Queen Mary University of London

Queen Mary School Hainan
Queen Mary University of London

QHP5701 Exploratory Data Analysis

Time Series Models

Nikesh Bajaj, PhD
Lecturer in Data Science,
Queen Mary University of London
nikesh.bajaj@qmul.ac.uk
https://nikeshbajaj.in

#Ref: Chapter 1 to 4, +
Time Series Analysis by William WS Wei

Recap

So far, we have learned:

- Basic concepts of Stochastic Processes, Random Variables
- Time series as Stochastic Process
- Characteristics of time series: Stationarity, Seasonality, Non-stationarity

- Autocovariance Function AVF: γ_k and estimations
- Autocorrelation Function ACF: ρ_k and estimations

- Partial Autocorrelation Function PACF - will be covered after models

Random Walk

Model: $$X_n = X_{n-1} + Z_n$$ $$Z_n \sim N(\mu, \sigma^2)$$

$$X_n = \sum_{i=0}^{n} Z_i \quad + X_0$$

X

$$Z_n \sim N(\mu, \sigma^2)$$

Random Walk

Model:

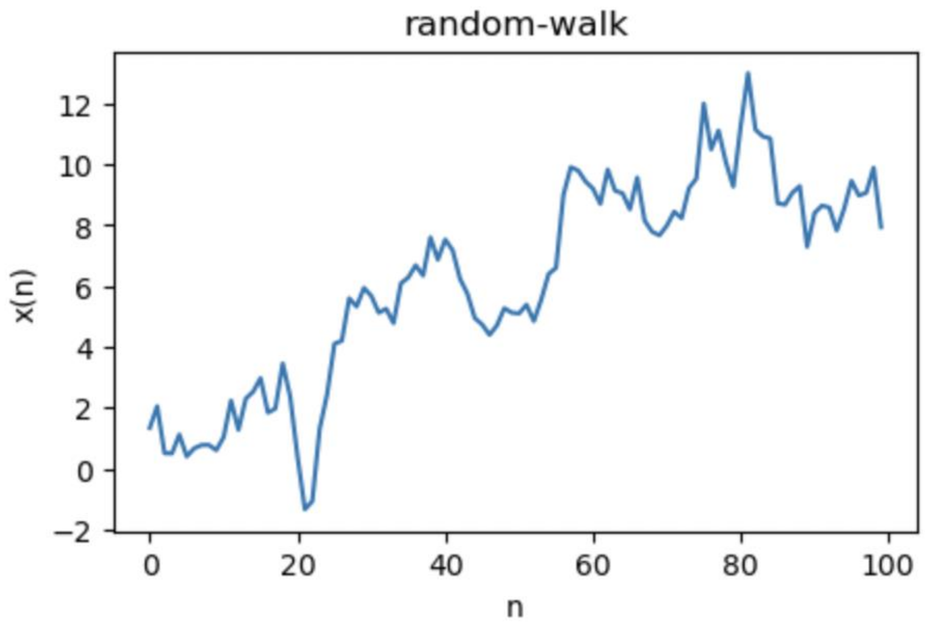
$$X_n = X_{n-1} + Z_n \qquad X_n = \sum_{i=0}^{n} Z_i$$

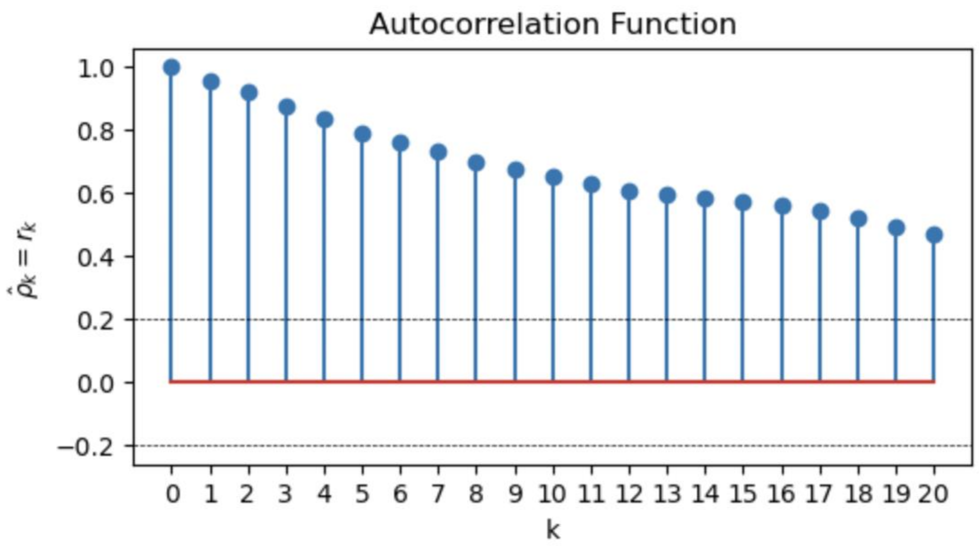
$$E[X_n] = E\left[\sum_{i=0}^{n} Z_i\right] = \sum_{i=0}^{n} E[Z_i] = \mu n$$

$$V[X_n] = V\left[\sum_{i=0}^{n} Z_i\right] = \sum_{i=0}^{n} V[Z_i] = \sigma^2 n$$

Stationary or not??

Random Walk: Simulate and compute AVF and ACF


random-walk
x(n)
n
12
10
8
6
4
2
0
−2
0
20
40
60
80
100


Autocorrelation Function
$\hat{\rho}_k = r_k$
k
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

Difference Operator on random walk

Difference Operator

- Difference Operator $\nabla X = X_n - X_{n-1}$

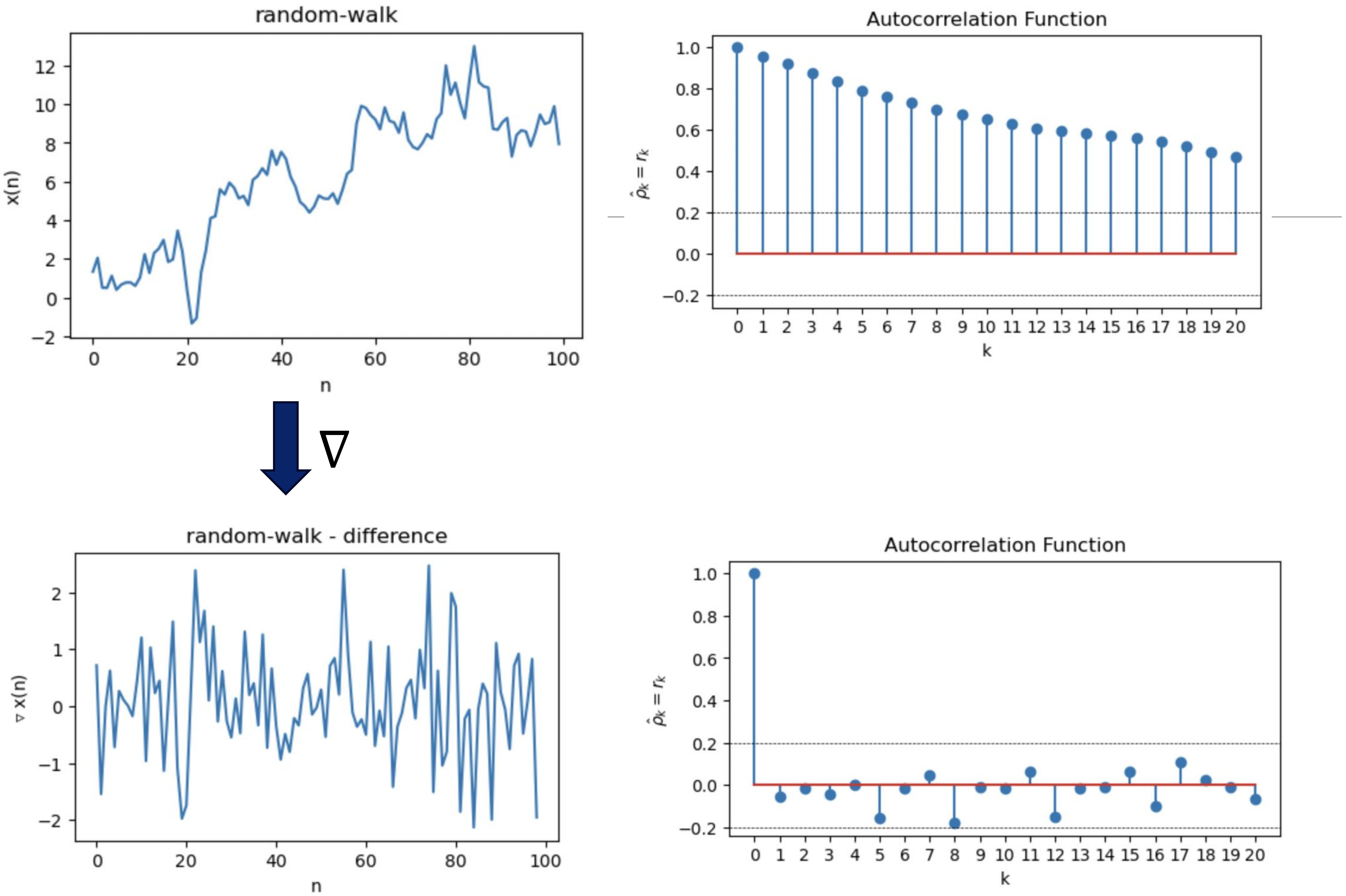
- $X_n = X_{n-1} + Z_n$
- $X_n - X_{n-1} = Z_n = \nabla X$ noise

- Now we can compute AVF and ACF of ∇X

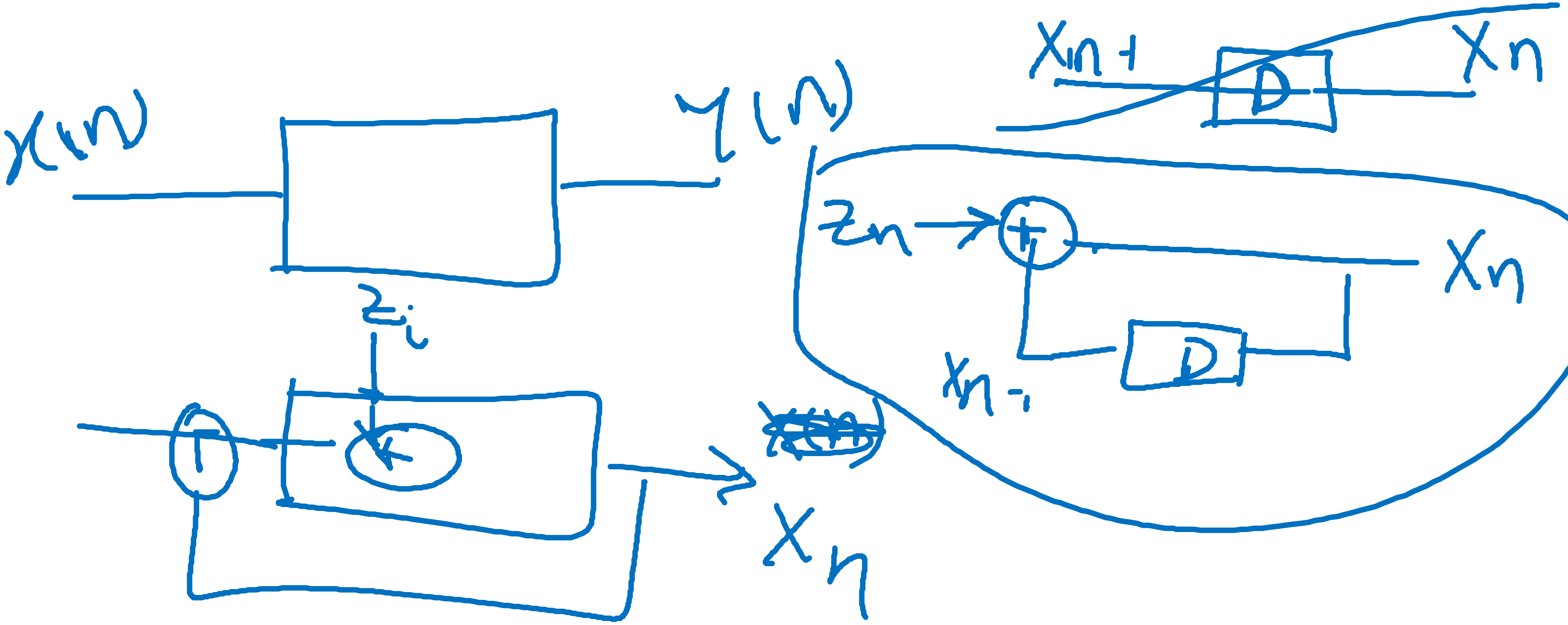
$DX_n = X_n - X_{n-1}$

Stationary or not??

Simulate


random-walk
x(n)
n
Autocorrelation Function
$\hat{\rho}_k = r_k$
k
∇
random-walk - difference
∇ x(n)
n
Autocorrelation Function
$\hat{\rho}_k = r_k$
k

Connection to Signal and Systems



Moving Average processes: MA(q)

Stationary Time Series Models

$Z_n \sim N(\mu, \sigma^2)$ iid noise

Model: order q

$$X_n = Z_n + \theta_1 Z_{n-1} + \theta_2 Z_{n-2} \ldots \theta_q Z_{n-q}$$

MA(1): $X_n = Z_n + \theta_1 Z_{n-1}$

MA(2): $X_n = Z_n + \theta_1 Z_{n-1} + \theta_2 Z_{n-2}$

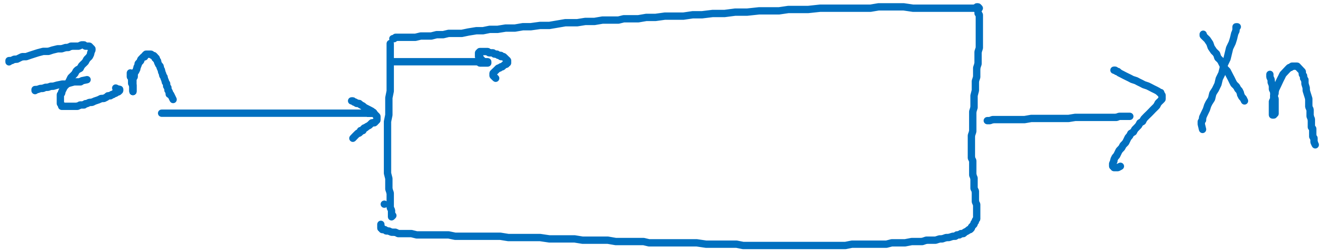
FIR

MA(q)

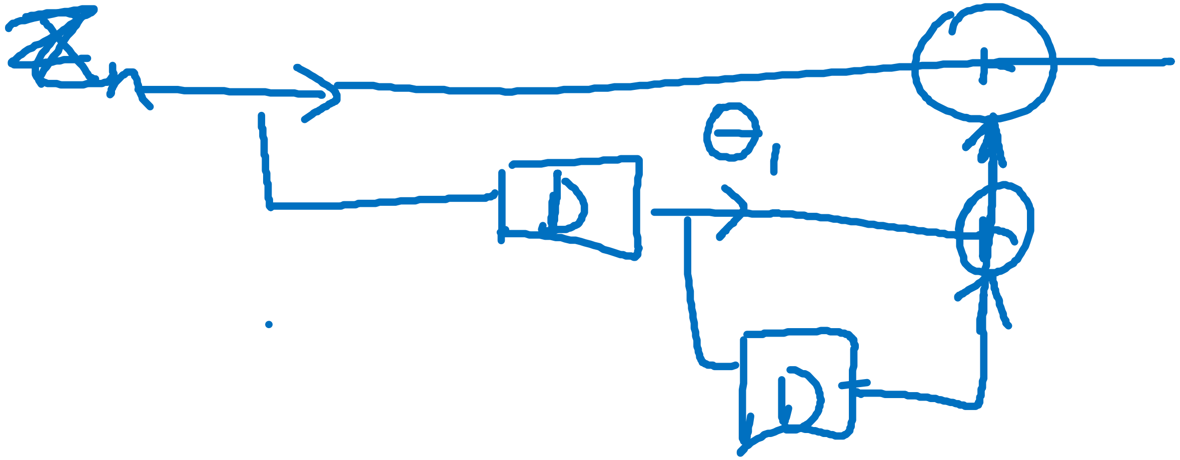
$X_n = Z_n + 0.7 Z_{n-1}$

Example

$$x(n) = z(n) + \theta_1 z(n-1) + \theta_2 z(n-2)$$



$$h(n) = \delta(n) + \theta_1 \delta(n-1) + \theta_2 \delta(n-2)$$



$$h = [1 \ \theta_1 \ \theta_2]$$

Simulate

h = [1 0.7 0.2]

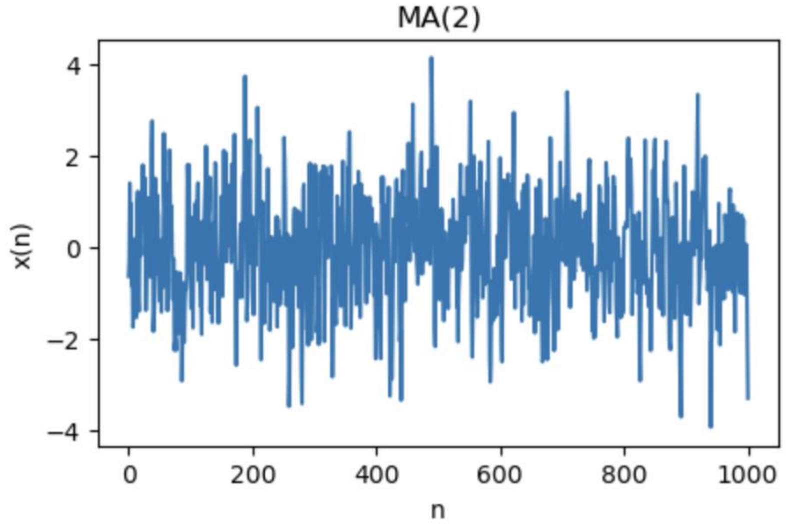
$$MA(2)$$
$$X_n = Z_n + 0.7Z_{n-1} + 0.2Z_{n-2}$$
$$Z_n \sim N(\mu = 0, \sigma^2 = 1)$$
$$Z_n \; are \; iid$$

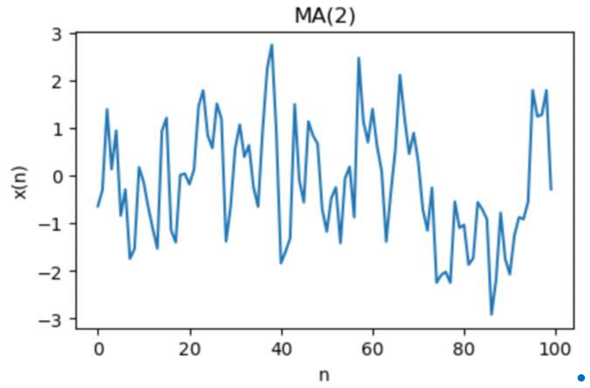
Compute ACF

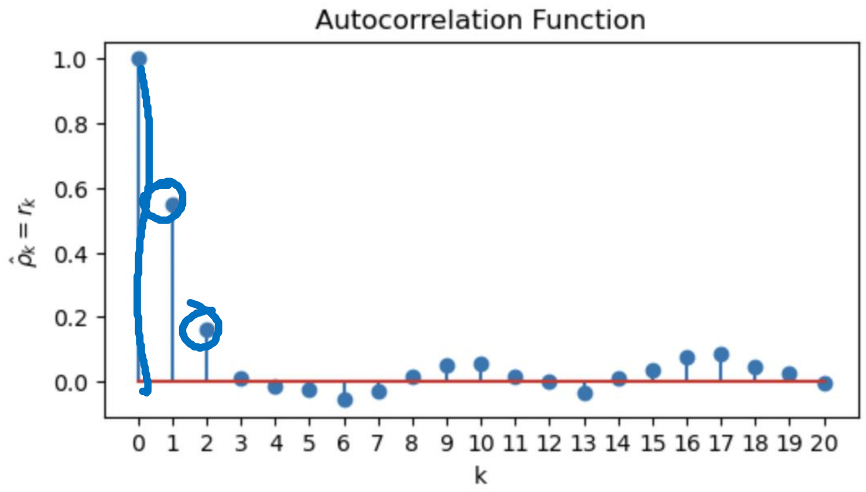
Simulate

$$MA(2)$$
$$X_n = Z_n + 0.7Z_{n-1} + 0.2Z_{n-2}$$
$$Z_n \sim N(\mu = 0, \sigma^2 = 1)$$
$$Z_n \text{ are iid}$$

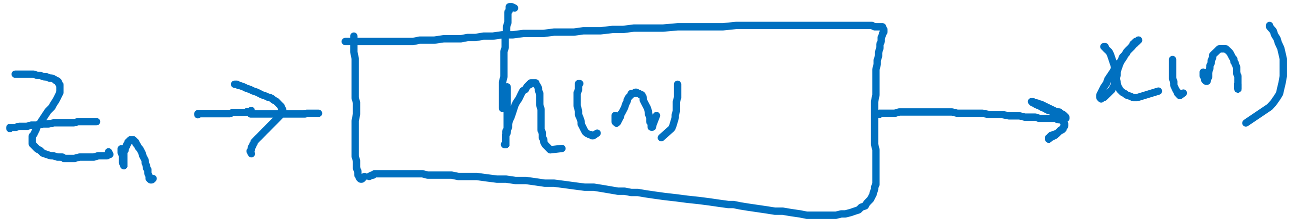
Compute ACF





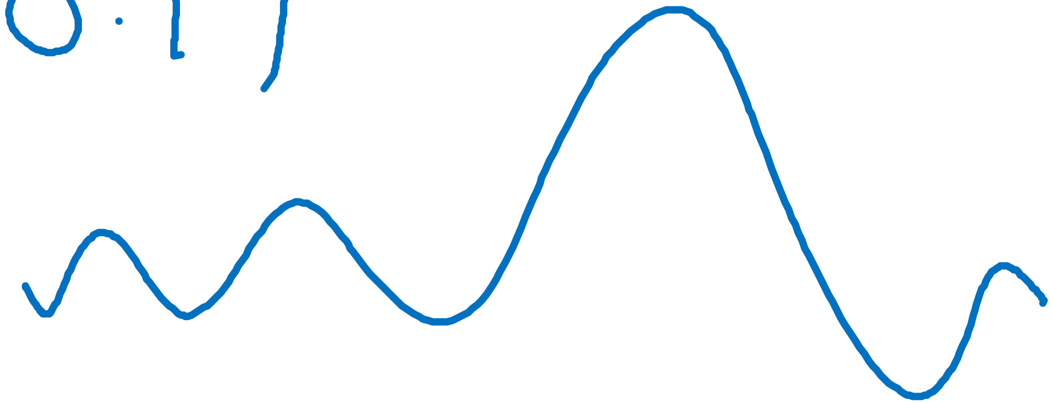


Connection to Signal and Systems



$h(n) = [1 \;\; 0.7 \;\; 0.2 \;\; 0.1]$

FIR — Finite Impulse response



Autoregressive Processes: AR(p)

Stationary Time Series Models

$Z_n \sim N(\mu, \sigma^2)$ iid noise

Model: order p

$$X_n = Z_n + history$$

$$X_n = Z_n + \phi_1 X_{n-1} + \phi_2 X_{n-2} \; \ldots \; \phi_q X_{n-p}$$

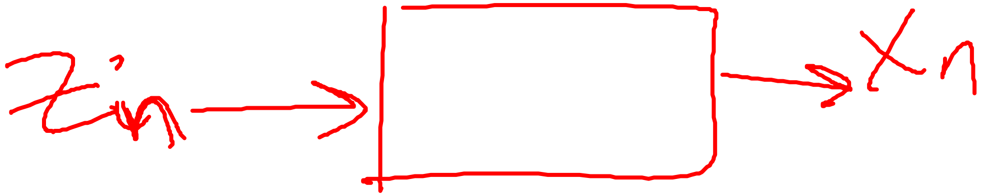
history of x

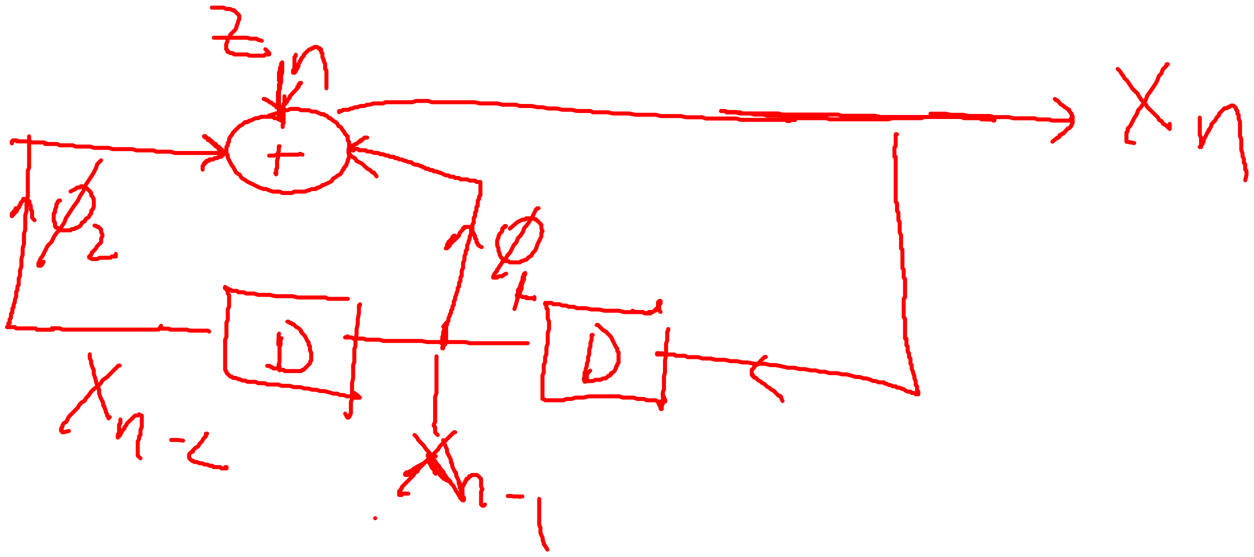
AR(1): $X_n = Z_n + \phi_1 X_{n-1}$ -- random walk for ϕ_1=1

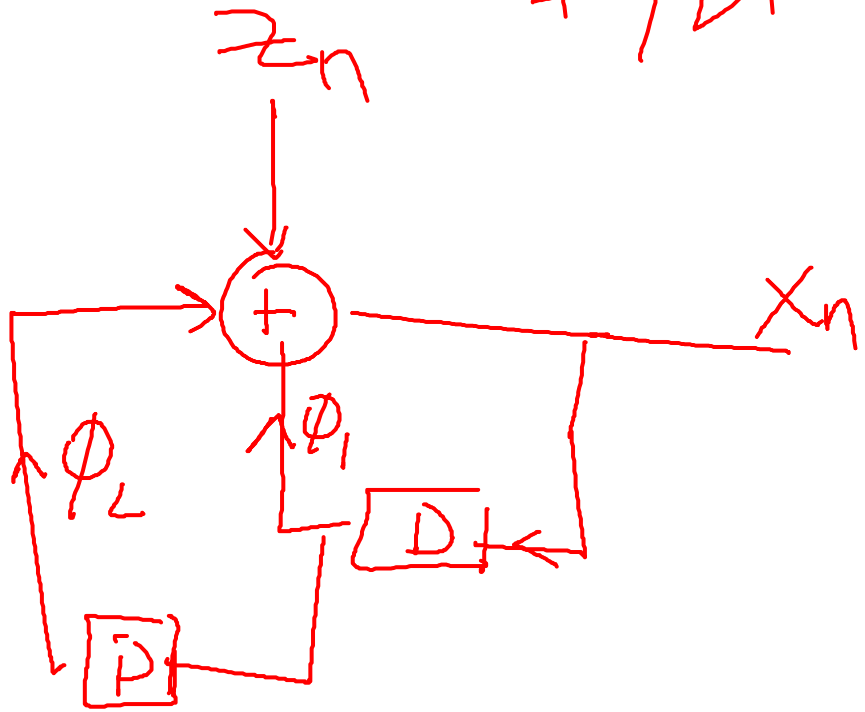
AR(2): $X_n = Z_n + \phi_1 X_{n-1} + \phi_2 X_{n-2}$

Example

AR(2) $X_n = Z_n + \phi_1 X_{n-1} + \phi_2 X_{n-2}$







Simulate

$$AR(2)$$
$$X_n = Z_n + 0.7X_{n-1} + 0.2X_{n-2}$$
$$Z_n \sim N(\mu = 0, \sigma^2 = 1)$$
$$Z_n \text{ are iid}$$

Compute ACF

$$x(n) = z(n) + 0.7x(n-1) + 0.3x(n-2)$$

$$h(n) = \delta(n) + 0.7h(n-1) + 0.3h(n-2)$$

$$h(n) - 0.7h(n-1) + 0.3h(n-2) = \delta(n)$$

$$h(n) - 0.7Dh(n) + 0.3D^2h(n) = \delta(n)$$

Connection to Signal and Systems

$$h_n - 0.7Dh_n - 0.3D^2h_n = \delta_n$$

$$h_n[1 - 0.7D - 0.3D^2] = \delta_n$$

$$h_n = \frac{\delta_n}{1 - 0.7D - 0.3D^2}$$

*D: **Delay** Operator (**B** backshift operator)*

There are different notations for this operator, we will use ***D*** for this operator taking from Signal and Systems lectures

$$DX_n = X_{n-1}$$
$$D^2X_n = X_{n-2}$$
$$D^pX_n = X_{n-p}$$

$$D^{-2}X_n = X_{n+2}$$

AR(1) $X_n = Z_n + 0.7 X_{n-1}$

$$\frac{1}{1-\alpha} = \alpha + \alpha^2 + \alpha^3$$

$$X_n = Z_n + 0.7 D X_n$$

$$X_n - 0.7 D X_n = Z_n$$

$$X_n (1 - 0.7D) = Z_n$$

$$X_n = \frac{1}{1-0.7D} Z_n$$

MA(∞)

$$X_n = (\alpha^1 + \alpha^2 \cdots) Z_n$$

$$X_n = z_n + 0.7 x_{n-1} \quad AR(1)$$

$$MA(\infty)$$

IIR Filters

Infinite Impulse Response

$$X_n = \frac{1}{1 - 0.7D} z_n$$

$$= (\alpha + \alpha^2 - \alpha \ldots) z_n$$

$$= \left(\sum_{k=0}^{\infty} \alpha^k\right) z_n$$

$$\alpha = 0.7D$$

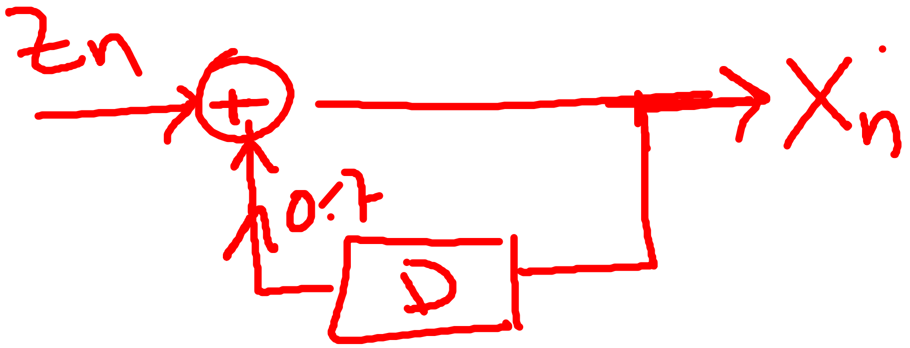
$$h(n) = \sum_{k=-\infty}^{\infty} a_k \delta(n-k)$$

$$X_n = \ldots \alpha^{100} z_n \ldots$$

$$(0.7D)^{100} z_n$$

$$z_{n-100}$$

$$X_n = Z_n + 0.7\,X_{n-1}$$



$X_n = 100 \longrightarrow X_0$

AR and MA

AR: $X_n = Z_n + \phi_1 X_{n-1} + \phi_2 X_{n-2} \ \dots \ \phi_q X_{n-p}$

MA: $X_n = Z_n + \theta_1 Z_{n-1} + \theta_2 Z_{n-2} \ \dots \ \theta_q Z_{n-q}$

$$X_n = Z_n + \phi_1 D X_n + \phi_2 D^2 X_n \ \dots \ \phi_q D^p X_n$$

$$X_n(1 - (\phi_1 D + \phi_2 D^2 \ \dots + \phi_q D^p)) = Z_n$$

$$X_n \phi(D) = Z_n$$

$$X_n = \frac{1}{\phi(D)} Z_n$$

$$X_n = \frac{1}{1-(\phi_1 D + \phi_2 D^2 \ \dots \ + \phi_q D^p)} Z_n$$

AR and MA Processes

MA(q): $X_n = Z_n + \theta_1 Z_{n-1} + \theta_2 Z_{n-2} \;\ldots\; \theta_q Z_{n-q}$

$$X_n = \theta^q(D) Z_n \qquad \theta^q(D) = (1 + \theta_1 D + \theta_2 D^2 + \ldots + \theta_q D^q)$$

AR(p) : $X_n = Z_n + \phi_1 X_{n-1} + \phi_2 X_{n-2} \;\ldots\; \phi_q X_{n-p}$

$$Z_n = \phi^p(D) X_n \qquad \phi^p(D) = 1 - (\phi_1 D + \phi_2 D^2 + \ldots + \phi_p D^p)$$

Duality & Invertibility

$MA(q) \Rightarrow AR(\infty)$: *invertibility condition*
$AR(q) \Rightarrow MA(\infty)$: *stationarity condition*

$MA(q)$: $X_n = \theta^q(D)Z_n$

$$X_n = \theta^q(D)Z_n$$
$$\frac{1}{\theta^q(D)} X_n = Z_n$$

$AR(p)$: $Z_n = \phi^p(D)X_n$

$$\frac{1}{\phi^p(D)} Z_n = X_n$$

AR(1) Process, Stationary?

$$X_n = Z_n + \phi_1 X_{n-1}$$

AR(1) Process, Stationary?

$$X_n = Z_n + \phi X_{n-1}$$

$$X_n - \phi X_{n-1} = Z_n$$
$$X_n - \phi D X_n = Z_n$$

$$\phi(D) X_n = Z_n$$

$$X_n = \frac{1}{1 - \phi D} Z_n$$

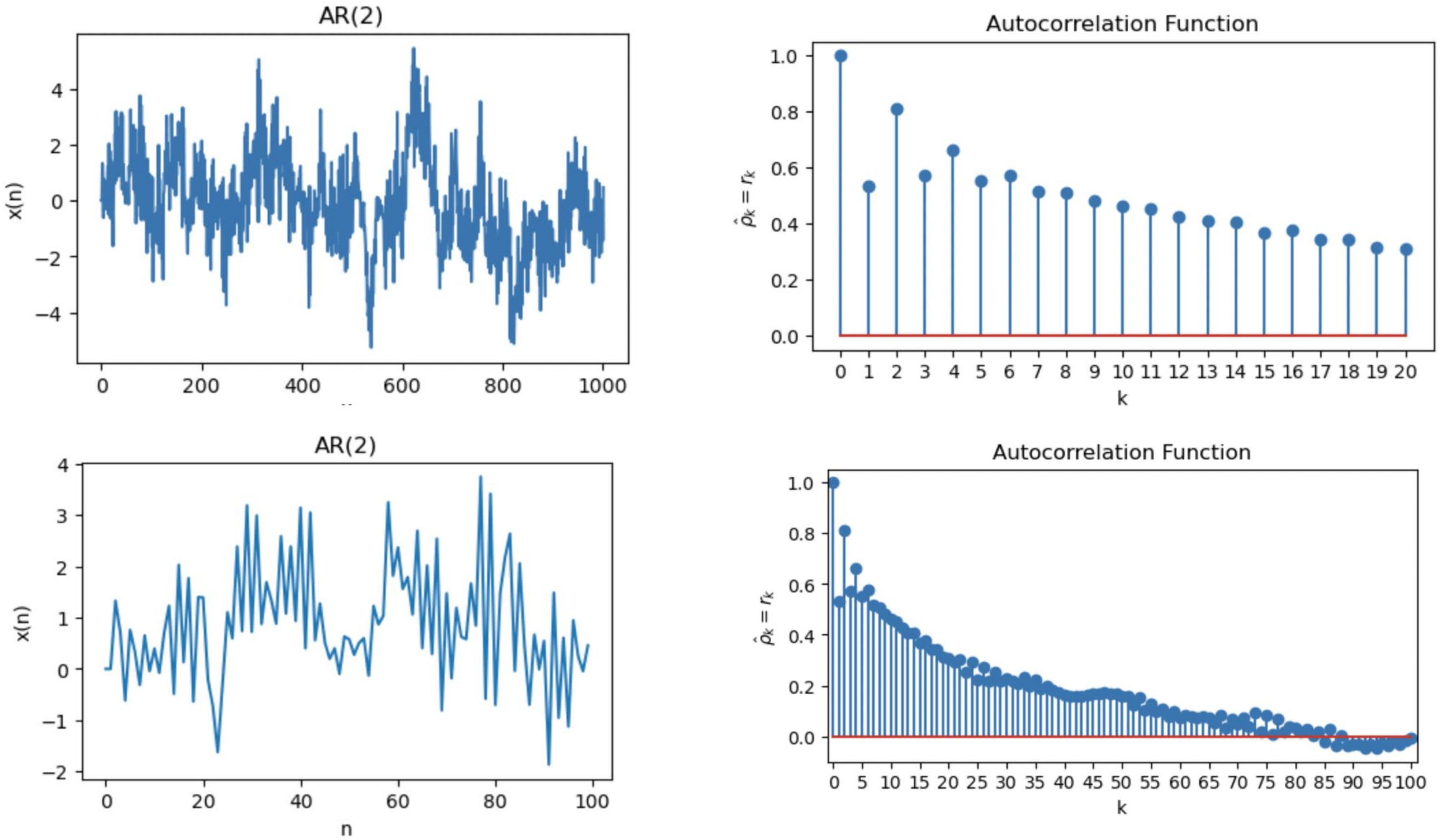
$$X_n = [1 + \phi D + \phi^2 D^2 + \ \ldots] Z_n$$

$$E[X_n] = 0$$
$$Var[X_n] = \sigma^2 \sum_i \phi^i$$

Example

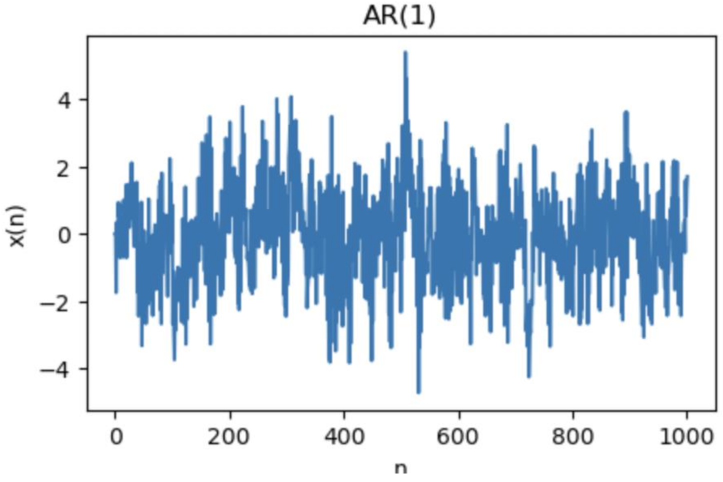
Simulate

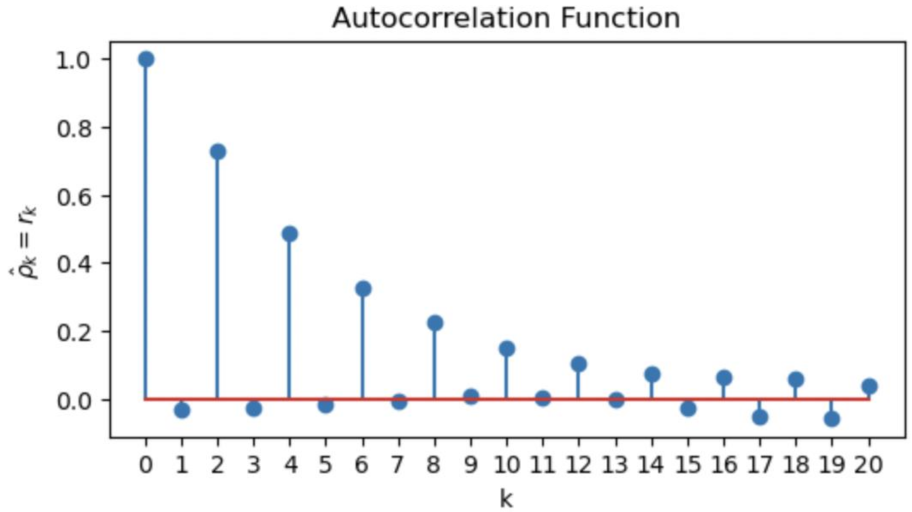
$$AR(2)$$
$$X_n = Z_n + 0.7X_{n-1} + 0.2X_{n-2}$$
$$Z_n \sim N(\mu = 0, \sigma^2 = 1)$$
$$Z_n \ are \ iid$$

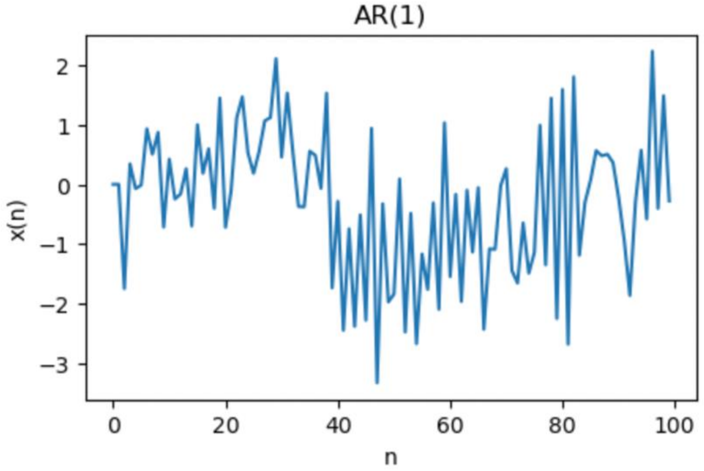


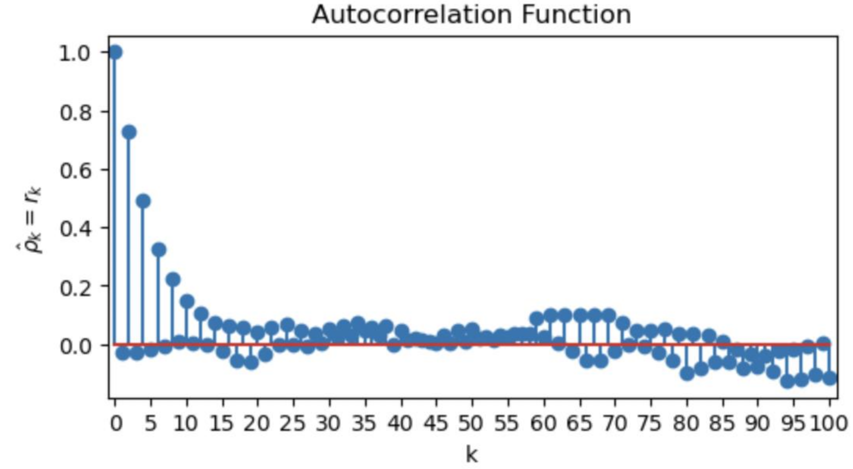
Simulate

$$AR(1)$$
$$X_n = Z_n + 0.7X_{n-1}$$
$$Z_n \sim N(\mu = 0, \sigma^2 = 1)$$
$$Z_n \; are \; iid$$









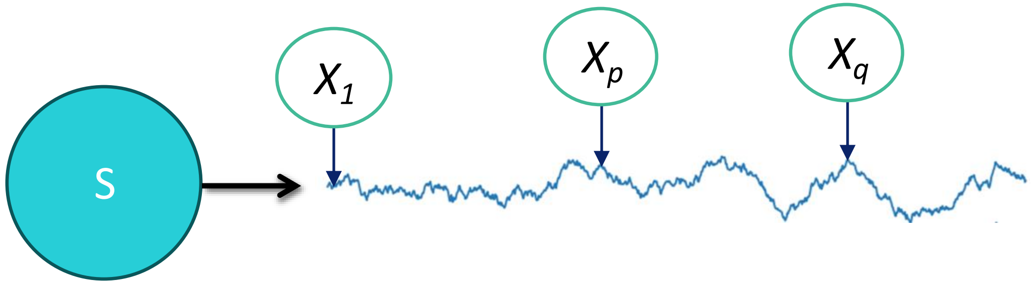
Summary: Problem

For MA(q) processes, ACF reveals the order of model - FIR

For AR(p) processes, ACF shows infinite order - IIR

Partial Autocorrelation Function: PACF

- With models like AR, there are more random variables that are highly correlated to each other.
  - For example, X_p might be highly correlated to other X_{p+1} and X_{p+5} too
- Computing correlation of X_p and X_q after removing the correlations from other variables is called **conditional correlation** or **partial correlation**



$$Corr(X_p, X_q \mid X_{p+1}, X_{p+2} \ldots X_{p+q-1})$$

$$P_k = Corr(X_p, X_{p+k} \mid X_{p+1}, X_{p+2} \ldots X_{p+k-1})$$

$$P_k = Corr(X_p, X_{p+k} \mid X_{p+1}, X_{p+2} \dots X_{p+k-1})$$

Partial Autocorrelation Function: PACF

Partial autocorrelation is defined as:

$$P_k = \frac{Cov[(X_p - \hat{X}_p)(X_{p+k} - \hat{X}_{p+k})]}{\sqrt{Var(X_p - \hat{X}_p)}\sqrt{Var(X_{p+k} - \hat{X}_{p+k})}}$$

Where

$$\hat{X}_p = \beta_1 X_{p+1} + \beta_2 X_{p+2} + \dots + \beta_{k-1} X_{p+k-1}$$

Best estimation with Linear regression fit

$$P_k = Corr(X_p, X_{p+k} \mid X_{p+1}, X_{p+2} \dots X_{p+k-1})$$

Partial Autocorrelation Function: PACF

Partial autocorrelation is defined as:

$$P_k = \frac{Cov[(X_p - \hat{X}_p)(X_{p+k} - \hat{X}_{p+k})]}{\sqrt{Var(X_p - \hat{X}_p)}\sqrt{Var(X_{p+k} - \hat{X}_{p+k})}}$$

In terms of linear regression

$$X_{p+k} = \phi_{k1}X_{p+k-1} + \phi_{k2}X_{p+k-2} + \cdots + \phi_{kk}X_p + e_{p+k}$$

Estimation Partial Autocorrelation Function: PACF

Partial Autocorrelation can be estimated using recursive equations

$$\hat{\phi}_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^{k} \hat{\phi}_{kj}\hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^{k} \hat{\phi}_{kj}\hat{\rho}_{kj}}$$

$$\hat{\phi}_{k+1,j} = \hat{\phi}_{kj} - \hat{\phi}_{k+1,k+1}\hat{\phi}_{k,k+1-j} \qquad for\ j = 1 \ \dots k$$

Example: ACF, PACF

$$X = [13, 8, 15, 4, 4, 12, 11, 7, 14, 12]$$

Solution

$\hat{\rho}_1 = -0.188,$
$\hat{\rho}_2 = -0.201,$
$\hat{\rho}_3 = 0.181,$

$\hat{\phi}_{11} = \hat{\rho}_1 = -0.188,$
$\hat{\phi}_{22} = -0.245$
$\hat{\phi}_{21} = -0.234,$
$\hat{\phi}_{33} = 0.097,$

t	Z_t	Z_{t+1}	Z_{t+2}	Z_{t+3}	$\cdots;$	Z_{t-1}	Z_{t-2}
1	13	8	15	4			
2	8	15	4	4		13	
3	15	4	4	12		8	13
4	4	4	12	11		15	8
5	4	12	11	7		4	15
6	12	11	7	14		4	4
7	11	7	14	12		12	4
8	7	14	12			11	12
9	14	12				7	11
10	12					14	7

Partial Autocorrelation Function: PACF

Questions for practice

ACF: 2.5.3,

Chapter 2:

Example 2.4

Example 2.5

Checked solved examples

Chapter 3:

Checked solved examples

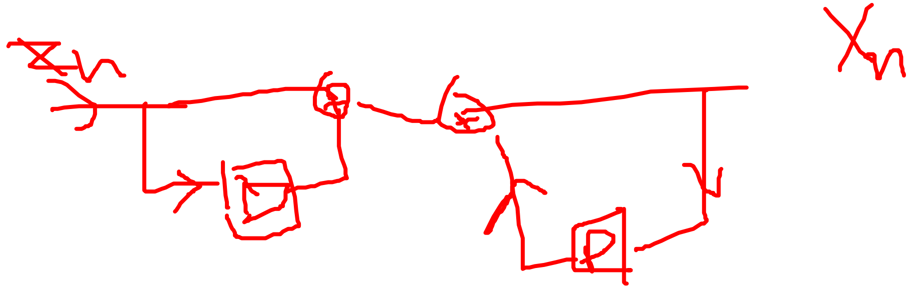
Autoregressive Moving Average: ARMA(p,q)

A generalized for for AR, MA or combination

$MA(q)\colon\ X_n = \theta^q(D)Z_n$

$AR(p)\colon Z_n = \phi^p(D)X_n$

$ARMA(p,q)\colon \phi^p(D)X_n = \ \theta^q(D)Z_n$



$ARMA(p,q)$

$ARMA(1,1)$

$$(1+\phi_1 D)X_n = (1+\theta_1 D)Z_n$$

$$(1+0.3D)X_n = (1+0.7D)Z_n$$

$$X_n + 0.3X_{n-1} = Z_n + 0.7Z_{n-1}$$

Example

Autoregressive Integrated Moving Average: ARIMA(p,d,q)

$$ARIMA(p,d,q): \phi^p(D)(1-D)^d X_n = \theta^q(D)Z_n$$

$$ARIMA(0,1,1) \;\; or \;\; IMA(1,1)$$

$$(1-D)X_n = (1-\theta D)Z_n$$

$$(1-D)X_n = (1-0.7D)Z_n$$

$$X_n - X_{n-1} = Z_n - 0.7Z_{n-1}$$

$$X_n = X_{n+1} + Z_n - 0.7Z_{n-1}$$

Identifying & Fitting a model

- Given a time series, first step is to analyse it to identify the possible stochastic model (processes) it comes from.
  - Is it a MA, AR, ARMA or ARIMA processes?
- For Identifying potential model, use ACF, PACF.
- Once a potential model is identified, We fit a model on time series to estimate the coefficients of model θ_i and ϕ_i.
  - Linear Regression is one approach
  - There are python libraries that can do that
  - The fundamentals of fitting a model and utilizing it will be covered Principles of Machine Learning Module.
- Once a model is fit, we test the quality of fitness using some criteria

Quality Criteria: Self-reading Section

- Sum of Squared Error
- R square Error
- Akaike Information Criterion
- Bayesian Information Criterion

Quality Criteria: Self-reading Section

- Sum of Squared Error
- R square Error

AIC: Self-reading Section

The AIC tries to help you assess the relative quality of several competing models, just like adjusted R square in linear regression, by giving credit for models which reduce the error sum of squares and at the same time by building in a penalty for models which bring in too many parameters

$$AIC = \log(\hat{\sigma}^2) + \frac{n + 2p}{n}$$

$$\hat{\sigma}^2 = \frac{SSE}{n}$$

p-terms in the model

n-data points

Alternative definition for AIC

$$AIC = -2\log(ML) + 2K$$

$ML = \ maximum \ - likelihood$
$K = \ \ number\ of\ parameters\ in\ model$

BIC: Self-reading Section

$$BIC = \mathrm{K}\log(n) - 2\log(L(\widehat{\mathrm{w}}))$$

$L(\widehat{\mathrm{w}}) = \ maximum \ - likelihood$
$K \ = \ number\ of\ parameters\ in\ model$
$n = sample\ size$

Advices for TSA

- Go through the slides
- All the contents of the slides are relevant for the module and exam
- Jupyter-Notebooks are to help you understand the computations
- Be very careful with natation - different books/lectures use different notations
- Go through the Book examples and details
- Read the theory about **quality criteria**

Queen Mary
University of London