Practical Time Series Analysis – Course Notes

Yuchen Wang

July 17, 2019

Contents

1 Week 1 - Review

Introduction One usually thinks of a time-series as a realization derived from a mathematical object called the stochastic process.

Stationarity $\begin{cases} \text{Strong} \\ \text{Weak} \end{cases}$
We want strong stationarity, but we often find real datasets not exhibiting strong stationarity. But they do exhibit weak stationarity, or at least approximately so.

R histogram code

```
hist(small.size.dataset, xlab='My data points', main='Histogram of
    my data', freq=F, col='green', breaks=10)
lines(density(small.size.dataset), col='red', lwd=5)
```

Calculations: Formulas

1. $SSX = \sum(x_i - \bar{x})(x_i - \bar{x}) = \sum x_i^2 - \frac{1}{n}\sum x_i \sum x_i$
2. $SSY = \sum(y_i - \bar{y})(y_i - \bar{y}) = \sum y_i^2 - \frac{1}{n}\sum y_i \sum y_i$
3. $SSXY = \sum(x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{1}{n}\sum x_i \sum y_i$

$$\begin{aligned}
Cov(x,y) &= \frac{1}{n-1}\sum(\frac{x_i - \bar{x}}{S_x})(\frac{y_i - \bar{y}}{S_y}) \\
&= \frac{1}{n-1}\sum(\frac{x_i - \bar{x}}{\sqrt{\frac{SSX}{n-1}}})(\frac{y_i - \bar{y}}{\sqrt{\frac{SSY}{n-1}}}) \\
&= \sum(\frac{x_i - \bar{x}}{\sqrt{SSX}})(\frac{y_i - \bar{y}}{\sqrt{SSY}}) \\
&= \frac{1}{\sqrt{SSXSSY}}\sum(x_i - \bar{x})(y_i - \bar{y}) \\
&= \frac{SSXY}{\sqrt{SSX}\sqrt{SSY}}
\end{aligned}$$

2 Week 2 - Visualizing Time Series, and Beginning to Model Time Series

Definition Time series is a data set collected through time.

Correlation Sampling adjacent points in time introduce a correlation. (classical statistical inference might not work in this setting.)

First intuitions on (Weak) Stationarity

Intuitions

1. No systematic change in mean (no trend)
2. No systematic change in variation
3. No periodic fluctuations
4. The properties of one section of a data are much like the properties of the other sections of the data
5. For an non-stationary time series, we will do some transformations to get stationary time series

2.1 Autocovariance function

Definition of Random Variables A random variable is a function that goes from sample space to real numbers:

$$X : S \to \mathbb{R}$$

Remarks Sample space contains all possible outcomes of the experiment, and if we map each possible outcome of the experiment to a real number, we get a random variable.

Stochastic Processes A collection of a random variables

$$X_1, X_2, X_3, \dots$$

Each one of these random variables might have their own distribution

Remarks Opposite to deterministic process (e.g. when solving an ODE you know exactly what is next), at every step we have some randomness

a Second Definition of Time Series A time series is a realization of a stochastic process

Autocovariance function

$$\gamma(s,t) = Cov(X_s, X_t) = E[(X_s - \mu_s)(X_t - \mu_t)]$$

$$\gamma_k = \gamma(t, t+k) \approx c_k$$

c_k: autocovariance coefficient, approximates γ_k

Remarks Only depends on time difference (k). We assume we are working with stationary time series.

Estimation of the covariance We have a paired dataset

$$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$$

Estimation of covariance (cov() in R)

$$s_{xy} = \frac{\sum_{t=1}^{n}(x_t - \bar{x})(y_t - \bar{y})}{n-1}$$

Autocovariance coefficient

$$c_k = \frac{\sum_{t=1}^{n-k}(x_t - \bar{x})(x_{t+k} - \bar{x})}{n}$$

where

$$\bar{x} = \frac{\sum_{t=1}^{n} x_t}{n}$$

R code

```
(acf(time_series, type='covariance'))
```

The parentheses also gives a table.

2.2 Autocorrelation Function (ACF)

Assumptions

1. We assume weak stationarity
2. The autocorrelation coefficient between x_t and x_{t+k} is defined to be
$$-1 \leq \rho_k = \frac{\gamma_k}{\gamma_0} \leq 1$$
3. Estimation of autocorrelation coefficient at lag k
$$r_k = \frac{c_k}{c_0} = \frac{\sum_{t=1}^{N-K}(x_t - \bar{x})(x_{t+k} - \bar{x})}{\sum_{t=1}^{N}(x_t - \bar{x})^2}$$

R code

```
(acf(time_series))
```

no specification for type.
It plots autocorrelation coefficients at different lags: Correlogram, and it always starts at 1 since $r_0 = \frac{c_0}{c_0} = 1$. Dotted lines are showing the significance level.

2.3 Random Walk

Model

$$X_t = X_{t-1} + Z_t$$

where $Z_t \sim Normal(\mu, \sigma^2)$

$$X_t = \sum_{i=1}^{t} Z_i$$

You accumulate noises.

Inference

$$E[X_t] = E[\sum_{i=1}^{t} Z_i] = \sum_{i=1}^{t} E[Z_i] = \mu t$$
$$Var[X_t] = Var[\sum_{i=1}^{t} Z_i] = \sum_{i=1}^{t} Var[Z_i] = \sigma^2 t$$

assume Z_i's are independent

Remarks There is high auto-correlation and no stationarity.

Difference diff() to remove the trend, get stationary time series from a random walk.

2.4 Moving Average Processes

Intuition X_t is a stock price of a company. Each daily announcement of the company is modeled as a noise. Effect of the daily announcements (noises Z_t) on the stock price (X_t) might last few days (say 2 days). Then stock price is linear combination of the noises that affects it

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2}$$

This model is basically one example of a moving average processes, called moving average model of order 2 [MA(2)].
Generalization

$$X_t = Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \ldots + \theta_q Z_{t-q}$$

where Z_i are i.i.d. and $Z_i \sim Normal(\mu, \sigma)$.
This is called MA(q) where q is the order.

3 Week 3 - Stationarity, MA(q) and AR(p) processes

Reading Materials

1. Examples - White Noise, Random Walks, and Moving Averages
2. Stationarity - Intuition and Definition

Definition of Stochastic Processes A stochastic process is a family of random variables structured with a time index, denoted X_t for discrete processes and $X(t)$ for continuous processes.

Remarks To fully specify the structure of a stochastic process, we need to know the joint distribution of the full set of random variables. So we usually just have one sequentially observed realization and infer properties of the generating process from this single trajectory.

Definition of Strict Stationarity We say a process is strictly stationary if the joint distribution of

$$X(t_1), X(t_2), \ldots, X(t_k)$$

is the same as the joint distribution of

$$X(t_1 + \tau), X(t_2 + \tau), \ldots, X(t_k + \tau)$$

for any τ

Implication

1. The random variables are identically distributed (constant mean and constant variance), though not necessarily independent.

2. Joint distribution of $X(t_1), X(t_2)$ is the same as joint distribution of $X(t_1 + \tau), X(t_2 + \tau)$. That is, the joint distribution depends only on the lag spacing.

Definition of Weak Stationarity We say a process is weakly stationary if it has constant mean and the joint distribution of two timepoints depends only on the lag spacing.

Implication Constant variance

Example 1: White Noise Stationary
Consider a discrete family of independent, identically distributed normal random variables (often Gaussian)

$$X_t \sim N(0, \sigma^2)$$

then

$$\gamma(t_1, t_2) = \begin{cases} 0 & t_1 \neq t_2 \\ \sigma^2 & t_1 = t_2 \end{cases}$$

Example 2: Random Walks Not Stationary
Start with iid random variables $Z_t \overset{i.i.d}{\sim} (\mu, \sigma)$
Build a walk with t steps

$$X_t = X_{t-1} + Z_t = \sum_{i=1}^{t} Z_i$$

$$E[X_t] = E[\sum_{i=1}^{t} Z_i] = \sum_{i=1}^{t} E[Z_i] = t \cdot \mu$$

$$V[X_t] = V[\sum_{i=1}^{t} Z_i] = \sum_{i=1}^{t} V[Z_i] = t \cdot \sigma^2$$

We can see the expectation is 0 or increasing; the variance is increasing.

Moving Average Process MA(q) Stationary

Start with i.i.d random variables $Z_t \overset{i.i.d}{\sim} (\mu, \sigma)$
MA(q) process:

$$X_t = \beta_0 Z_t + \beta_1 Z_{t-1} + \ldots + \beta_q Z_{t-q}$$

when $k \leq q$,

$$Cov[X_t, X_{t+k}] = \sigma^2 \cdot \sum_{i=0}^{q-k} \beta_i \beta_{i+k}$$

3.1 Backward shift operator

Definition We have a stochastic process $\{X_1, X_2, X_3, \ldots\}$. Backward shift operator is defined as $BX_t = X_{t-1}$

Remarks

$$B^2 X_t = BBX_t = BX_{t-1} = X_{t-2}$$
$$B^k X_t = X_{t-k}$$

Example1 - Rewrite Random Walk

$$\begin{aligned} X_t &= X_{t-1} + Z_t \\ X_t &= BX_t + Z_t \\ (1 - B)X_t &= Z_t \\ \phi(B)X_t &= Z_t \qquad (\phi(B) = 1 - B) \end{aligned}$$

Example2 - MA(2) process

$$\begin{aligned} X_t &= Z_t + 0.2Z_{t-1} + 0.04Z_{t-2} \\ X_t &= Z_t + 0.2BZ_t + 0.04B^2 Z_t \\ X_t &= (1 + 0.2B + 0.04B^2)Z_t \\ X_t &= \beta(B)Z_t \qquad (\beta(B) = 1 + 0.2B + 0.04B^2) \end{aligned}$$

Example3 - MA(q) process

$$
\begin{aligned}
X_t &= \mu + \beta_0 Z_t + \beta_1 Z_{t-1} + \ldots + \beta_q Z_{t-q} \\
X_t &= \mu + \beta_0 Z_t + \beta_1 B Z_t + \ldots + \beta_q B^q Z_t \\
X_t - \mu &= \beta(B) Z_t \qquad\qquad (\beta(B) = \phi_0 + \phi_1 B + \ldots + \phi_q B^q)
\end{aligned}
$$

3.2 Invertibility of a stochastic process

Definition $\{X_t\}$ is a stochastic process.
$\{Z_t\}$ is innovations, i.e., random disturbances or white noise. $\{X_t\}$ is called invertible, if $Z_t = \sum_{k=0}^{\infty} \pi_k X_{t-k}$ where $\sum_{k=0}^{\infty} |\pi_k|$ is convergent.

Example - Model 1

$$X_t = Z_t + 2Z_{t-1}$$

model 1 is not invertible since

$$\sum_{k=0}^{\infty} |\pi_k| = \sum_{k=0}^{\infty} 2^k$$

which is divergent

$$\gamma(k) = Cov[X_{t+k}, X_t] = Cov[Z_{t+k} + 2Z_{t+k-1}, Z_t + 2Z_{t-1}]$$

1. If $k > 1$, then $t + k - 1 > t$, so all Z's are uncorrelated, thus $\gamma(k) = 0$.
2. If $k = 0$, then $\gamma(0) = 5\sigma_Z^2$
3. If $k = 1$, then $\gamma(1) = 2\sigma_Z^2$
4. If $k < 0$, then $\gamma(k) = \gamma(-k)$

Then

$$\gamma(k) = \begin{cases} 0 & k > 1 \\ 2\sigma_Z^2 & k = 1 \\ 5\sigma_Z^2 & k = 0 \\ \gamma(-k) & k < 0 \end{cases}$$

$$\rho(k) = \begin{cases} 0 & k > 1 \\ \frac{2}{5} & k = 1 \\ 1 & k = 0 \\ \gamma(-k) & k < 0 \end{cases}$$

Inverting through backward substitution MA(1) process

$$X_t = Z_t + \beta Z_{t-1}$$

$$Z_t = X_t - \beta Z_{t-1} = X_t - \beta(X_{t-1} - \beta Z_{t-2}) = X_t - \beta X_{t-1} + \beta^2 Z_{t-2}$$

$$X_t = \beta(B) Z_t$$

where $\beta(B) = 1 + \beta B$ Then, we find Z_t by inverting the polynomial operator $\beta(B)$

$$\beta(B)^{-1} X_t = Z_t$$

$$\beta(B)^{-1} = \frac{1}{1 + \beta B}$$

In this manner,

$$Z_t = X_t - \beta X_{t-1} + \beta^2 X_{t-2} - \beta^3 X_{t-3} + \dots$$

i.e.

$$X_t = Z_t + \beta X_{t-1} - \beta^2 X_{t-2} + \beta^3 X_{t-3} + \dots$$

$$\beta(B)^{-1} = 1 - \beta B + \beta^2 B^2 - \beta^3 B^3 + \dots$$

Thus we obtain,

$$\beta(B)^{-1} X_t = 1 - \beta X_{t-1} + \beta^2 X_{t-2} - \beta^3 X_{t-3} + \dots$$

$$Z_t = \sum_{n=0}^{\infty} (-\beta)^n X_{t-n}$$

In order to make sure that the sum on the right is convergent, we need $|\beta| < 1$.

Example - Model 2

$$X_t = Z_t + \frac{1}{2} Z_{t-1}$$

model 2 is invertible since

$$\sum_{k=0}^{\infty} |\pi_k| = \sum_{k=0}^{\infty} \frac{1}{2^k}$$

is convergent (geometric series)

3.3 Mean Square Convergence

Let

$$X_1, X_2, X_3, \ldots$$

be a sequence of random variables (i.e. a stochastic process). We say X_n converge to a random variable X in the mean-square sense if

$$E[(X_n - X)^2] \to 0 \text{ as } n \to \infty$$

3.4 Autoregressive Processes

3.4.1 Definition

We are looking at some general stochastic processes that are useful in understanding the driving mechanisms behind the Time Series that we encounter. We've already seen the Random Walk. We can generalize this to an autoregressive process of order p, denoted AR(p).

$$X_t = Z_t + history$$

Let's take the Z_t's to be white noise $Z_t \sim iid(0, \sigma^2)$.
By history we mean that we include previous terms in the process as

$$history = \Phi_1 X_{t-1} + \ldots + \Phi_p X_{t-p}$$

So we then have

$$AR(p)\, process : \quad X_t = Z_t + \Phi_1 X_{t-1} + \ldots + \Phi_p X_{t-p}$$

AR(p) vs MA(q)

$$MA(q)\, process : \quad X_t = \theta_0 Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$$

1. We build an MA(q) from a finite set of innovations (the Z's)
2. We build an AR(p) from a current innovation Z_t together with knowledge of a finite set of prior states (the X's)

Example: Random Walk Take $p = 1$ and $\Phi_1 = 1$

$$X_t = Z_t + X_{t-1}$$

Caution An autoregressive process isn't necessarily stationary!

R code

```
set.seed(2017)
X.ts <- arima.sim(list(ar = c(.7,.2)), n=1000)
par(mfrow=c(2,1))
plot(X.ts, main="AR(2) Time Series, phi1=.7, phi2=.2")
X.acf = act(X.ts, main="Autocorrelation of AR(2) Time Series")
```

Stationarity of an AR(2) In order for an AR(2) to be stationary, we need

$$-1 < \Phi_2 < 1$$
$$\Phi_2 < 1 + \Phi_1$$
$$\Phi_2 < 1 - \Phi_1$$

3.4.2 Backshift Operator and the ACF

Reading Materials

1. Backshift Operator and the ACF

Definition

$$X_{t-1} = BX_t$$
$$X_{t-2} = B^2X_t$$
$$\vdots$$
$$X_{t-p} = B^pX_t$$

Now from the expression

$$X_t = Z_t + \Phi_1 BX_t + \ldots + \Phi_p B^p X_t = Z_t + (\Phi_1 B + \ldots + \Phi_p B^p)X^t$$

Express AR(p) as an Infinite Order Moving Average

$$Z_t = (1 - \Phi_1 B - \ldots - \Phi_p B^p)X_t = \Phi(B)X_t$$

We can write

$$X_t = \frac{1}{1 - (1 - \Phi_1 B - \ldots - \Phi_p B^p)} Z_t = (1 + \theta_1 B + \theta_2 B^2 + \ldots)Z_t$$

Example: p = 1

$$\begin{aligned}
X_t &= Z_t + \Phi B X_t \\
(1-\Phi B)X_t &= Z_t \\
X_t &= \frac{1}{(1-\Phi B)} Z_t = (1 + \Phi B + \Phi^2 B^2 + \ldots) Z_t \\
X_t &= Z_t + \Phi Z_{t-1} + \Phi^2 Z_{t-2} + \ldots \\
X_t &= Z_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \ldots
\end{aligned}$$

Inference

$$\begin{aligned}
E[X_t] &= E[(1 + \theta_1 B + \theta_2 B^2 + \ldots) Z_t] \\
&= E[Z_t] + \theta_1 E[Z_{t-1}] + \ldots + \theta_k E[Z_{t-k}] + \ldots \\
&= 0 \\
V[X_t] &= V[(1 + \theta_1 B + \theta_2 B^2 + \ldots) Z_t] \\
&= V[Z_t] + \theta_1^2 V[Z_{t-1}] + \ldots + \theta_k^2 V[Z_{t-k}] + \ldots \\
&= \sigma_Z^2 (1 + \theta_1^2 + \ldots + \theta_k^2 + \ldots) \\
&= \sigma_Z^2 \sum_{i=0}^{\infty} \theta_i^2
\end{aligned}$$

Necessary condition for stationarity: the sum must converge.

Results: Autocovariance Compare:
It can be shown that, for an MA(q) process,

$$\gamma(k) = \sigma_Z^2 \cdot \sum_{i=0}^{q-k} \theta_i \theta_{i+k}$$

For an AR(p) process,

$$\gamma(k) = \sigma_Z^2 \cdot \sum_{i=0}^{\infty} \theta_i \theta_{i+k}$$

$$\rho(k) = \frac{\sum_{i=0}^{\infty} \theta_i \theta_{i+k}}{\sum_{i=0}^{\infty} \theta_i^2}$$

3.5 Duality

3.5.1 Invertibility condition for MA(q)

MA(q) process

$$X_t = \theta_0 Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$$

is invertible if the roots of the polynomial

$$\beta(B) = \beta_0 + \beta_1 B + \ldots + \beta_q B^q$$

all lie outside the unit circle, where we regard B as a complex variable (not an operator).
(Proof is done using mean-square convergence)

Example: MA(1) process

$$X_t = Z_t + \beta Z_{t-1}$$

$$\beta(B) = 1 + \beta B$$

In this case only one (real) root $B = -\frac{1}{\beta}$

$$|-\frac{1}{\beta}| > 1 \implies |\beta| < 1$$

Then, $Z_t = \sum_{k=0}^{\infty} (-\beta)^k B^k X_t = \sum_{k=0}^{\infty} (-\beta)^k X_{t-k}$

Example - MA(2) process

$$X_t = Z_t + \frac{5}{6} Z_{t-1} + \frac{1}{6} Z_{t-1}$$

Then

$$X_t = \beta(B) Z_t$$

where

$$\beta(B) = 1 + \frac{5}{6} B + \frac{1}{6} B^2$$

$$\begin{aligned}
\beta(B)^{-1} &= \frac{1}{1 + \frac{5}{6}B + \frac{1}{6}B^2} = \frac{3}{1 + \frac{1}{2}B} - \frac{2}{1 + \frac{1}{3}B} \qquad \text{(two geometric series)} \\
&= \sum_{k=0}^{\infty} [3(-\frac{1}{2})^k - 2(-\frac{1}{3})^k] B^k \\
Z_t &= \sum_{k=0}^{\infty} [3(-\frac{1}{2})^k - 2(-\frac{1}{3})^k] B^k X_t
\end{aligned}$$

MA(2) process $\Rightarrow$ AR(∞) process

3.5.2 Stationary condition for AR(p)

AR(p) process

$$X_t = \Phi_1 X_{t-1} + \Phi_2 X_{t-1} + \ldots + \Phi_p X_{t-p} + Z_t$$

is (weakly) stationary if the roots of the polynomial

$$\Phi(B) = 1 - \Phi_1 B - \Phi_2 B^2 - \ldots - \Phi_p B^p$$

all lie outside the unit circle, where we regard B as a complex variable (not an operator).

Example: AR(1) process

$$\begin{aligned}
X_t &= \Phi_1 X_{t-1} + Z_t \Rightarrow (1 - \Phi_1 B) X_t = Z_t \\
\Phi(b) &= 1 - \Phi_1 B = 0 \\
\Rightarrow B &= \frac{1}{\Phi_1} \\
|\frac{1}{\Phi_1}| &> 1 \\
\Rightarrow |\Phi_1| &< 1
\end{aligned}$$

Thus, when $|\Phi_1| < 1$, the AR(1) process is stationary.

$$\begin{aligned}
X_t &= \frac{1}{1 - \Phi_1 B} Z_t \\
&= (1 + \Phi_1 B + \Phi_1^2 B^2 - \ldots) Z_t \\
&= \sum_{k=0}^{\infty} \Phi_1^k Z_{t-k}
\end{aligned}$$

Take the variance from both side,

$$Var[X_t] = \sigma_Z^2 \sum_{k=0}^{\infty} \Phi_1^{2k}$$

which is a convergent geometric series if $|\Phi_1^2| < 1$, i.e.,

$$|\Phi_1| < 1$$

AR(1) process $\Rightarrow$ MA(∞) process

Duality Under invertibility condition of MA(q),

$$MA(q) \Rightarrow AR(\infty)$$

Under stationarity condition of AR(p),

$$AR(p) \Rightarrow MA(\infty)$$

3.6 Intro to Yule-Walker Equations

3.6.1 Difference Equations

Example: $a_n = 5a_{n-1} - 6a_{n-2}$

k-th Order Difference Equation

$$a_n = \beta_1 a_{n-1} + \beta_2 a_{n-2} + \ldots + \beta_k a_{n-k}$$

Its characteristic equation

$$\lambda^k - \beta_1 \lambda^{k-1} - \ldots - \beta_{k-1}\lambda - \beta_k = 0$$

Then we look for the solutions of the characteristic equation. Say, all k solutions are distinct real numbers, $\lambda_1, \lambda_2, \ldots, \lambda_k$, then

$$a_n = c_1\lambda_1^n + c_2\lambda_2^n + \ldots + c_n\lambda_k^n$$

Coefficients c_j's are determined using intial(given) values.

Example - Fibonacci Sequence We are looking for a sequence $\{a_n\}_{n=0}^{\infty}$ such that

$$a_n = a_{n-1} + a_{n-2}$$

where $a_0 = 1, a_1 = 1$ Characteristic equation becomes

$$\lambda^2 - \lambda - 1 = 0$$

Then $\lambda_1 = \frac{1-\sqrt{5}}{2}$ and $\lambda_2 = \frac{1+\sqrt{5}}{2}$. Thus

$$a_n = c_1\left(\frac{1-\sqrt{5}}{2}\right)^n + c_2\left(\frac{1+\sqrt{5}}{2}\right)^n$$

Use initial data

$$\begin{cases} c_1 + c_2 = 1 \\ c_1(\frac{1-\sqrt{5}}{2})^n + c_2(\frac{1+\sqrt{5}}{2})^n = 1 \end{cases}$$

We obtain

$$c_1 = \frac{5-\sqrt{5}}{10} = -\frac{1}{\sqrt{5}}(\frac{1-\sqrt{5}}{2})$$

$$c_2 = \frac{5+\sqrt{5}}{10} = \frac{1}{\sqrt{5}}(\frac{1+\sqrt{5}}{2})$$

$$a_n = -\frac{1}{\sqrt{5}}(\frac{1-\sqrt{5}}{2})^{n+1} + \frac{1}{\sqrt{5}(\frac{1+\sqrt{5}}{2})^{n+1}}$$

3.6.2 Yule- Walker Equations

Example We have an AR(2) process

$$X_t = \frac{1}{3}X_{t-1} + \frac{1}{2}X_{t-2} + Z_t \quad (*)$$

Polynomial

$$\phi(B) = 1 - \frac{1}{3}B - \frac{1}{2}B^2$$

has real roots $\frac{-2\pm\sqrt{76}}{6}$ both of which has magnitude greater than 1, so roots are outside of the unit circle in $\mathbb{R}^2$. Thus, this AR(2) process is a stationary process.
Multiply both side of (*) with X_{t-k}, and take expectation

$$E(X_{t-k}X_t) = \frac{1}{3}E(X_{t-k}X_{t-1}) + \frac{1}{2}E(X_{t-k}X_{t-2}) + E(X_{t-k}Z_t)$$

Since $\mu = 0$, and assume $E(X_{t-k}Z_t) = 0$,

$$\gamma(-k) = -\frac{1}{3}\gamma(-k+1) + \frac{1}{2}\gamma(-k+2)$$

Since $\gamma(k) = \gamma(-k)$ for any k,

$$\gamma(k) = \frac{1}{3}\gamma(k-1) + \frac{1}{2}\gamma(k-2)$$

Divide by $\gamma(0) = \sigma_X^2$,

$$\rho(k) = \frac{1}{3}\rho(k-1) + \frac{1}{2}\rho(k-2)$$

This set of equations is called Yule-Waker equations.
We look for a solution in the format of $\rho(k) = \lambda^k$.

$$\lambda^2 - \frac{1}{3}\lambda - \frac{1}{2} = 0$$

Roots are $\lambda_1 = \frac{2+\sqrt{76}}{12}$ and $\lambda_2 = \frac{2-\sqrt{76}}{12}$, thus

$$\rho(k) = c_1(\frac{2+\sqrt{76}}{12})^k + c_2(\frac{2-\sqrt{76}}{12})^k$$

Use constraints to obtain coefficients

$$\rho(0) = 1 \rightarrow c_1 + c_2 = 1$$

And for $k = p - 1 = 2 - 1 = 1$,

$$\rho(k) = \rho(-k)$$

Thus,

$$\rho(1) = \frac{1}{3}\rho(0) + \frac{1}{2}\rho(-1)$$

$$\rightarrow \rho(1) = \frac{2}{3} \rightarrow c_1(\frac{2+\sqrt{76}}{12}) + c_2(\frac{2-\sqrt{76}}{12}) = \frac{2}{3}$$

We have

$$\begin{cases} c_1 + c_2 = 1 \\ c_1(\frac{2+\sqrt{76}}{12}) + c_2(\frac{2-\sqrt{76}}{12}) = \frac{2}{3} \end{cases}$$

Then $c_1 = \frac{4+\sqrt{6}}{8}$ and $c_2 = \frac{4-\sqrt{6}}{8}$.
For any $k \geq 0$,

$$\rho(k) = \frac{4+\sqrt{6}}{8}(\frac{2+\sqrt{76}}{12})^k + \frac{4-\sqrt{6}}{8}(\frac{2-\sqrt{76}}{12})^k$$

and

$$\rho(k) = \rho(-k)$$