 \\ \vdots \\ \phi_p \end{pmatrix}}_{\Phi} \quad \text{or } P = R\Phi \text{ with } R \text{ invertible}$$

This system can be solved to obtain either the **theoretical** ρ_i **or the estimated** ϕ_i based on the situation.

It can be shown that the ACF **decays** exponentially as the lag increases but remains **non-zero** for all lags.

because of left over correlation of ρ

Partial auto-correlation function (PACF)

only works with stationary time series

The PACF π_h measures the correlation between X_t and X_{t+h} after **removing** the influence of the **intermediate lags**:

$$\pi_h = cor(X_{t+h}, X_t | X_{t+1}, \dots X_{t+h-1})$$

This influence can be represented as a linear **regression** of the intermediate lags.
We denote $\hat{X}_{t+h}$ (respectively $\hat{X}_t$) as the linear regression of X_{t+h} (respectively X_t) on $X_{t+1}, \dots X_{t+h-1}$.

$$\pi_h = \begin{cases} 1 & h = 0 \\ \rho_1, & h = 1 \\ cor\left(X_{t+h} - \hat{X}_{t+h}, X_t - \hat{X}_t\right) & h > 1 \end{cases}$$

For $h = p$, we have $\hat{X}_{t+p} \approx X_{t+p} - \phi_p X_t - W_{t+p} \Longrightarrow \pi_p \approx \phi_p$

For $h > p$, the regression results in $\hat{X}_{t+h} \approx X_{t+h} - W_{t+h} \Rightarrow$ the **PACF is zero for all lags greater than the AR order p**.

To compute π_i, consider the Yule-Walker equation system $\mathrm{P}^{(i)} = R^{(i)}\Phi^{(\mathrm{i})}$ where (i) denotes setting $p = i$ then

π_i is the i^{th} component of the vector $\Phi^{(\mathrm{i})} = \left(R^{(i)}\right)^{-1}\mathrm{P}^{(i)}$
(see Levinson-Durbin method for efficient computation)

Partial auto-correlation function (PACF)

order 1 coefficient $-\frac{1}{7}$

last significant coefficient

non stationary ?

order 3 coefficient $-\frac{1}{6}$

normally consider $\frac{n}{4}$ lags

Fitting an AR model

Given a **stationary** time series realization $\{x_1, x_2, \dots x_n\}$ with **zero-mean**,

1. Is the autoregressive model **appropriate**?
   - ACF shows an exponential (sinusoidal) decay pattern.
   - PACF shows a cut-off at some lag p after which the PACF is close to zero.
2. Select model order based on the PACF, favoring smaller p values.
3. Estimate the AR(p) model **coefficients**.
   - Burg’s algorithm
   - Maximum-likelihood estimation
4. Analyze **residuals** to evaluate model goodness of fit (discussed in following lectures).

Estimating the coefficients of an AR(p) process

Different approaches exist. They are asymptotically equivalent but can produce variations in practice.
Considering a **zero-mean stationary AR(p)** realization $\{x_1, x_2, \dots x_n\}$,

Ordinary least squares (OLS): $\widehat{\boldsymbol{\phi}} = \begin{pmatrix} \widehat{\phi}_1 \\ \widehat{\phi}_2 \\ \vdots \\ \widehat{\phi}_p \end{pmatrix} = \left(X^T X\right)^{-1} X^T Y$ with $Y = \begin{pmatrix} x_{p+1} \\ x_{p+2} \\ \vdots \\ x_n \end{pmatrix}$ and $X = \begin{pmatrix} x_p & x_{p-1} & \cdots & x_1 \\ x_{p+1} & x_p & \cdots & x_2 \\ \vdots & \vdots & \ddots & \vdots \\ x_{n-1} & x_{n-2} & \cdots & x_{n-p} \end{pmatrix}$

- The first p observations can only be used as predictors, leaving $\mathrm{n} - p$ observations to evaluate fit $(x_t - \hat{x}_t)^2$.

Burg's algorithm: minimize forward and backward squared prediction errors $(x_t - \hat{x}_t)^2$.

- Solves OLS asymmetry by considering that any AR(p) process is also an AR(p) if t is reversed.

Yule-Walker Equations: for finite samples, $\widehat{\boldsymbol{\phi}}$ typically have lower likelihood → prefer **Burg's algorithm**.

Maximum-likelihood estimation (MLE): find $\widehat{\boldsymbol{\phi}}$ that maximizes the likelihood of the data under the AR(p) model.

- residual should be close to zero

$\widehat{\boldsymbol{\phi}} = \max_{\boldsymbol{\phi}} \prod_{i=p+1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\widehat{w}_i^{\,2}}{2\sigma^2}\right)$ with $\widehat{w}_i = x_i - \sum_{j=1}^{p} \phi_i x_{i-j}$ and assuming $w_i \sim \mathcal{N}(0, \sigma^2)$

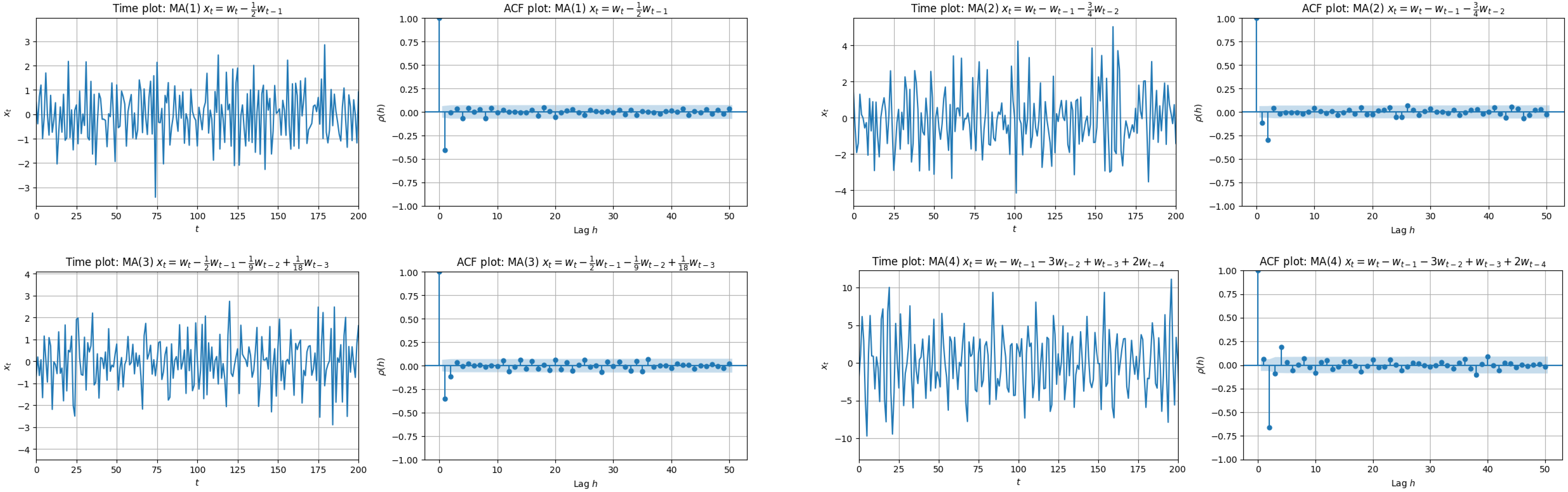
gaussian distribution formula

SCW02

Moving average (MA) model (≠ moving average filter)

Moving average model of order q, MA(q): $X_t = c + \sum_{i=1}^{q} \theta_i W_{t-i} + W_t$ with $W_t \sim WN(0, \sigma^2)$ and c a constant.

We focus on the **zero-mean** MA model $(c = 0)$ since we can always choose to model $\mathrm{Y_t} = X_t - \mu_X$.



Stationarity of MA(q)

The zero-mean MA(q) can be written as $X_t = \sum_{i=1}^{q} \theta_i W_{t-i} + W_t = \sum_{\boldsymbol{i=0}}^{q} \theta_i W_{t-i}$ with $\theta_0 = 1$.

Mean: $E[X_t] = E\left[\sum_{i=0}^{q} \theta_i W_{t-i}\right] = \sum_{i=0}^{q} \theta_i E[W_{t-i}] = 0$

Auto-covariance: $cov(X_{t+h}, X_t) = cov\left(\sum_{j=0}^{q} \theta_j W_{t+h-j}, \sum_{i=0}^{q} \theta_i W_{t-i}\right) = \sum_{i,j=0}^{q} \theta_j \theta_i cov\left(W_{t+h-j}, W_{t-i}\right)$

$$= \sum_{\substack{0 \le i \\ i+h \le q}}^{q} \theta_{i+h} \theta_i cov\left(W_{t+h-(i+h)}, W_{t-i}\right) = \begin{cases} \sigma^2 \sum_{i=0}^{q-h} \theta_{i+h} \theta_i & 0 \le h \le q \\ 0 & h > q \end{cases} = \gamma_h$$

MA(q) is stationary for any coefficients $\boldsymbol{\theta_i}$.

The **ACF is zero for all lags greater than the MA order q and $\boldsymbol{\rho_q = \theta_q}$.**

Duality between the MA(q) and AR(p) models

MA(q) processes can be represented as **AR(∞) processes**. Taking a MA(1) as example,

$$X_t = \theta W_{t-1} + W_t = \theta(X_{t-1} - \theta W_{t-2}) + W_t = \sum_{i=1}^{\infty} (-1)^{i+1} \theta^i X_{t-i} + W_t = \sum_{i=1}^{\infty} \phi_i X_{t-i} + W_t$$

with $\phi_i = (-1)^{i+1}\theta^i$

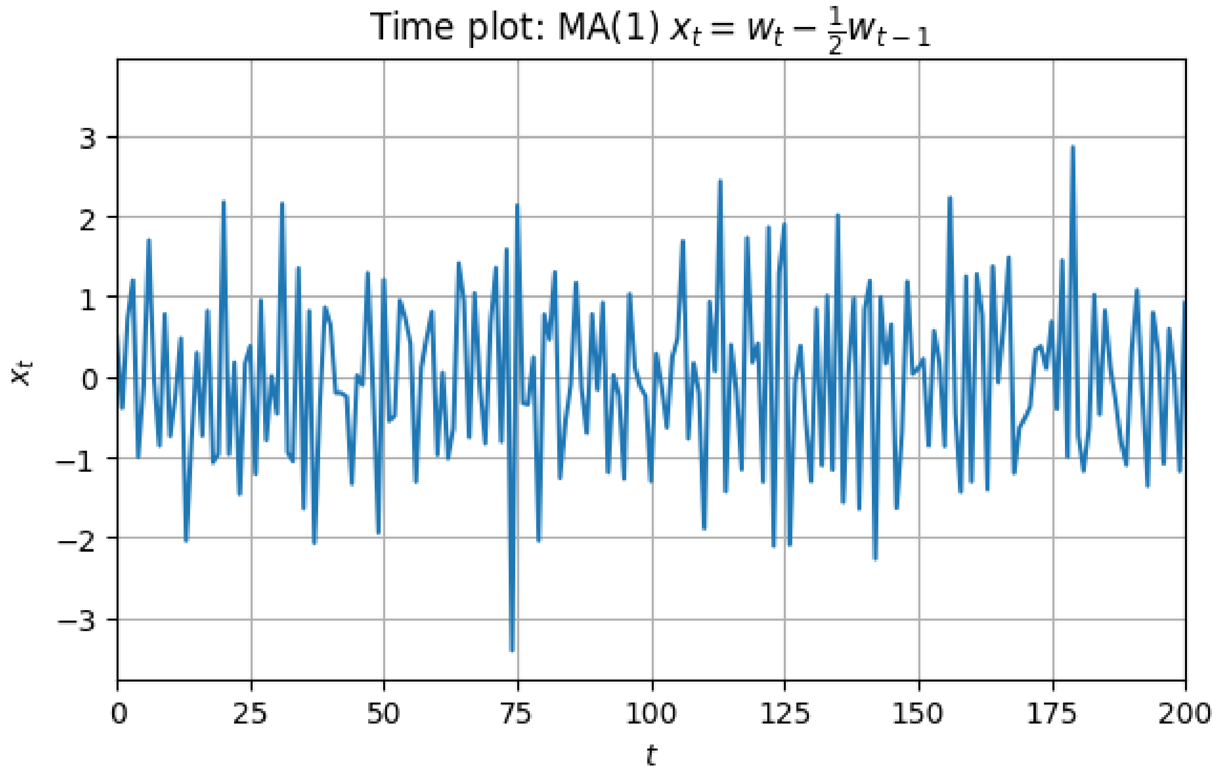
AR(p) processes can be represented as an **MA(∞) processes**. Taking an AR(1) as example,

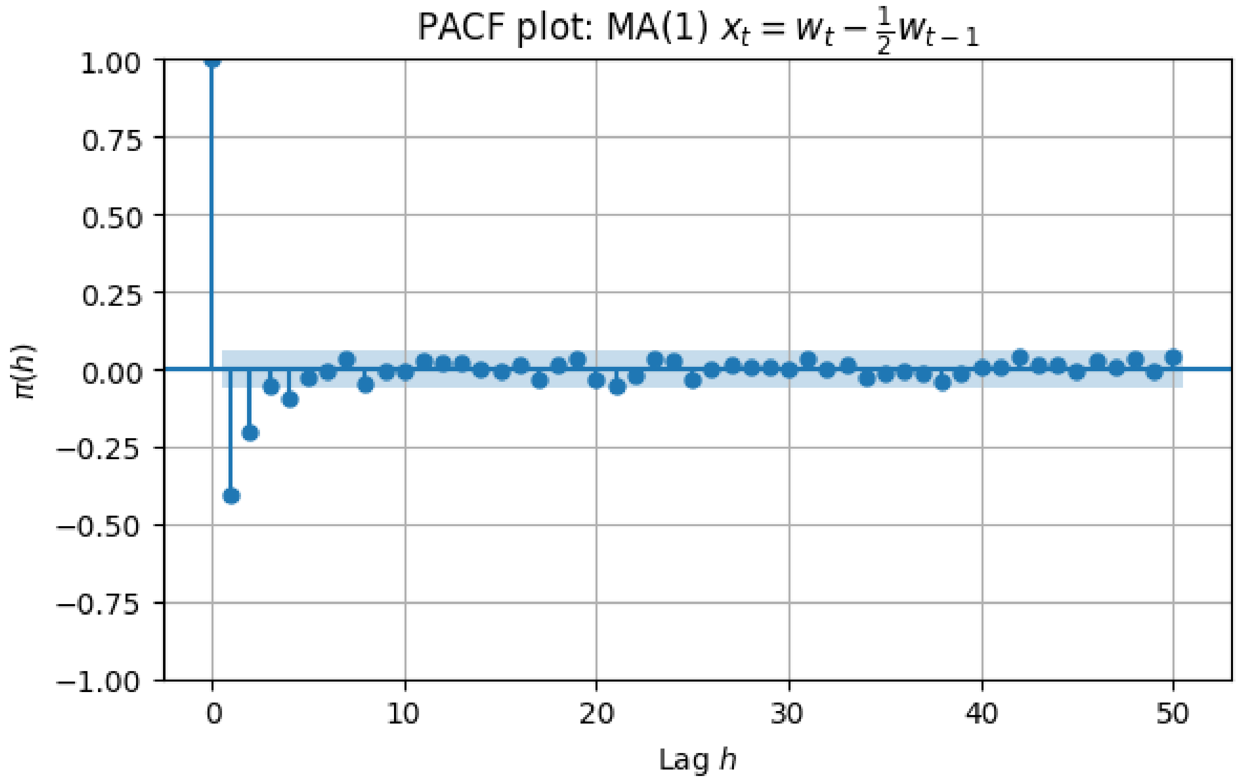
$$X_t = \phi X_{t-1} + W_t = \phi(\phi X_{t-2} + W_{t-1}) + \mathrm{W_t} = \sum_{i=0}^{\infty} \phi^i W_{t-i} = \sum_{i=1}^{\infty} \theta_i W_{t-i} + W_t$$

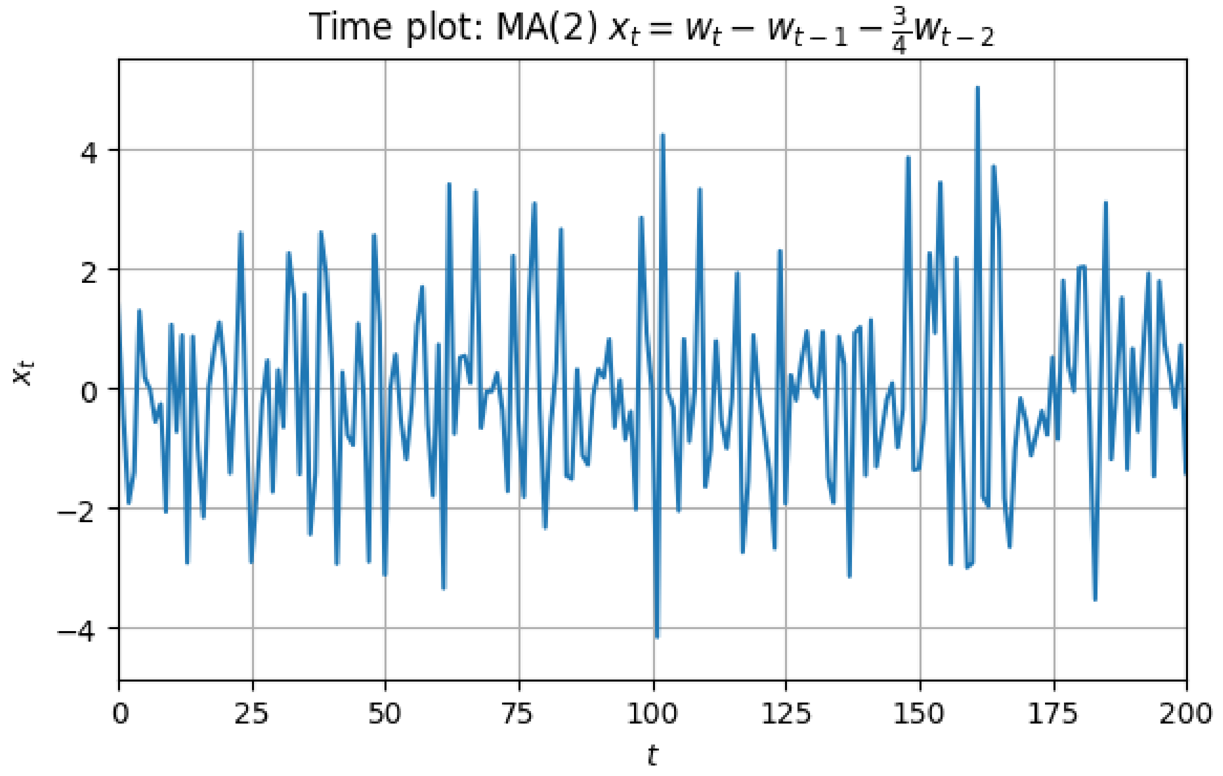
with $\theta_i = \phi^i$

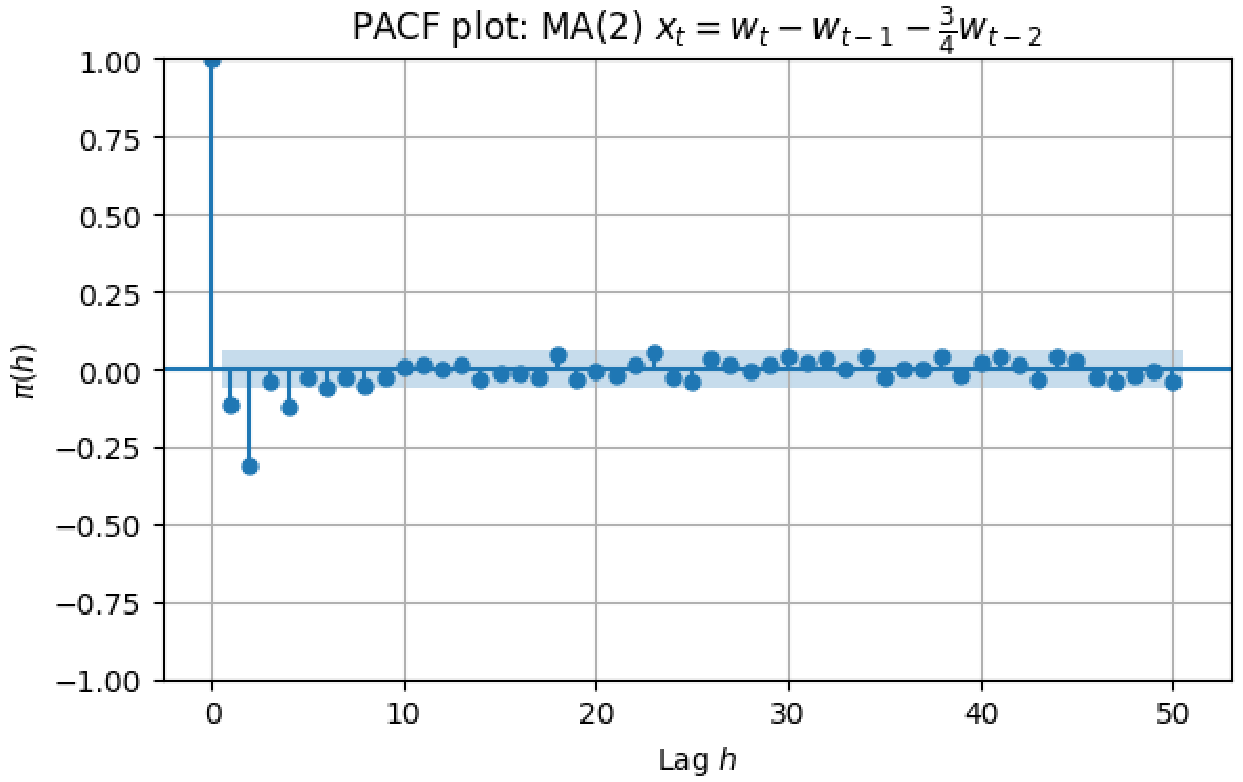
Model	**ACF** $\boldsymbol{\rho_h}$	**PACF** π_h
AR(p)	Infinite exponential decay	Cut-off at $h = p$
MA(q)	Cut-off at $h = q$	Infinite exponential decay

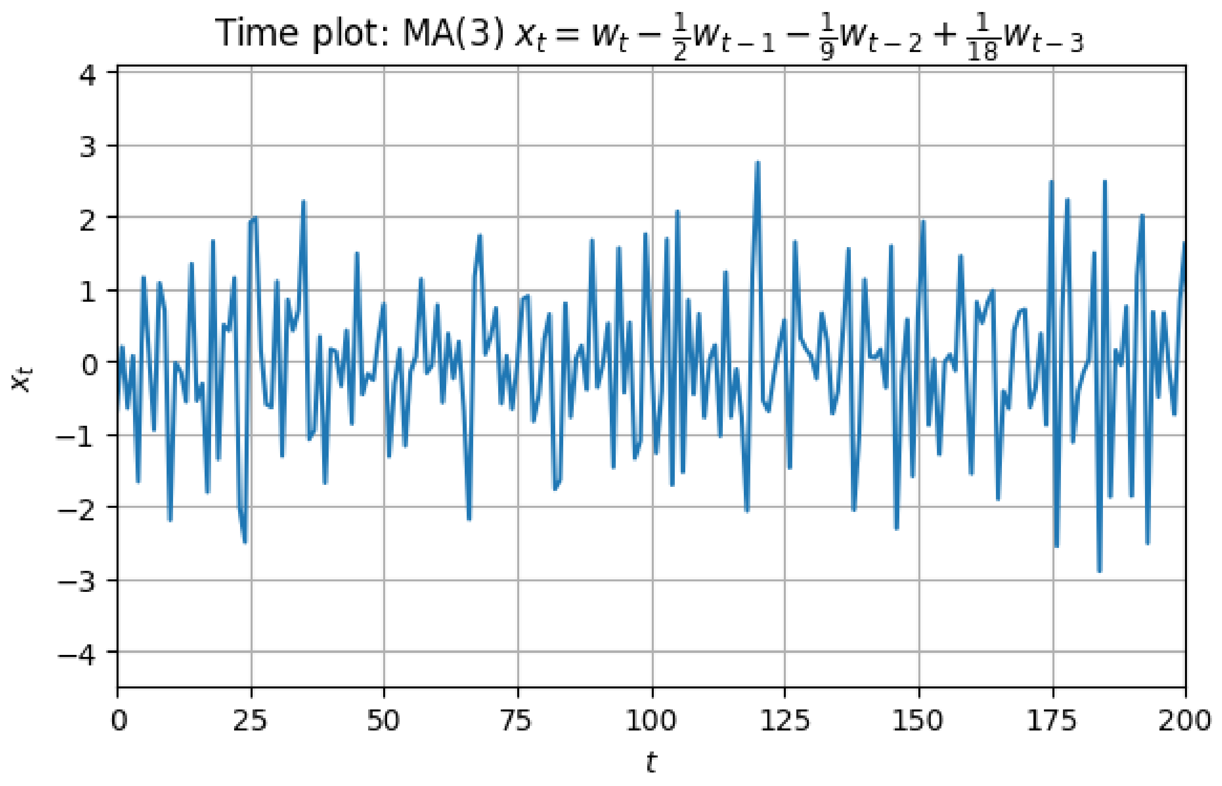
Partial auto-correlation function (PACF)

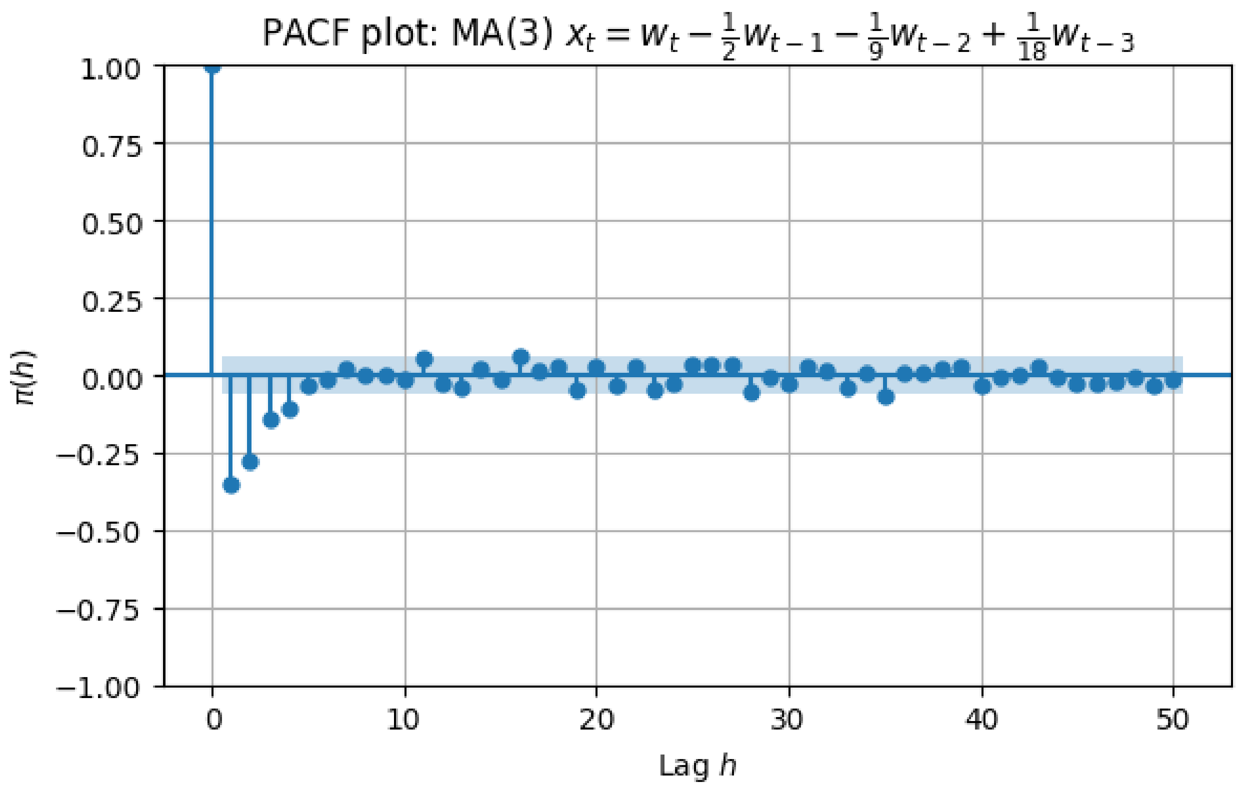


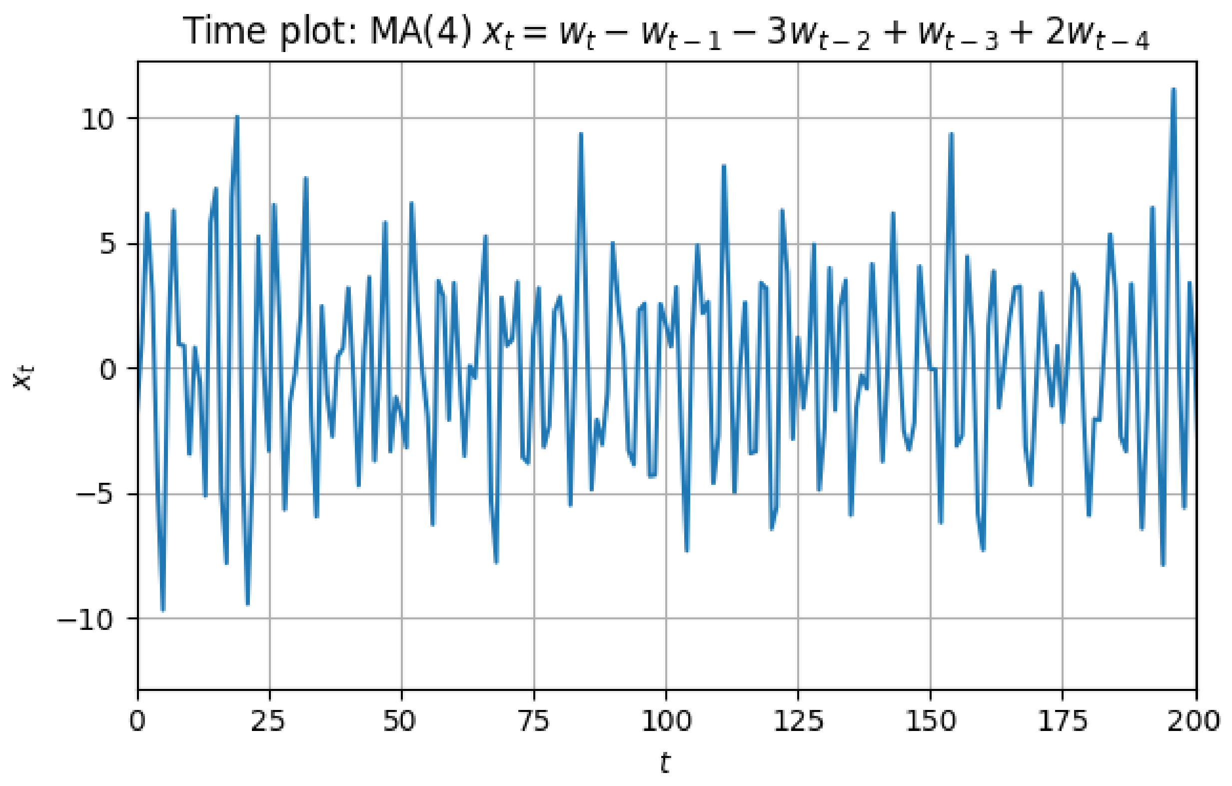


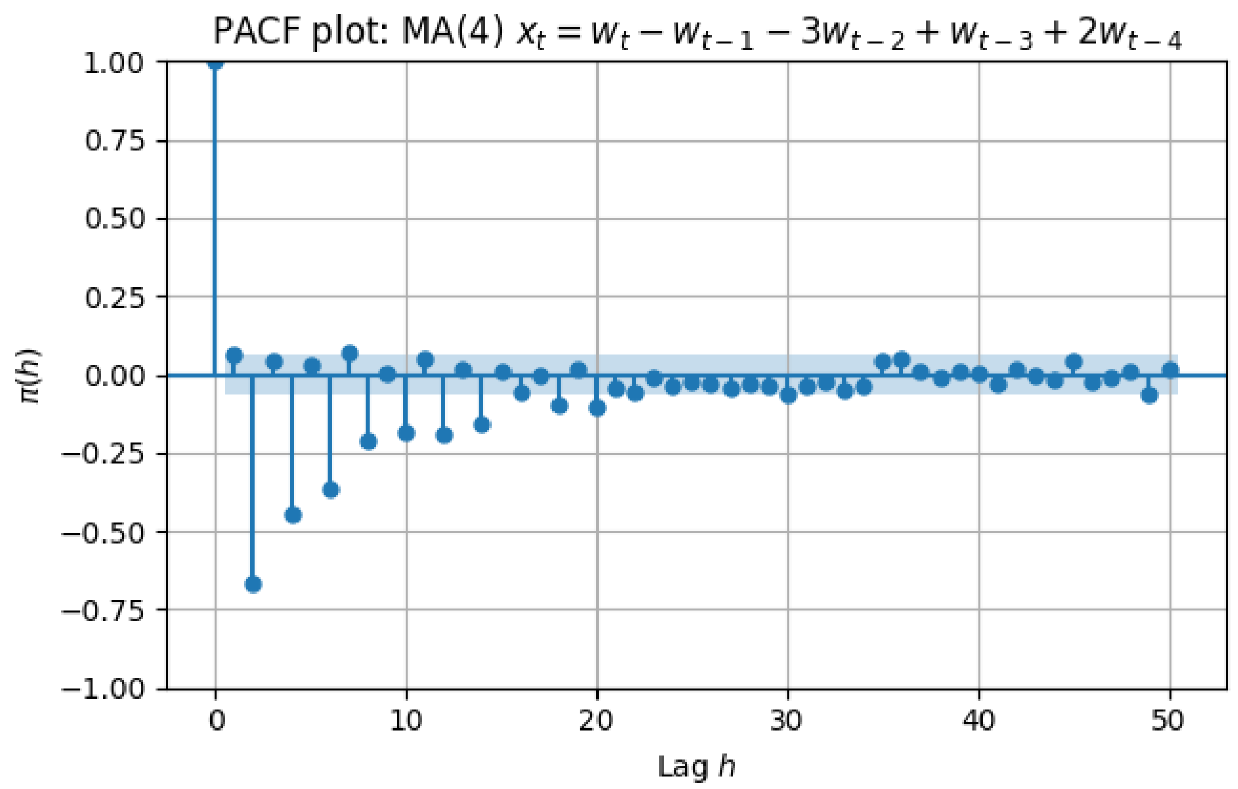












Invertibility of MA(1)

Consider the MA(1) processes $X_t = \theta W_{t-1} + W_t$ and $Y_t = \frac{1}{\theta} W_{t-1} + W_t$ with $W_t \sim WN\left(0, \sigma^2\right)$

Both result in the same ACF $\rho_h = \begin{cases} 1 & h = 0 \\ \frac{\theta}{\theta^2+1} & h = 1 \\ 0 & h > 1 \end{cases}$, which **prevents the unique identification of the MA parameters**.

Rewrite both processes as AR(∞) processes:

- $X_t = \theta(X_{t-1} - \theta W_{t-2}) + W_t = \sum_{i=1}^{\infty} (-1)^{i+1} \theta^i X_{t-i} + W_t = \sum_{i=1}^{\infty} -\phi^i X_{t-i} + W_t$ with $\phi = -\theta$
- $Y_t = \frac{1}{\theta}\left(Y_{t-1} - \frac{1}{\theta} W_{t-2}\right) + W_t = \sum_{i=1}^{\infty} \frac{(-1)^{i+1}}{\theta^i} Y_{t-i} + W_t = \sum_{i=1}^{\infty} -\phi^i Y_{t-i} + W_t$ with $\phi = -\frac{1}{\theta}$

Only the AR(∞) satisfying $|\phi| < 1$ will **converge**. We say the corresponding MA process is **invertible**.

For a given ACF, we ensure the unique MA representation by considering **only invertible MA processes**.

Invertibility of MA(q)

Different MA(q) parameters can lead to the same ACF, **preventing their unique identification.**

The zero-mean MA(q) equation can be rewritten as $X_t = \sum_{i=1}^{q} \theta_i W_{t-i} + W_t = \left(1 + \sum_{i=1}^{q} \theta_i B^i\right) W_t = \Theta(B) W_t$
where $\Theta(B) = \left(1 + \sum_{i=1}^{q} \theta_i B^i\right)$ is called the **characteristic polynomial**.

Theorem: A MA(q) process is invertible $\Leftrightarrow$ all roots of $\boldsymbol{\Theta(B)}$ are strictly greater than 1 in absolute value.

By imposing the **invertibility condition**, we ensure that there is a unique MA(q) process for a given ACF.

Examples (cf. figure with MA examples):

- MA(1) with $\Theta(B) = 1 - \frac{1}{2}B$, the root is $r_1 = 2 \Rightarrow$ invertible
- MA(2) with $\Theta(B) = 1 - B - \frac{3}{4}B^2 = \frac{3}{4}(B+2)(B-\frac{2}{3})$, the roots are $r_1 = -2$ and $r_2 = \frac{2}{3} \Rightarrow$ not invertible
- MA(3) with $\Theta(B) = 1 - \frac{1}{2}B - \frac{1}{9}B^2 + \frac{1}{18}B^3 = \frac{1}{18}(B-3)(B+3)(B-2)$, then $r_{1,2} = \pm 3$ and $r_3 = 2 \Rightarrow$ invertible
- MA(4) with $\Theta(B) = 1 - B - 3B^2 + B^3 + 2B^4 = (2B-1)(B+1)^2(B-1)$, then $r_{1,2} = \pm 1$, and $r_3 = \frac{1}{2} \Rightarrow$ not invertible

(r_4 is a complex root, which we will discuss in following lectures)

Fitting a MA model

Given a **stationary** time series realization $\{x_1, x_2, \dots x_n\}$ with **zero-mean**,

1. Is the moving average model **appropriate**?
   - ACF shows a cut-off at some lag q after which the ACF is close to zero.
   - PACF shows an exponential (sinusoidal) decay pattern.
2. Select model order based on the ACF, favoring smaller q values.
3. Estimate the MA(q) model **coefficients**.
   - Conditional sum of squares
   - Maximum-likelihood estimation
4. Analyze **residuals** to evaluate model goodness of fit (discussed in following lectures).

Estimating the coefficients of an MA(q) process

No (efficient) explicit estimators available to estimate the MA(q) model coefficients since W_t are **unobservable**. Exploiting the relation between the model parameters and the ACF is **inefficient**.

Considering a **zero-mean stationary MA(q)** realization $\{x_1, x_2, \dots x_n\}$,

Conditional sum of squares (CSS): $\hat{\boldsymbol{\theta}} = \begin{pmatrix} \hat{\theta}_1 \\ \hat{\theta}_2 \\ \vdots \\ \hat{\theta}_q \end{pmatrix} = \min_{\boldsymbol{\theta}} \sum_{i=q+1}^{n} \left(x_i - \sum_{j=1}^{q} \theta_i \hat{w}_{i-j}(\boldsymbol{\theta}) \right)$

- Express $\hat{w}_t(\boldsymbol{\theta})$in terms of x_t and θ_t only.
- Recursive computation of the error terms $\hat{w}_t(\boldsymbol{\theta})$ starting with $\hat{w}_0 = 0$.
- No closed-form solution, use numerical optimization.

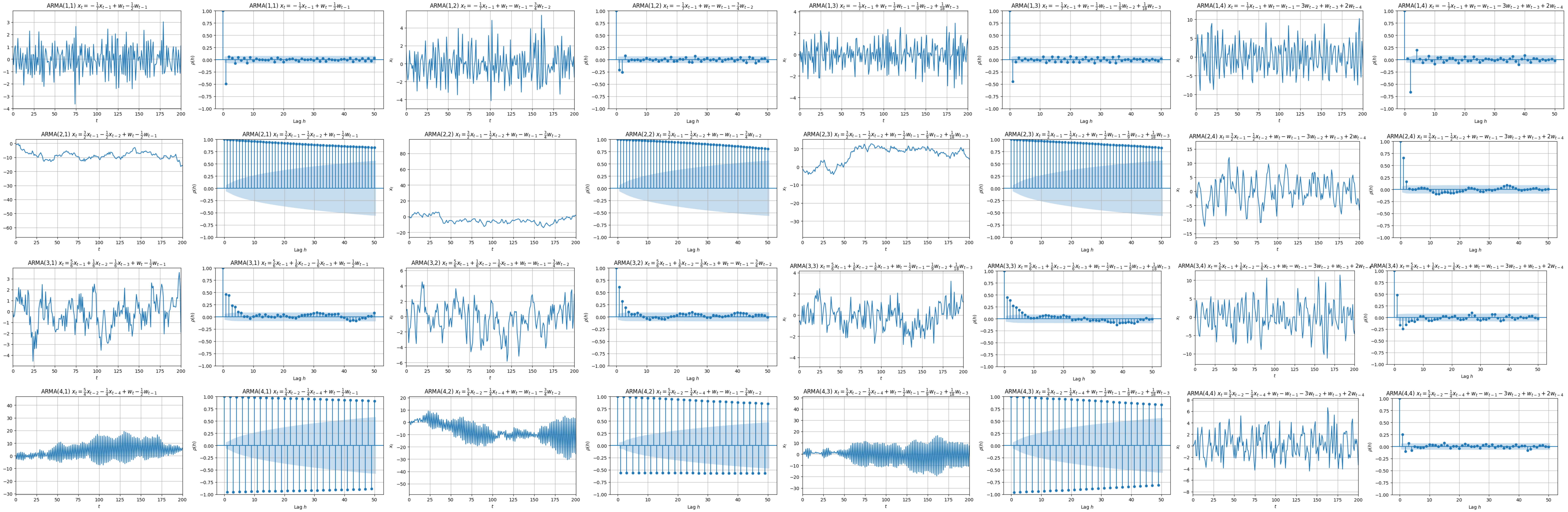
Maximum-likelihood estimation (MLE): find $\hat{\boldsymbol{\theta}}$ that maximizes the likelihood of the data under the MA(q) model, using CSS to obtain a first estimate of the coefficients.

$$\hat{\boldsymbol{\theta}} = \max_{\boldsymbol{\theta}} \prod_{i=q+1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{\hat{w}_i(\boldsymbol{\theta})^2}{2\sigma^2}\right) \text{ with } \hat{w}_i(\boldsymbol{\theta}) = x_i - \sum_{j=1}^{q} \theta_i \hat{w}_{i-j}(\boldsymbol{\theta}) \text{ and assuming } w_i \sim \mathcal{N}(0, \sigma^2)$$

Autoregressive moving average (ARMA) model

ARMA(p,q): $X_t = c + \sum_{i=1}^{p} \phi_i X_{t-i} + \sum_{i=0}^{q} \theta_i W_{t-i}$ with $\theta_0 = 1$ and $W_t \sim WN\left(0, \sigma^2\right)$ and c a constant.

We focus on the **zero-mean** ARMA model ($c = 0$) since we can always choose to model $\mathrm{Y_t} = X_t - \mu_X$.



Properties of ARMA(p,q)

ARMA processes can capture both **long-term dependencies** (AR part) and **short-term shocks** (MA part).

The zero-mean ARMA(p,q) can be rewritten with the **characteristic polynomials** of the AR and MA parts:

$$\Phi(B)X_t = \left(1 - \sum_{i=1}^{p} \phi_i B^i\right) X_t = \left(1 + \sum_{i=1}^{q} \theta_i B^i\right) W_t = \Theta(B) W_t$$

We **only** consider ARMA models **without common roots** between $\Phi(\mathrm{B})$ and $\Theta(\mathrm{B})$ (**no parameter redundancy**).

Theorem: An ARMA(p,q) process is stationary $\Leftrightarrow$ all roots $\mathbf{r}$ of $\boldsymbol{\Phi}(\mathbf{B})$ lie outside the unit circle ($|\mathbf{r}| > \mathbf{1}$).

- Stationarity depends on the AR part since any MA(q) is stationary.

Theorem: An ARMA(p,q) process is invertible $\Leftrightarrow$ all roots $\mathbf{r}$ of $\boldsymbol{\Theta}(\mathbf{B})$ lie outside the unit circle ($|\mathbf{r}| > \mathbf{1}$).

- Invertibility depends on the MA part since any AR(p) is uniquely represented.
- The invertibility condition ensures that an ARMA(p, q) process can be uniquely represented.

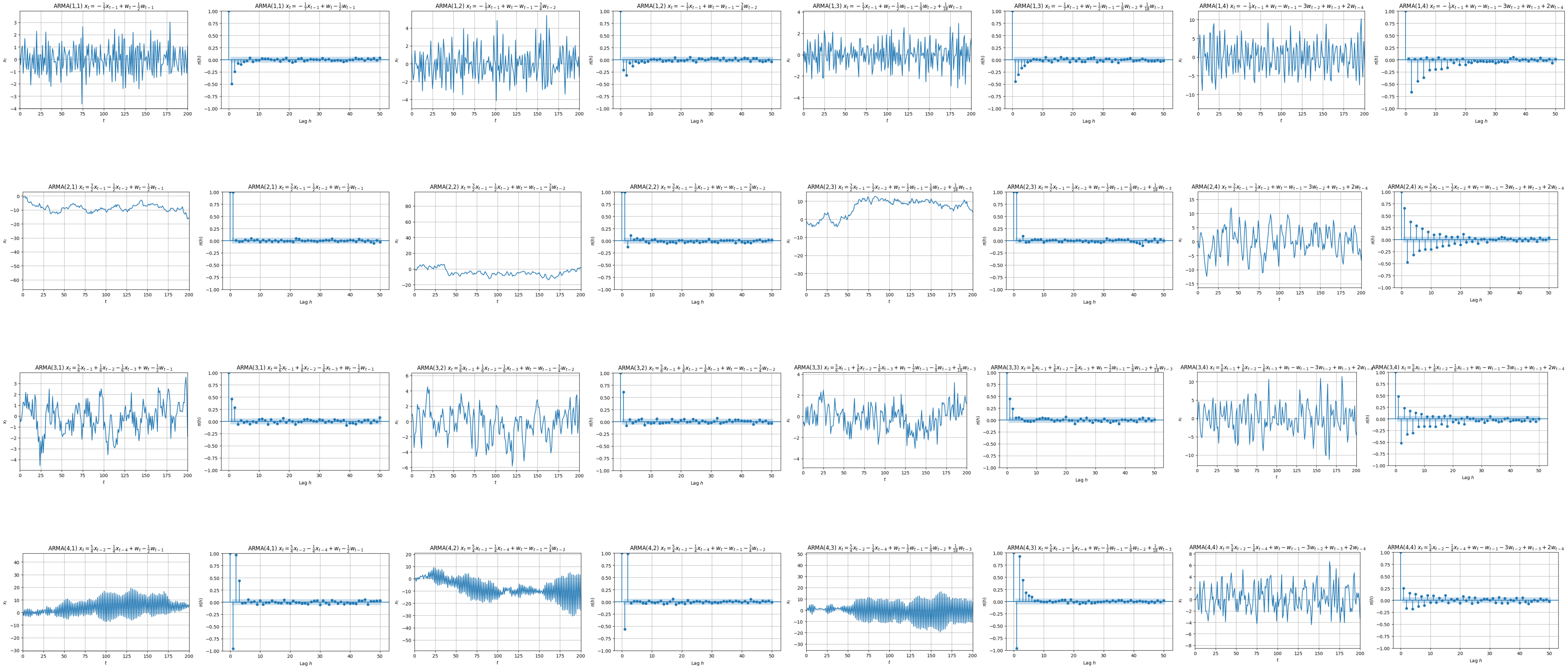
ACF and PACF of ARMA(p,q)

Theorem: an ARMA(p, q) process can be expressed as either an AR(∞) process or an MA(∞) process.

- Expand either the MA(q) part as an AR(∞) process or the AR(p) part as an MA(∞) process.
- Typically, ARMA(p,q) requires fewer parameters than pure AR or MA processes would.

Model	**ACF** $\boldsymbol{\rho_h}$	**PACF** π_h
AR(p)	Infinite exponential decay	Cut-off at $h = p$
MA(q)	Cut-off at $h = q$	Infinite exponential decay
ARMA(p,q)	Infinite exponential decay with drop-off for $h > q$	Infinite exponential decay with drop-off for $h > p$

Partial auto-correlation function (PACF)



Fitting an ARMA model

$AR(p) = ARMA(p, 0)$
$MA(q) = ARMA(0, q)$

$p = 2 \quad q = 1$
$(2,1), (2,0), (0,1), (1,1), (0,1)$

Given a **stationary** time series realization $\{x_1, x_2, \dots x_n\}$ with **zero-mean**,

1. Is the ARMA model **appropriate**?
   - Neither the ACF nor the PACF show a pronounced cut-off.
   - ACF shows a drop-off after some lag q and an exponential decay pattern.
   - PACF shows a drop-off after some lag p and an exponential decay pattern.
2. Select p and q based on the ACF and PACF, favoring smaller values.
   - With unclear patterns, test multiple pairs and calculate information criteria (discussed in following lectures)
3. Estimate the ARMA(p,q) model **coefficients**.
   - Conditional sum of squares
   - Maximum-likelihood estimation
4. Analyze **residuals** to evaluate model goodness of fit (discussed in following lectures).

Autoregressive integrated moving average (ARIMA) model

if not stationary
differencing is required

X_t follows an ARIMA(p,d,q) model if its d$^{\text{th}}$ order **differencing** $\nabla^d X_t$ is an ARMA(p,q) process:

$$\Phi(B)\nabla^d X_t = \Phi(B)(1-B)^d X_t = c + \Theta(B)W_t$$

We say that X_t is given by the d$^{\text{th}}$ order **integration** of an ARMA(p,q).

Differencing enables to handle non-stationarity. Setting $d = 1$ or 2 is typically sufficient to achieve stationarity.

Model	ARIMA equivalent
White noise	ARIMA(0,0,0) with $c = 0$
Random walk	ARIMA(0,1,0) with $c = 0$
Random walk with drift	ARIMA(0,1,0) with $c \neq 0$
AR(p)	ARIMA(p,0,0)
MA(q)	ARIMA(0,0,q)
ARMA(p,q)	ARIMA(p,0,q)

ARIMA	Long-term forecast
$c = 0$ and $d = 0$	Converge to 0
$c = 0$ and $d = 1$	Converge to a non-zero constant
$c = 0$ and $d = 2$	Follow a linear trend
$c \neq 0$ and $d = 0$	Converge to the historical mean
$c \neq 0$ and $d = 1$	Follow a linear trend
$c \neq 0$ and $d = 2$	Follow a quadratic trend

Seasonal autoregressive integrated moving average (SARIMA) model

Seasonal ARMA with **seasonal orders** P, Q and **period** s, ARMA(p,q)(P,Q)$_s$:

$$\Phi(B)\acute{\Phi}(B^s)X_t = c + \Theta(B)\acute{\Theta}(B^s)W_t$$

$$\text{with } \acute{\Phi}(B^s) = \left(1 - \sum_{i=1}^{p} \acute{\phi}_i B^{i\cdot s}\right) \text{ and } \acute{\Theta}(B^s) = \left(1 + \sum_{i=1}^{q} \acute{\theta}_i B^{i\cdot s}\right)$$

Seasonal orders can be inferred from ACF and PACF but for **lags on the seasonal scale** i.e., in multiples of s.

To **handle non-stationarity**, we add integration on both the original and seasonal scales, ARIMA(p,d,q)(P,D,Q)$_s$:

$$\Phi(B)\acute{\Phi}(B^s)\nabla^d\nabla_s^D X_t = \Phi(B)\acute{\Phi}(B^s)(1-B)^d(1-B^s)^D X_t = c + \Theta(B)\acute{\Theta}(B^s)W_t$$

Comments

- Setting $d = D = 1$ is typically sufficient to achieve stationarity.
- As with other models, the ACF and PACF help guide initial estimates for the orders p,q,P,Q.
- In practice different combinations of these values are tested.
- Best model is selected based on information criteria (discussed in following lectures).

Exercise

Simulate and analyze AR, MA, ARMA, ARIMA, SARIMA models

- Generate time series data then compute ACF and PACF.
- Determine model orders.
- Compare empiric ACF and PACF with theoretical counterpart.
- Estimate the model parameters from the generated data.

Model real-world time series

- Is differencing required?
- Determine model order from ACF and PACF (review this summary).
- Estimate the model parameters.