MATH 7339 - Machine Learning and Statistical Learning Theory 2

Section ARMA

1. AR, MA, ARMA
2. Stationarity, causality and invertibility
3. linear process representation of ARMA
4. Autocovariance of an ARMA process
5. Homogeneous linear difference equations

❑ AR(p): Autoregressive Models of order p.

A class of models closely related to the random walk are the **autoregressive models** (**AR**). An autoregressive model is defined so that the current location is a linear combination of previous locations plus a random term (Gaussian white noise).

The **AR(p)** model is

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + W_t$$

where $W_t \sim WN(0, \sigma^2)$

Let B be the **backshift** operator, $BX_t := X_{t-1}$. The above two AR(p) descriptions can be written as

$$(1 - \phi_1 B - \cdots - \phi_p B^p) X_t = W_t \qquad \text{or } \phi(B) X_t = W_t$$

Question: Under what condition(s) is the random walk a special case of an AR model?

❑ MA(q): Moving Average Models or order q

One way to think about **Moving Average models** is to take a sliding window and take a weighted average of a white noise process for everything in the window. So, start with a white noise process, $\{W_t\}$. Then a moving average of order q is of the following form

$$X_t = W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q}$$

where $W_t \sim WN(0, \sigma^2)$

Use backshift operator

$$X_t = \left(1 + \theta_1 B + \cdots \theta_q B^q\right) W_t$$ or $X_t = \theta(B) W_t$

❑ ARMA(p,q): Autoregressive moving average models

An **ARMA(p,q)** process $\{X_t\}$ is a *stationary* process that satisfies

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q}$$

where $W_t \sim WN(0, \sigma^2)$

or $\phi(B)X_t = \theta(B)W_t$

ARMA processes can accurately approximate many stationary processes.

Theorem: For any stationary process with autocovariance γ and any $k > 0$, there is an ARMA process $\{X_t\}$ for which

$$\gamma_X(h) = \gamma(h) \qquad \text{for } h = 0,1,\ldots,k$$

AR(p) = ARMA(p,0).

MA(q) = ARMA(0,q).

Usually, we insist that $\phi_p, \ \theta_q \neq 0$ and that the polynomials

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p$$

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q$$

have no common factors. This implies it is not a lower order ARMA model.

Example of parameter redundancy

A white noise process $X_t = W_t$

So $$X_t - X_{t-1} + 0.5X_{t-2} = W_t - W_{t-1} + 0.5W_{t-2}$$

So, $$(1 - B + 0.5B^2)X_t = (1 - B + 0.5B^2)W_t$$

This is in the form of an ARMA(2,2) process. But it is white noise.

There are a few issues with ARMA models:

- Parameter redundancy in models.
- AR models that depend on the future.
- MA models that are not unique..

To overcome these issues, we require some restrictions on the model parameters

AR(1) in terms of the back-shift operator

Assume that X_t is **stationary** solution to

$$X_t = \phi X_{t-1} + W_t$$

If $|\phi| < 1$, then

$$X_t = \sum_{j=0}^{\infty} \phi^j W_{t-j}$$

Let B be the **backshift** operator, $BX_t := X_{t-1}$. The above two AR(1) descriptions can be written as

$$(1 - \phi B)X_t = W_t \quad \text{and} \quad X_t = \sum_{j=0}^{\infty} \phi^j B^j W_t = \pi(B)W_t$$

Denote

$$\phi(B) = 1 - \phi B \quad \text{and} \quad \pi(B) = \sum_{j=0}^{\infty} \phi^j B^j$$

So, $\pi(B) = \phi(B)^{-1}$ as in the Taylor expansion of $\frac{1}{1-\phi B}$

❑ Causality

A linear process $\{X_t\}$ is **causal** (strictly, a **causal function** of $\{W_t\}$) if there is a

$$\psi(B) = \psi_0 + \psi_1 B + \cdots + \psi_t B^2 + \cdots$$

with $\displaystyle\sum_{j=0}^{\infty} |\psi_j| < \infty$ and $X_t = \psi(B) W_t$

Example.

- AR(1) is causal if $|\phi| < 1$.
- MA(q) is causal.

❑ Invertibility

A linear process $\{X_t\}$ is **Invertible** (strictly, a **invertible function** of $\{W_t\}$) if there is a

$$\pi(B) = \pi_0 + \pi_1 B + \cdots + \pi_t B + \cdots$$

with $\sum_{j=0}^{\infty} |\pi_j| < \infty$ and $W_t = \pi(B) X_t$

- Causality and Invertibility are properties of $\{X_t\}$ *and* $\{W_t\}$

Example.

- AR(1) is causal if $|\phi| < 1$.

Consider the MA(1) process defined by

$$X_t = W_t + \theta W_{t-1} = (1 + \theta B)W_t$$

So, if $|\theta| < 1$, we have the Tayler series expansion

$$W_t = \frac{1}{1 + \theta B} X_t = \sum_{j=0}^{\infty} (-\theta)^j B^j X_t$$

So, MA(1) is invertible if $|\theta| < 1$.

❑ AR(p): Stationarity and causality

Theorem: A (**unique**) **stationary** solution to $\phi(B)X_t = W_t$ exists iff

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p = 0 \text{ implies } |z| \neq 1$$

This AR(p) process is **causal** iff

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p = 0 \text{ implies } |z| > 1$$

❑ Calculating ψ for an AR(p): matching coefficients.

AR(p) model:

$$X_t = \phi_1 X_{t-1} + \cdots + \phi_p X_{t-p} + W_t$$

Equivalently, $\phi(B)X_t = W_t$, where

$$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B^p$$

Let us solve $\psi(B)$ such that

$$X_t = \psi(B)W_t$$

So, $\psi(B)\phi(B) = 1$

Equivalently,

$$(\psi_0 + \psi_1 B + \cdots + \psi_t B^t + \cdots)\left(1 - \phi_1 B - \cdots - \phi_p B^p\right) = 1$$

Equivalently, $\psi_0 = 1$

$$\psi_1 - \phi_1\psi_0 = 0$$

$$\psi_2 - \phi_1\psi_1 - \phi_2\psi_0 = 0$$

……

$$\psi_k - \sum_{s=1}^{p} \phi_s \psi_{k-s} = 0 \qquad \text{for } k > 0$$

Equivalently, $\psi_j = 0 \text{ for } j < 0$

$$\psi_0 = 1$$

$$\phi(B)\psi_j = 0$$

We can solve these **linear difference equations** in several ways:

- numerically, or
- by guessing the form of a solution and using an inductive proof, or
- by using the theory of linear difference equations.

Calculating ψ for an ARMA(p,q): matching coefficients

Example: $\phi(B)X_t = \theta(B)W_t$

$$(1 + 0.25B^2)X_t = (1 + 0.2B)W_t$$

Let us solve $\psi(B)$ such that

$$X_t = \psi(B)W_t$$

So,

$$(1 + 0.2B) = (1 + 0.25B^2)(\psi_0 + \psi_1 B + \cdots + \psi_t B^t + \cdots)$$

Compare the same degree of B, we have the **first order differential equation of** ψ_i

as $\theta_j = \phi(B)\psi_j$ with $\theta_0 = 1 \; and \; \theta_j = 0 \; otherwise.$

We can use the θ_j to give the initial conditions and solve it using the theory of homogeneous difference equations.

$$\psi_j = (1, \frac{1}{5}, -\frac{1}{4}, -\frac{1}{20}, \frac{1}{16}, \frac{1}{80}, -\frac{1}{64}, -\frac{1}{320}, \ldots)$$

The method is the same for the general case.

ARMA(p,q):Stationarity, causality, and invertibility

Theorem: If ϕ and θ have no common factors, the ARMA(p,q) process is the (unique) solution to

$$\phi(B)X_t = \theta(B)W_t$$

- This ARMA process is **stationary** iff the roots of $\phi(z)$ avoid the unit circle, i.e.,

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p = 0 \text{ implies } |z| \neq 1$$

- This ARMA process is **casual** iff the roots of $\phi(z)$ outside the unit circle, i.e.,

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p = 0 \text{ implies } |z| > 1$$

- This ARMA process is **invertible** iff the roots of $\theta(z)$ outside the unit circle, i.e.,

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q = 0 \text{ implies } |z| > 1$$

Example: (ARMA(1,1))

$$(1 - 1.5B)X_t = (1 + 0.2B)W_t$$

1. ϕ and θ have no common factors, and ϕ's root is at $\frac{2}{3}$, which is not on the unit circle, so $\{X_t\}$ is an ARMA(1,1) process.
2. ϕ's root (at 2/3) is inside the unit circle, so $\{X_t\}$ is not causal.
3. θ's root is at –5, which is outside the unit circle, $\{X_t\}$ is invertible.

Example: (ARMA(2,1))

$$(1 + 0.25B)X_t = (1 + 2B)W_t$$

1. ϕ and θ have no common factors, and ϕ's root is at $\pm 2i$, which is not on the unit circle, so $\{X_t\}$ is an ARMA(2,1) process.
2. ϕ's root (at $\pm 2i$) is outside the unit circle, so $\{X_t\}$ is causal.
3. θ's root is at –1/2, which is inside the unit circle, $\{X_t\}$ is not invertible.

❑ Autocovariance functions of linear processes

Suppose the mean of X_t is zero. Consider a **linear** process:

$$X_t = \psi(B)W_t$$

where $\psi(B) = \psi_0 + \psi_1 B + \cdots + \psi_n B^n + \cdots$, and $W_t = WN(0, \sigma^2)$

$$\gamma(h) = E(X_t X_{t+h})$$

$$= E[(\psi_0 W_t + \psi_1 W_{t-1} + \cdots + \psi_n W_{t-n} + \cdots)(\psi_0 W_{t+h} + \psi_1 W_{t+h-1} + \cdots)]$$

$$= \sigma^2(\psi_0\psi_h + \psi_1\psi_{h+1} + \psi_2\psi_{h+2} + \cdots)$$

$$= \sigma^2 \sum_{j=0}^{\infty} \psi_j \psi_{j+h} \qquad \text{for } h \geq 0$$

Example: MA(q) process

$$X_t = W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q} = \theta(B) W_t$$

The **autocovariance**:

$$\gamma(h) = \begin{cases} \sigma^2 \sum_{j=0}^{q-h} \theta_j \theta_{j+h} & \text{for } h \leq q \\ 0 & \text{for } h > q \end{cases}$$

The **autocorrelation** function (**ACF**) of an MA(q) model is

$$\rho(h) = \frac{\gamma(h)}{\gamma(0)} = \begin{cases} \dfrac{\sum_{j=0}^{q-h} \theta_j \theta_{j+h}}{1 + \theta_1^2 + \cdots + \theta_q^2} & \text{for } h \leq q \\ 0 & \text{for } h > q \end{cases}$$

The ACF will be zero for lags greater than q. Thus, the ACF provides information about the order of the dependence for a MA model.

❑ Autocovariance functions ARMA(p,q) process

$$\phi(B)X_t = \theta(B)W_t$$

Method 1. Write causal ARMA as $X_t = \psi(B)W_t$, then use the above result for $\gamma(h)$.

Method 2. $$\phi(B)X_t = \theta(B)W_t$$

$$X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p} = W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q}$$

So

$$E\left[\left(X_t - \phi_1 X_{t-1} - \cdots - \phi_p X_{t-p}\right)X_{t-h}\right] = E\left[\left(W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q}\right)X_{t-h}\right]$$

So,

$$\gamma(h) - \phi_1\gamma(h-1) - \cdots - \phi_p\gamma(h-p) = \sigma^2 \sum_{j=0}^{q-h} \theta_{h+j}\psi_j$$

This is a linear difference equation.

❑ Homogeneous linear difference equations.

Homogeneous linear difference equations of order k:

$$a_0 x_t + a_1 x_{t-1} + \cdots + a_k x_{t-k} = 0$$

Equivalently,

$$\left(a_0 + a_1 B + \cdots + a_k B^k\right) x_t = 0 \qquad \text{or} \quad a(B) x^t = 0$$

auxiliary equation: $a_0 + a_1 z + \cdots + a_k z^k = 0$

$$(z - z_1)(z - z_2) \cdots (z - z_k) = 0$$

The roots of this characteristic polynomial are $z_i \in \mathbb{C}$.

$$(B - z_1)(B - z_2) \cdots (B - z_k) x_t = 0$$

So, $(B - z_1) x_t = 0$

Three cases:

1. The z_i are real and distinct.
2. The z_i are complex and distinct.
3. Some z_i are repeated.

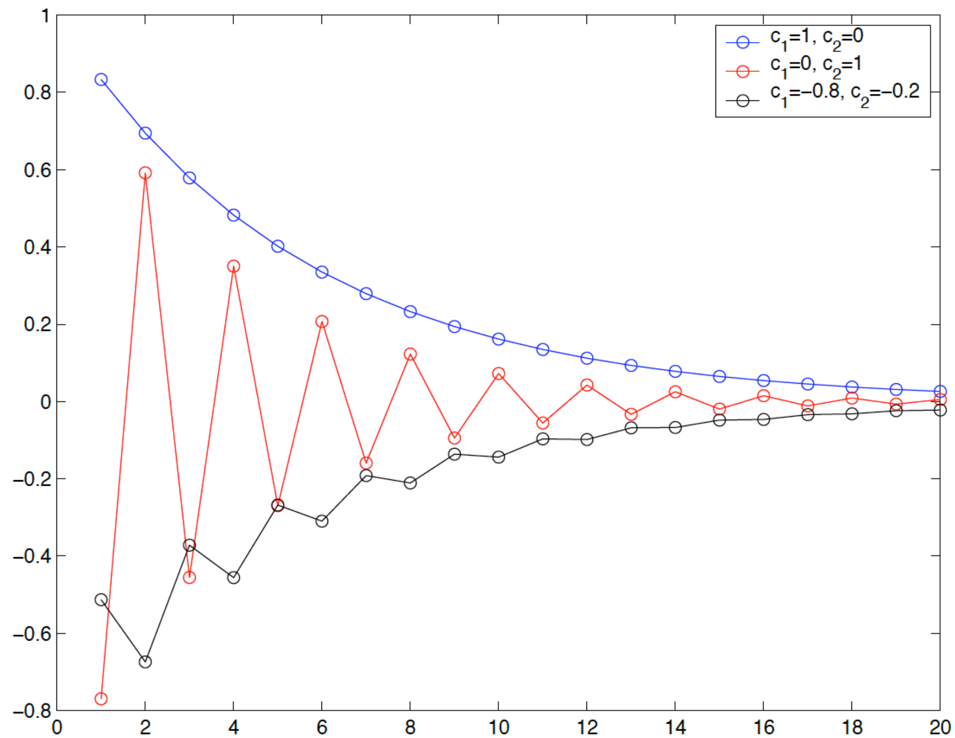
1. The z_i are real and distinct.

$$x_t = c_1 z_1^{-t} + c_2 z_2^{-t} + \cdots + c_k z_k^{-t}$$

Example:

$$z_1 = 1.2, z_2 = -1.3$$

$$x_t = c_1 z_1^{-t} + c_2 z_2^{-t}$$



2. The z_i are complex and distinct.

$$x_t = c_1 z_1^{-t} + c_2 z_2^{-t} + \cdots + c_k z_k^{-t}$$

Suppose z_1, then there is a complex conjugate root $z_j = \overline{z_1}$

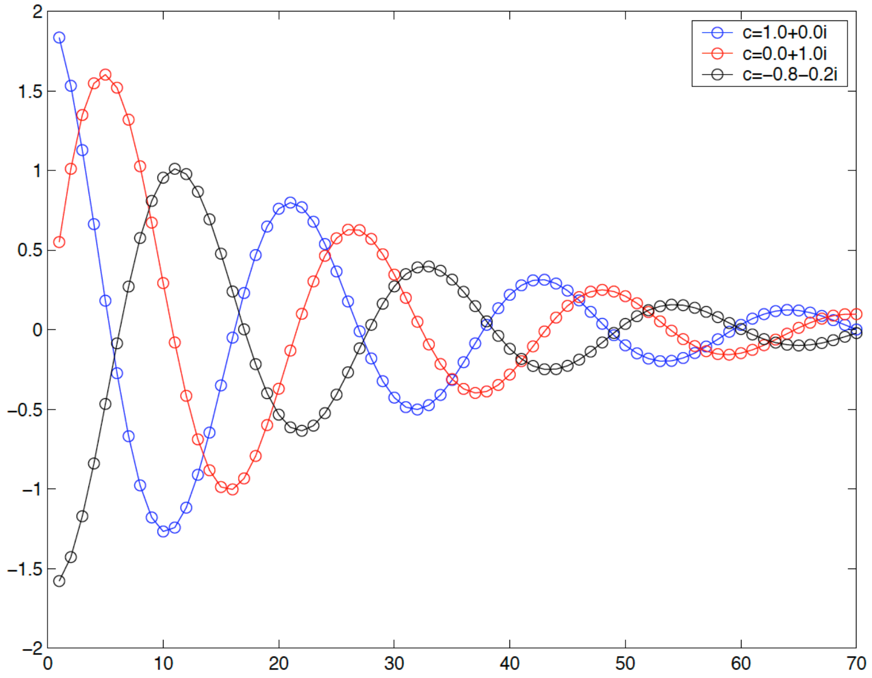
For example,

$$x_t = c_1 z_1^{-t} + \overline{c}_1 \overline{z}_1^{-t} = 2r|z_1|^{-t}\cos(\omega t - \alpha)$$

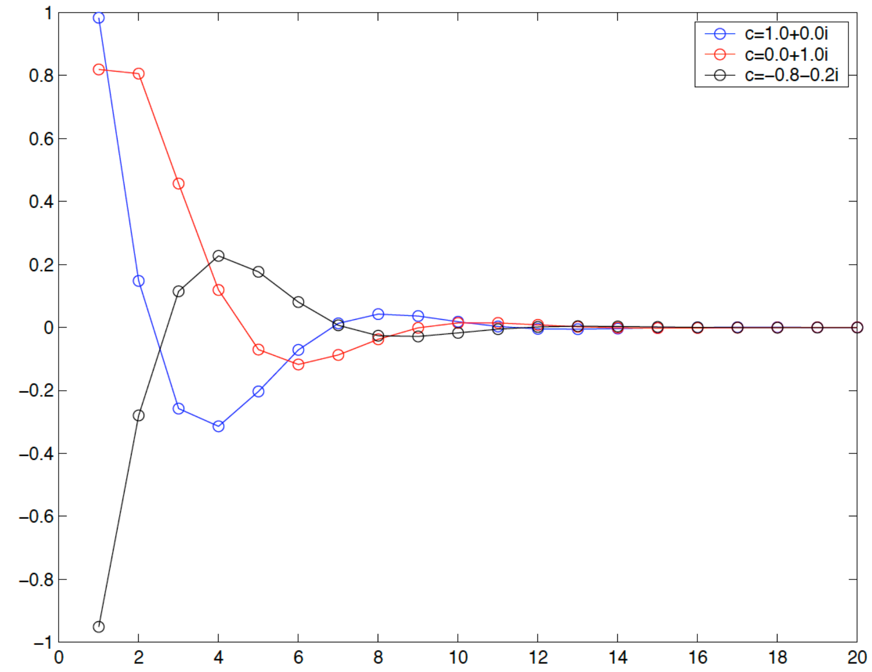
Here $z_1 = |z_1|e^{i\omega}$ and $c_1 = re^{i\alpha}$

Examples: $x_t = c_1 z_1^{-t} + \bar{c}_1 \bar{z}_1^{-t}$

$z_1 = 1 + 0.1i, z_2 = 1 - 0.1i$



$z_1 = 1.2 + i, z_2 = 1.2 - i$



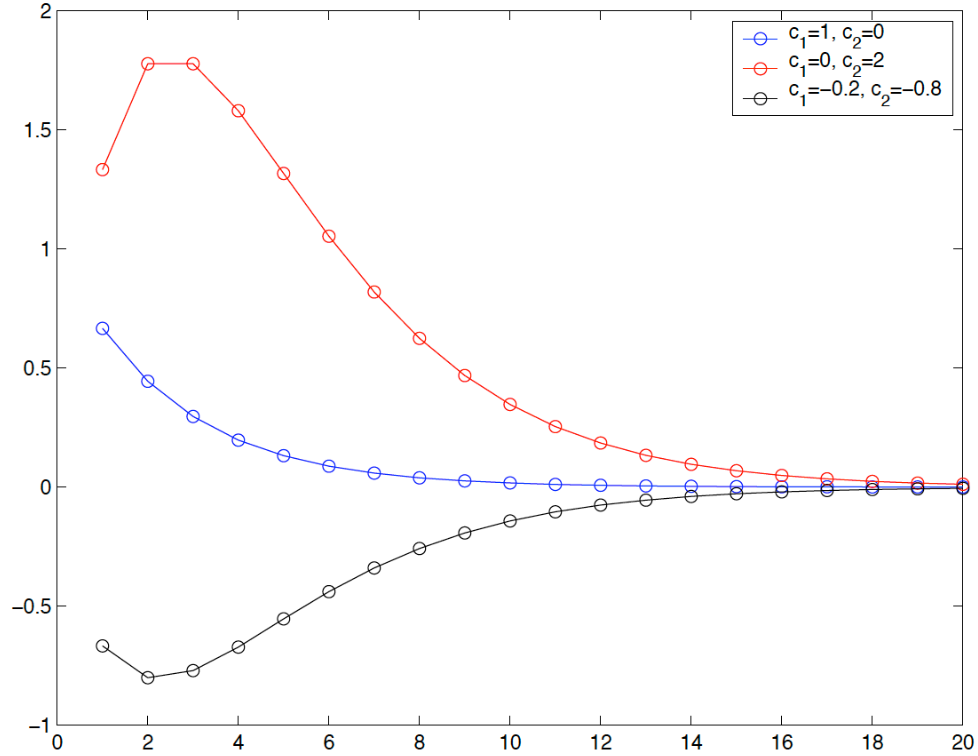
3. Some z_i are repeated.

$$(B - z_1)^m x_t = 0$$

Check $(c_1 + c_2 t + \cdots + c_{m-1} t^{m-1}) z_1^{-t}$ is a solution

Example $z_1 = z_2 = 1.5$.

$$(c_1 + c_2 t) z_1^{-t}$$



Find Autocovariance functions of ARMA processes

$$(1+0.25B^2)X_t = (1+0.2B)W_t \qquad \Leftrightarrow \qquad X_t = \psi(B)W_t$$

$$\psi_j = \left(1, \frac{1}{5}, -\frac{1}{4}, -\frac{1}{20}, \frac{1}{16}, \frac{1}{80}, -\frac{1}{64}, -\frac{1}{320}, \dots\right).$$

$$\gamma(h) - \phi_1\gamma(h-1) - \phi_2\gamma(h-2) = \sigma_w^2 \sum_{j=0}^{q-h} \theta_{h+j}\psi_j$$

$$\Leftrightarrow \gamma(h) + 0.25\gamma(h-2) = \begin{cases} \sigma_w^2\left(\psi_0 + 0.2\psi_1\right) & \text{if } h = 0, \\ 0.2\sigma_w^2\psi_0 & \text{if } h = 1, \\ 0 & \text{otherwise.} \end{cases}$$

We have the homogeneous linear difference equation

$$\gamma(h) + 0.25\gamma(h-2) = 0$$

for $h \geq 2$, with initial conditions

$$\gamma(0) + 0.25\gamma(-2) = \sigma_w^2\left(1 + 1/25\right)$$
$$\gamma(1) + 0.25\gamma(-1) = \sigma_w^2/5.$$

Homogeneous lin. diff. eqn:

$$\gamma(h) + 0.25\gamma(h-2) = 0.$$

The characteristic polynomial is

$$1 + 0.25z^2 = \frac{1}{4}\left(4 + z^2\right) = \frac{1}{4}(z-2i)(z+2i),$$

which has roots at $z_1 = 2e^{i\pi/2}, \bar{z_1} = 2e^{-i\pi/2}$.

The solution is of the form

$$\gamma(h) = cz_1^{-h} + \bar{c}\bar{z_1}^{-h}.$$

$z_1 = 2e^{i\pi/2}, \bar{z_1} = 2e^{-i\pi/2}, c = |c|e^{i\theta}$.

We have

$$\begin{aligned}\gamma(h) &= cz_1^{-h} + \bar{c}\bar{z_1}^{-h} \\ &= 2^{-h}\left(|c|e^{i(\theta-h\pi/2)} + |c|e^{i(-\theta+h\pi/2)}\right) \\ &= c_1 2^{-h}\cos\left(\frac{h\pi}{2} - \theta\right).\end{aligned}$$

And we determine c_1, θ from the initial conditions

$$\gamma(0) + 0.25\gamma(-2) = \sigma_w^2\left(1 + 1/25\right)$$
$$\gamma(1) + 0.25\gamma(-1) = \sigma_w^2/5.$$

We determine c_1, θ from the initial conditions:

We plug

$$\gamma(0) = c_1 \cos(\theta)$$

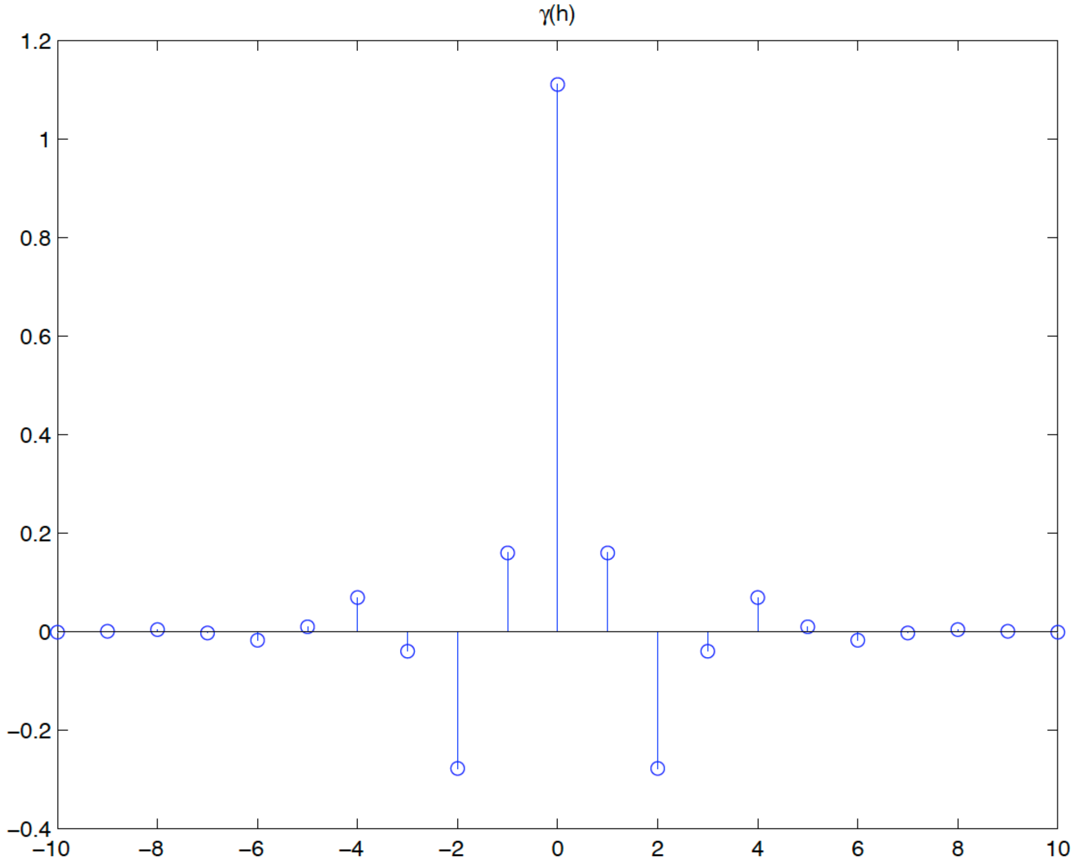
$$\gamma(1) = \frac{c_1}{2} \sin(\theta)$$

$$\gamma(2) = -\frac{c_1}{4} \cos(\theta)$$

into

$$\gamma(0) + 0.25\gamma(2) = \sigma_w^2 (1 + 1/25)$$

$$1.25\gamma(1) = \sigma_w^2/5.$$



Notations:

One way to remove linear connections is through linear regression.

Let $\hat{x}_{t+h}$ denote the regression of x_{t+h} on $\{x_{t+h-1}, x_{t+h-2}, \dots, x_{t+1}\}$

$$\hat{x}_{t+h} = \beta_1 x_{t+h-1} + \beta_2 x_{t+h-2} + \cdots + \beta_{h-1} x_{t+1}$$

Here we do not include the intercept assuming the mean of x_t is zero.
Otherwise, replace x_t with $x_t - \mu_x$.

Let $\hat{x}_t$ denote the regression of x_t on $\{x_{t+1}, x_{t+2}, \dots, x_{t+h-1}\}$

$$\hat{x}_t = \beta_1 x_{t+1} + \beta_2 x_{t+2} + \cdots + \beta_{h-1} x_{t+h-1}$$

Use OLS to estimate the parameters.

$$X_t, X_{t+1}, X_{t+2}, \dots, X_{t+h-1}, X_{t+h,}$$

➢ partial autocorrelation function (PACF)

ACF provides considerable information for MA(q). But for AR(p) and ARMA, ACF tells us little information. We will pursuing a similar function like ACF next.

The **Partial AutoCorrelation Function (PACF)** of a **stationary** time series $\{X_t\}$ is

$$\phi_{11} = Coor(X_{t+1}, X_t) = \rho(1)$$

$$\phi_{hh} = Coor\left(X_{t+h} - \hat{X}_{t+h}, X_t - \hat{X}_t\right) \text{ for } h = 2,3,4,\dots$$

This ϕ_{hh} is the correlation between X_0 and $X_{h,}$ which removes the linear effects of $X_1, \dots, X_{h-1}$

$$\dots, X_{-1}, X_0, X_1, X_2, \dots, X_{h-1}, X_{h,} X_{h+1}, \dots$$

Autocorrelation Function (ACF) of AR(1)

AR(1) model:

$$X_t = \phi_1 X_{t-1} + W_t$$

$$\gamma(1) = Cov(X_0, X_1) = \phi_1 \gamma(0)$$

$$\gamma(2) = Cov(X_0, X_2)$$

$$= Cov(X_0, \phi_1 X_1 + W_2)$$

$$= Cov(X_0, \phi_1^2 X_0 + \phi_1 W_1 + W_2)$$

$$= \phi_1^2 \gamma(0)$$

Clearly, X_0 and X_2 are correlated through X_1.

In the PACF, we remove this dependence by considering the covariance of the prediction errors of X_2^1 and X_0^1

Partial Autocorrelation Function: AR(1)

Calculate the PACF of a causal AR(1) model: $X_t = \phi X_{t-1} + W_t$, with $|\phi| < 1$

$$\phi_{11} = corr(X_1, X_0) = \rho(1) = \phi$$

Suppose regression $\hat{X}_{t+2} = \beta X_{t+1}$

We choose β to minimize

$$E\left((X_{t+2} - \beta X_{t+1})^2\right) = \gamma_X(0) - 2\beta\gamma_X(1) + \beta^2\gamma_X(0)$$

So, $\beta = \phi$. Similarly for the regression $\hat{X}_t = \alpha X_{t+1}$

Then $$\phi_{22} = Cov\left(X_{t+2} - \hat{X}_{t+2}, X_t - \hat{X}_t\right) = 0$$

Partial Autocorrelation Function: AR(p)

In general, for a causal **AR(p)** model

$$X_h = \phi_1 X_{h-1} + \cdots + \phi_p X_{h-p} + W_h$$

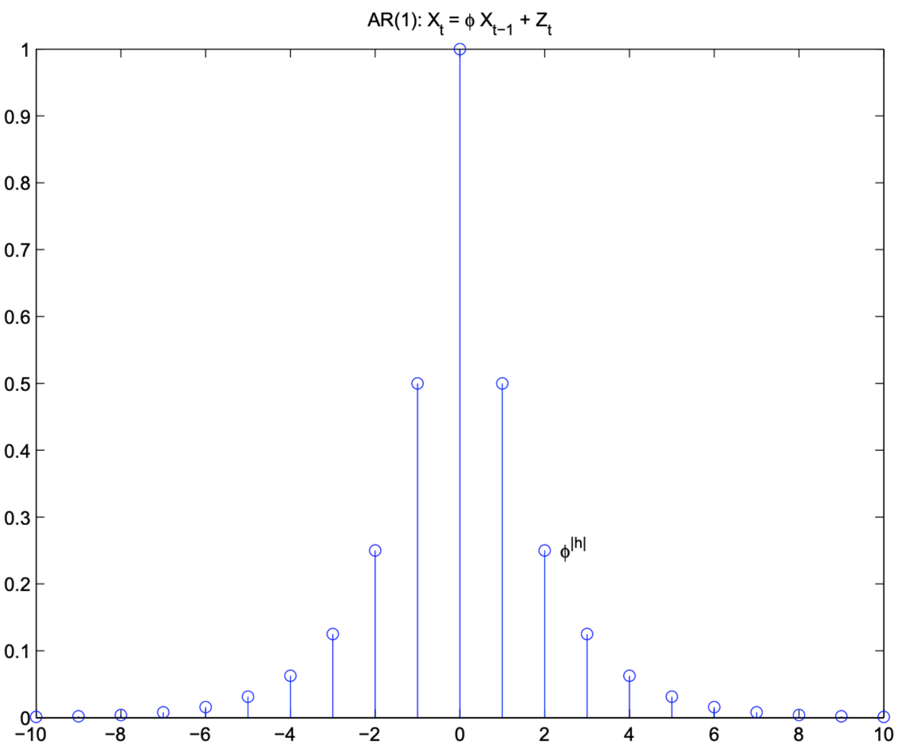
The regression of X_h on $X_{h-1}, \ldots, X_1$ is

$$\hat{X}_h = \phi_1 X_{h-1} + \cdots + \phi_p X_{h-p}$$

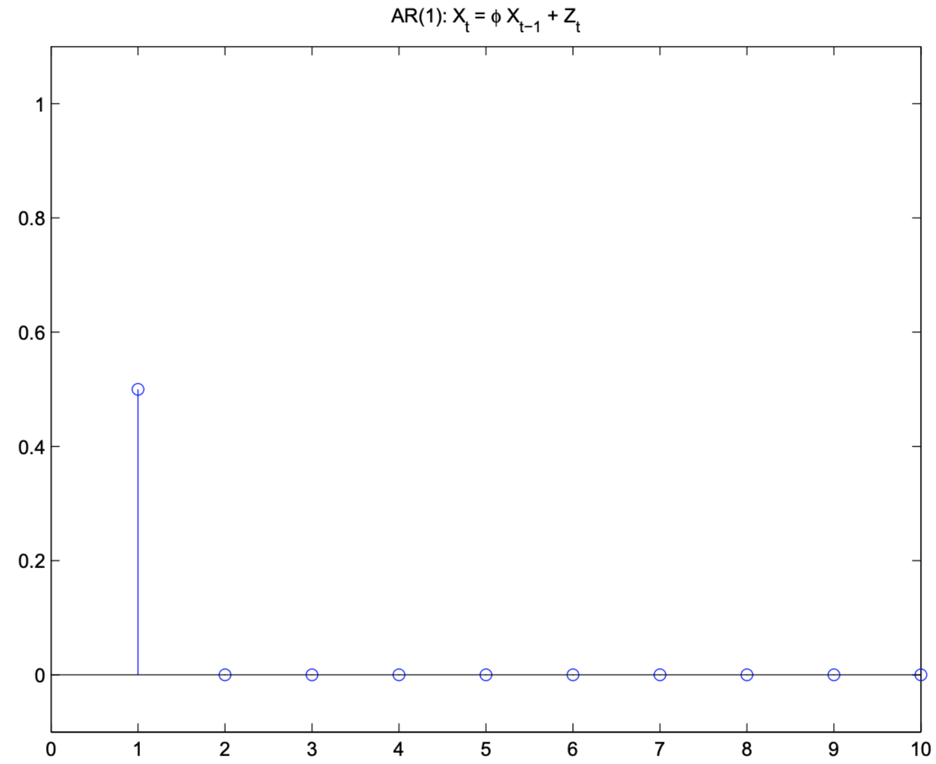
Thus, when $h > p$, by causality,

$$\phi_{hh} = Coor\left(X_h - \hat{X}_h, X_0 - \hat{X}_0\right) = \begin{cases} \phi_h & if\ 1 \le h \le p \\ 0 & otherwise \end{cases}$$

ACF and PACF of AR(1)

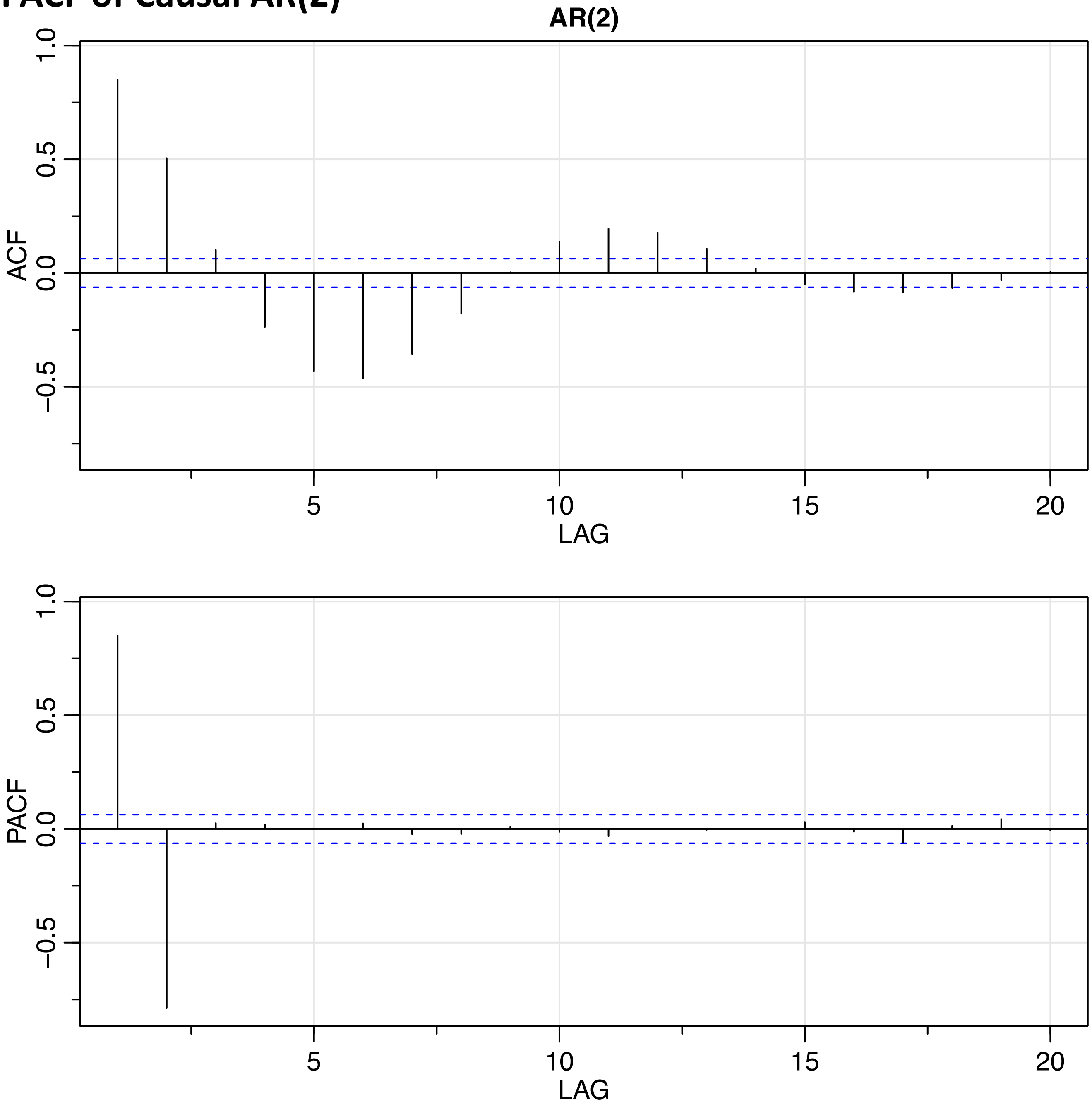


ACF



PACF

ACF and PACF of Causal AR(2)



PACF of an invertible MA(q)

$$X_t = W_t + \theta_1 W_{t-1} + \cdots + \theta_q W_{t-q} = \theta(B) W_t$$

$$X_t = -\sum_{i=1}^{\infty} \pi_i X_{t-i} + W_t$$

$$\hat{X}_{n+1} = P(X_{n+1} \,| X_1, \ldots, X_n)$$

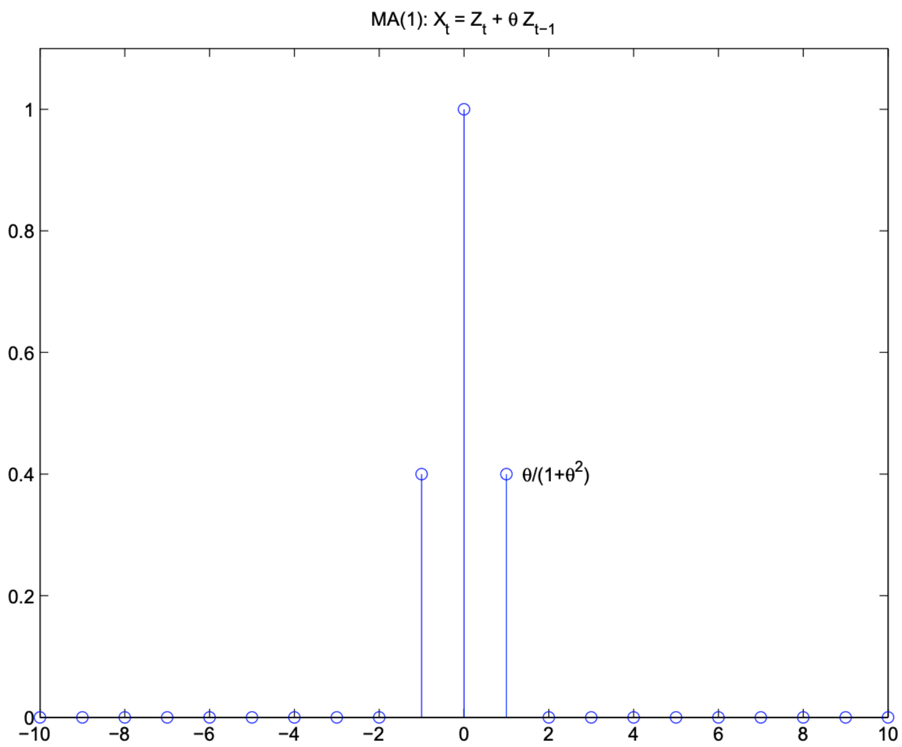
$$= P\left(-\sum_{i=1}^{\infty} \pi_i X_{n+1-i} + W_{n+1} \;|\, X_1, \ldots, X_n\right)$$

$$= -\sum_{i=1}^{\infty} \pi_i \, P(X_{n+1-i} \,|\, X_1, \ldots, X_n)$$

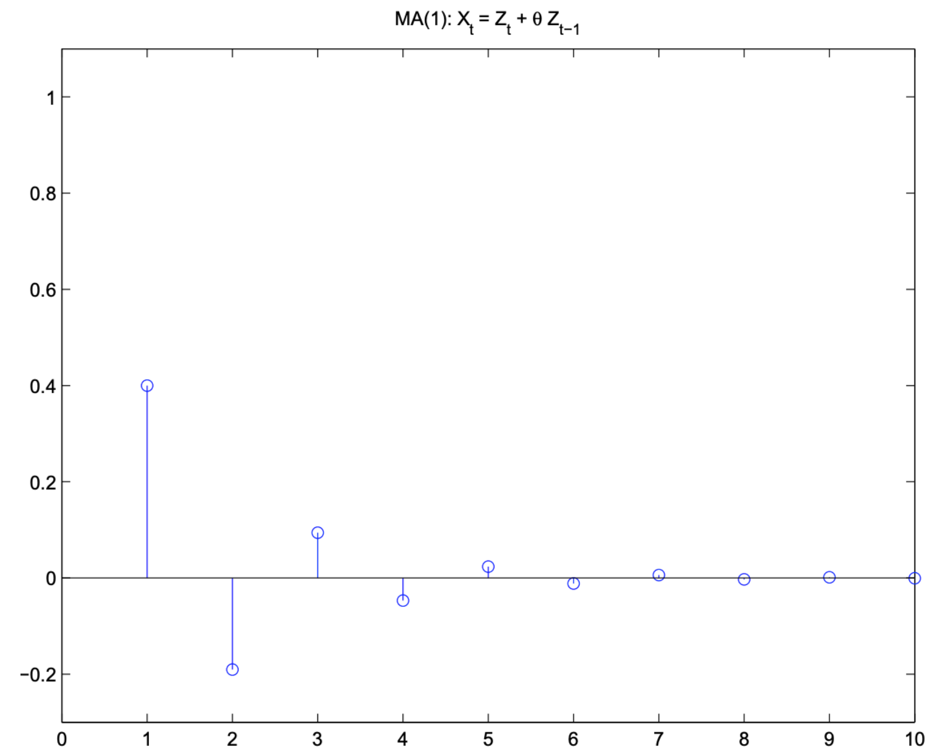
$$= -\sum_{i=1}^{n} \pi_i X_{n+1-i} - \sum_{i=n+1}^{\infty} \pi_i \, P(X_{n+1-i} \,|\, X_1, \ldots, X_n)$$

In general, $\phi_{hh} \neq 0$

ACF and PACF of MA(1)

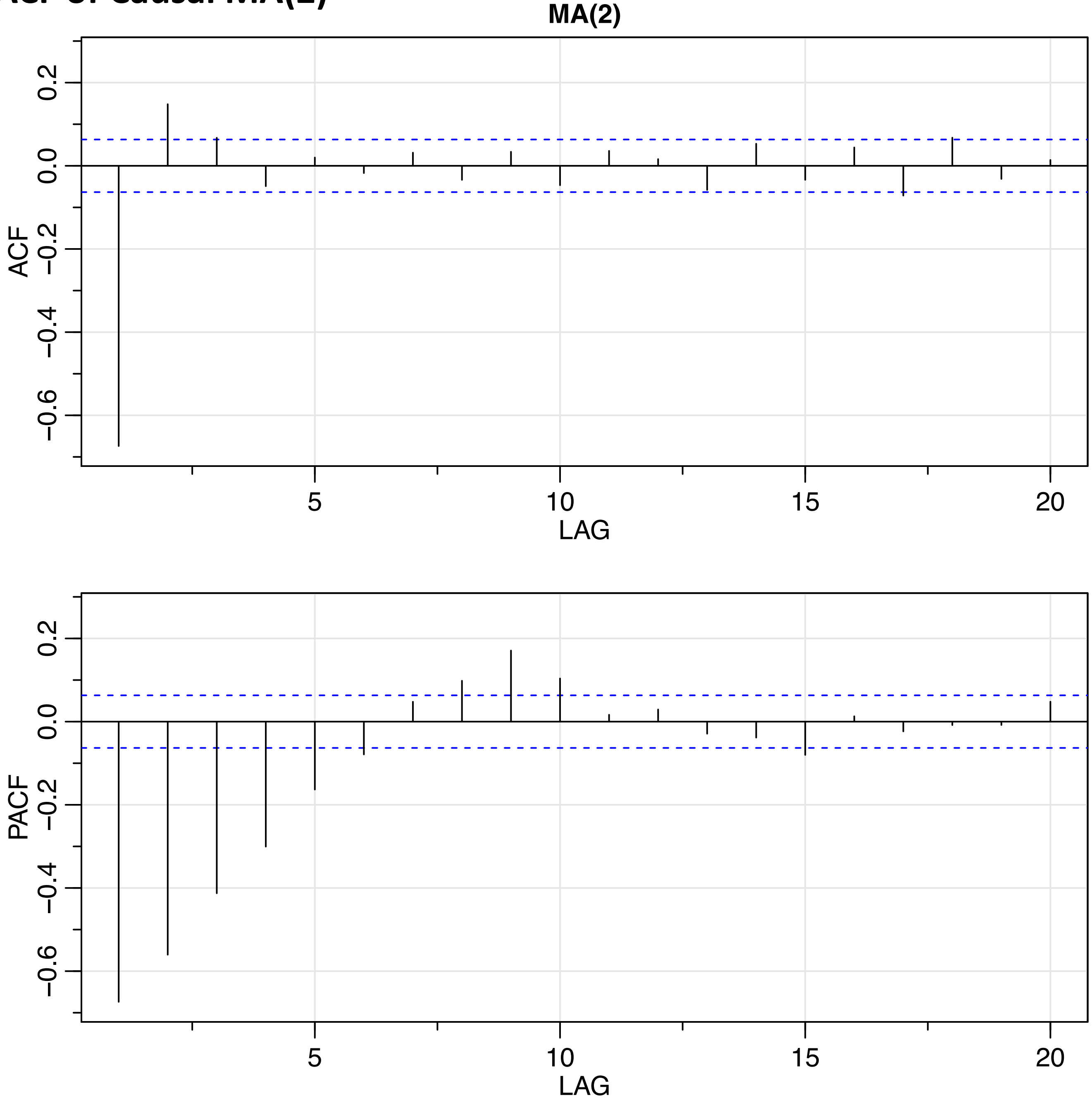


ACF

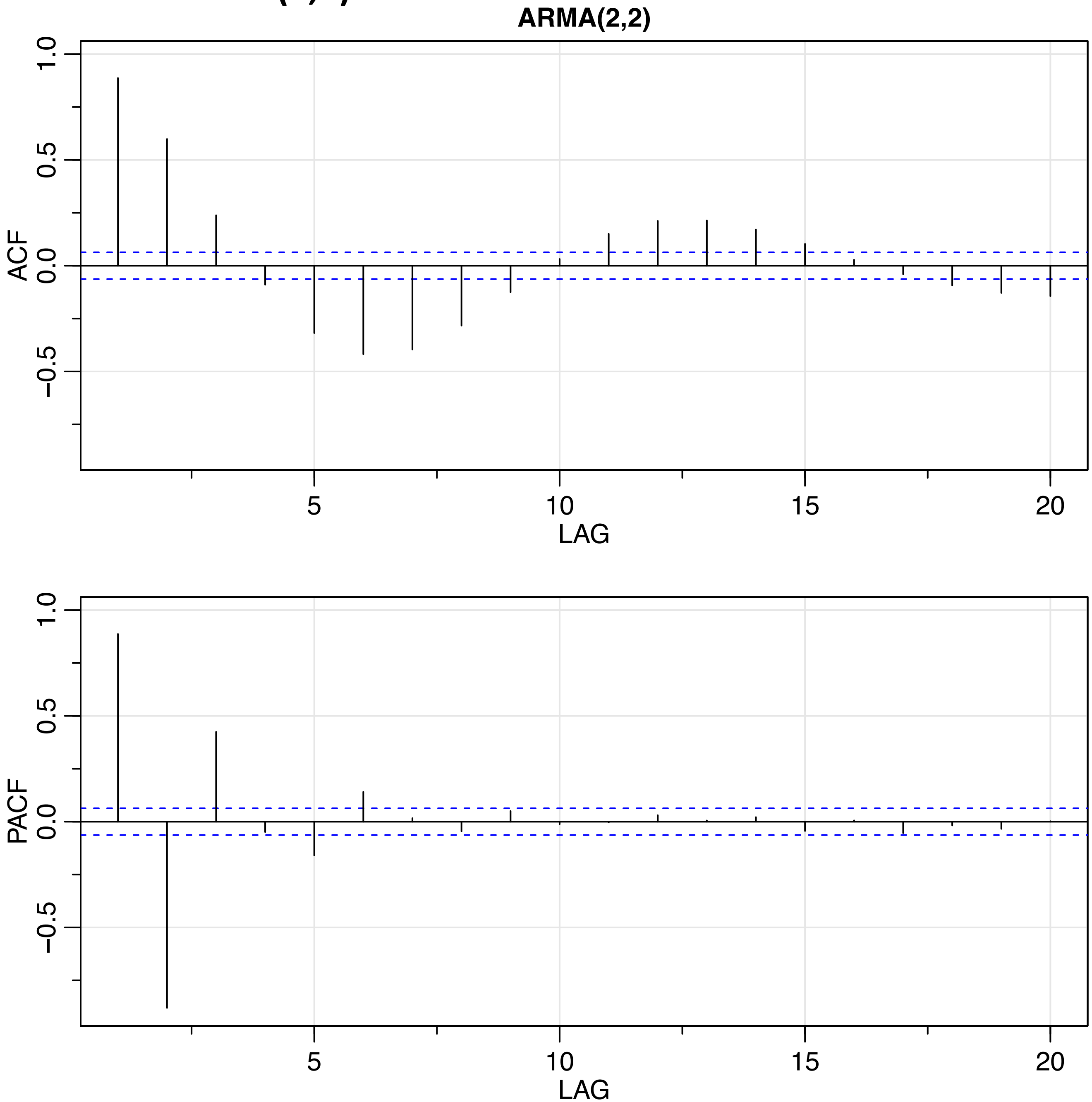


PACF

ACF and PACF of Causal MA(2)



ACF and PACF of Causal ARMA(2,2)



Sample PACF

For a realization $x_1, \ldots, x_n$ of a time series, the **sample PACF** is defined by

$$\hat{\phi}_{00} = 1$$

$$\hat{\phi}_{hh} = last\ component\ of\ \hat{\phi}_h$$

where $\hat{\phi}_h = \hat{\Gamma}_h^{-1}\,\hat{\gamma}_h$

Summary

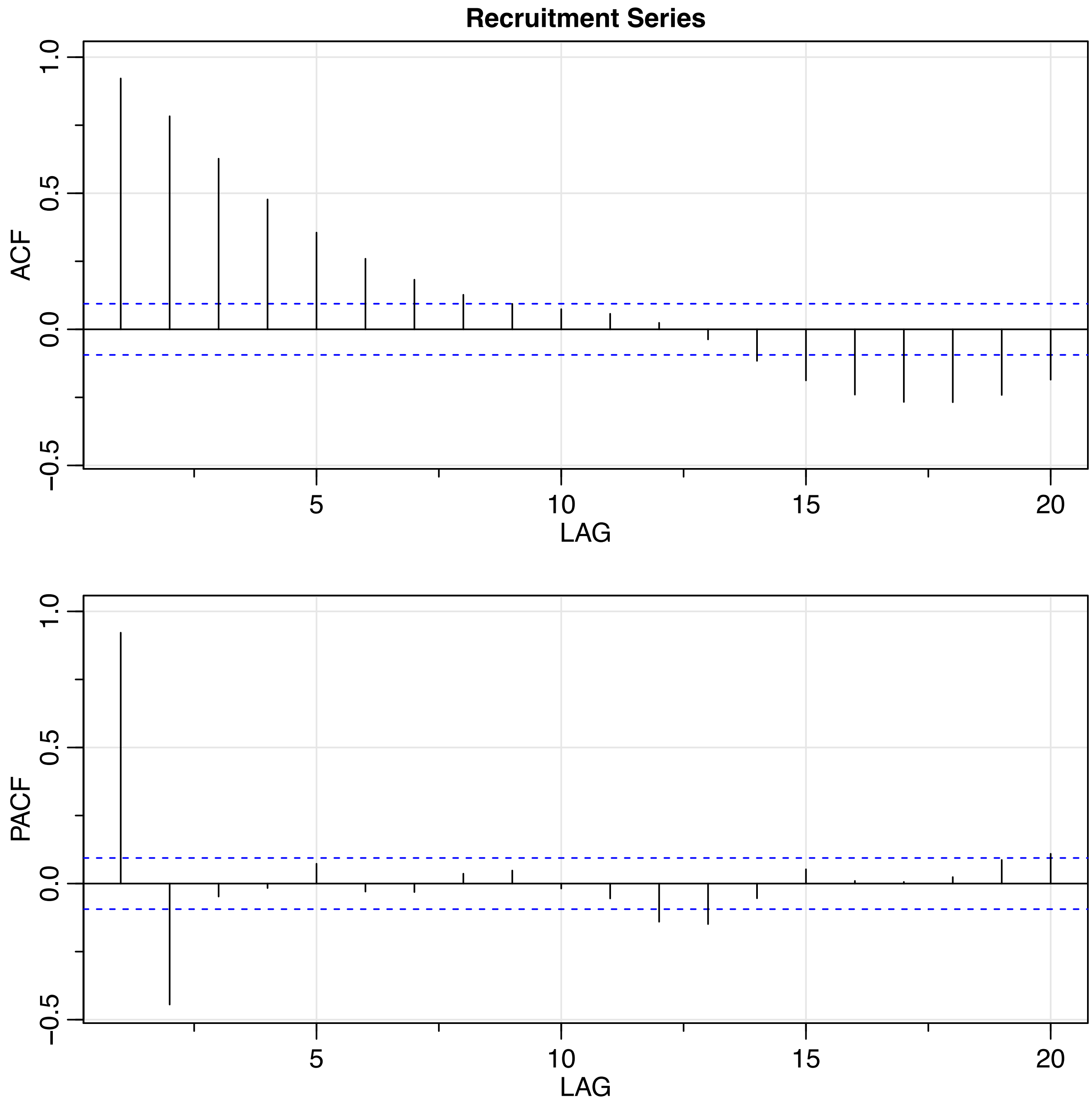
- The ACF of MA(q) model cuts off after lag q. The PACF of an AR(p) model cuts off after lag p.
- Identification of an MA(q) model is best done with ACF; identification of an AR(p) model is best done with PACF.
- The PACF between x_t and x_{t-h} is the correlation between $x_t - \hat{x}_t$ and $x_{t-h} - \hat{x}_{t-h}$. Think of it as taking the correlation between the residuals from two regression models. The dependence on all intermediate variables is removed.

ACF and PACF of Causal AR and Invertible MA

	AR(p)	**MA(q)**	**ARMA(p,q)**
ACF	Decay	0 after lag q	Decay
PACF	0 after lag p	Decay	Decay

Fish Population Example. This time series contains data on fish population in the central Pacific Ocean. The numbers represent the number of new fish in the years 1950-1987.

Question: Based on the ACF and PACF plots, what process do you think is most likely to describe this time series?



Recruitment Series
ACF
1.0
0.5
0.0
−0.5
5
10
15
20
LAG
PACF
1.0
0.5
0.0
−0.5
5
10
15
20
LAG

- "Time Series Analysis and Its Applications", 4th ed. 2017, by Shumway and Stoffer.

  Sections 3.1-3.3

Select ARIMA Model for Time Series:

https://www.mathworks.com/help/econ/box-jenkins-model-selection.html