ESE538 Assignment 1 Sep. 15, 2024

JieWang key: tonyw3 ID: 83749364

Exercises

1. Consider a random process defined by the recursion: $Y_{k+1} = \alpha k + \phi Y_k + \epsilon_k$, where $\alpha \in \mathbb{R}$, $|\phi| < 1$, ϵ_k is a white noise with mean zero and constant variance σ_ϵ^2, and the initial condition is $Y_0 = 0$. Answer the following questions:

leave it as a finite summation

   (a) Compute the theoretical mean and variance of the process as a function of k.

   (b) Compute the autocorrelation function of the process for large k.

   (c) Is the process wide-sense stationary? Explain your answer.

   (d) Program a piece of Python code to generate and plot 10 sample paths of length 100 of the random process. Include a shaded area indicating the 95% confidence interval of the process, assuming that the process is normally distributed. Use the values of $\phi = 0.8$ and ~~$\sigma_\epsilon^2 = 1$~~ in your simulations. $\alpha = 1$, $\sigma_\epsilon^2 = 10$

(a): $\because Y_0 = 0 \quad \therefore E[Y_0] = 0$

Since ϵ_k is white noise, $E[\epsilon_k] = 0$

$\therefore E[Y_{k+1}] = E[\alpha k + \phi Y_k + \epsilon_k] = \alpha k + \phi E[Y_k]$

Then: $E[Y_1] = 0 + \phi \cdot 0$

$E[Y_2] = \alpha \cdot 1 + \phi E[Y_1]$

$\vdots$

$E[Y_k] = \alpha(k-1) + \phi E[Y_{k-1}]$

$E[Y_{k+1}] = \alpha k + \phi E[Y_k]$

it's a 1st order non homogenous linear recurrence

$$\Rightarrow E[Y_k] = \alpha \sum_{i=0}^{k} \phi^i (k-i) + \phi^{k+1} E[Y_0]$$

$$= \alpha \sum_{i=0}^{k} \phi^i (k-i) + 0$$

$$= \frac{\alpha}{1-\phi}\left[k + \frac{\phi^k - 1}{1-\phi}\right]$$

$Var(Y_{k+1}) = Var(\alpha k + \phi Y_k + \epsilon_k)$

$Var(Y_{k+1}) = 0 + \phi^2 Var(Y_k) + \sigma_\epsilon^2$

$Var(Y_{k+1}) = \phi^2 Var(Y_k) + \sigma_\epsilon^2$

let $\mu_k = E[Y_k]$, then: $\mu_{k+1} = \alpha k + \phi \mu_k$, $\mu_0 = 0$

let $Z_k = Y_k - \mu_k$, then Z_k satisfies:

$Z_{k+1} = Y_{k+1} - \mu_{k+1} = (\alpha k + \phi Y_k + \epsilon_k) - (\alpha k + \phi \mu_k)$

$= \phi(Y_k - \mu_k) + \epsilon_k$

$\Rightarrow Z_{k+1} = \phi Z_k + \epsilon_k$

So Z_k is AR(1), $Var(Z_k) = \dfrac{\sigma_\epsilon^2 (1-\phi^2)}{1-\phi^2}$

Then: $Var(Y_k) = Var(Z_k + \mu_k)$

$= Var(Z_k) - Var(\mu_k)$

As $k \to \infty$, $\Rightarrow Var(Y_k) = \dfrac{\sigma_\epsilon^2}{1-\phi^2}$ is a constant

(b): Auto variance func of Y_k

is $Cov(Y_k, Y_{k-h}) = \phi^{|h|} \gamma(0)$

$\Rightarrow$ Auto correlation func

is $\rho(h) = \dfrac{\gamma(h)}{\gamma(0)} = \phi^{|h|}$

when $k \to \infty$, we have stationary ans $\phi^{|h|}$

(c): let's verify WSS condition:

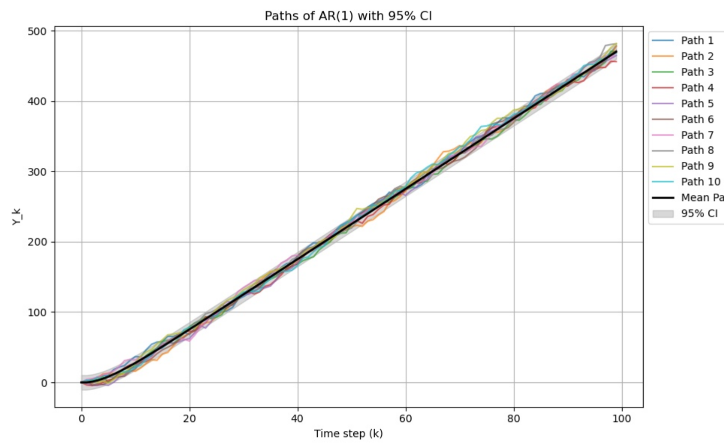
① Stationary Mean $\boxed{\times}$ → it's func of k

② Stationary Variance $\boxed{\checkmark}$

③ Auto covariance: $\boxed{\checkmark}$ ⇒ depends only on h

So r.p. is not WSS

(d): Here's simulation result



2. Consider a random process Y_k defined by the recursion:

$$Y_k = 1 + \epsilon_k + \theta_1 \epsilon_{k-1} \text{ with } Y_0 = 0, \tag{7}$$

where $\epsilon(k) \overset{\text{iid}}{\sim} \mathcal{N}(0, 1)$, i.e., independent and identically distributed standard Gaussians. Answer the following questions:

(a) Compute the theoretical mean $\mu(k)$ and variance $\sigma^2(k)$ as a function of k.

(b) Compute the theoretical autocovariance $\text{Cov}(Y_k, Y_{k-h})$.

(c) Is the random process strong-sense stationary? Explain your answer.

(d) Program a piece of Python code to generate and plot 10 sample paths of length 100 of the random process. Include a shaded area indicating the 95% confidence interval of the process, assuming that the process is normally distributed. Use the value $\theta_1 = 0.5$ in your simulations.

(a):

$\mu(k) = E[Y_k] = E[1 + \epsilon_k + \theta_1 \epsilon_{k-1}]$

$= 1 + E[\varepsilon_k] + \theta_1 E[\varepsilon_{k-1}]$

Since $\epsilon(k) \overset{iid}{\sim} N(0,1)$. we know $E[\varepsilon_k] \equiv 0$ for $\forall k \in \mathbb{Z}$

$\rightarrow \mu(k) \equiv 1$ for $\forall k \in \mathbb{Z}$

$\sigma^2(k) = Var(Y_k) = Var(1 + \varepsilon_k + \theta_1 \varepsilon_{k-1})$

$= 0 + 1 + \theta_1^2$

$= \theta_1^2 + 1$ for $\forall k \in \mathbb{Z}$

(b): if $h=0$, then:

$Cov(Y_k \cdot Y_{k-h}) = Var(Y_k) = \theta_1^2 + 1$

if $h=1$, then we have:

$\begin{cases} Y_k = 1 + \epsilon_k + \theta_1 \epsilon_{k-1} \\ Y_{k-1} = 1 + \varepsilon_{k-1} + \theta_1 \varepsilon_{k-2} \end{cases} \Rightarrow$ only common variable is ε_{k-1}

So $Cov(Y_k \cdot Y_{k-h}) = Cov(\varepsilon_{k-1}, \varepsilon_{k-1}) = Var(\varepsilon_{k-1}) = 1$

if $h \geq 2$: we have:

$\begin{cases} Y_k = 1 + \epsilon_k + \theta_1 \epsilon_{k-1} \\ Y_{k-2} = 1 + \varepsilon_{k-2} + \theta_1 \varepsilon_{k-3} \\ Y_{k-3} = \cdots \end{cases} \Rightarrow$ No common term between them

So Covariance $\equiv 0$

$\Rightarrow Cov(Y_k \cdot Y_{k-h}) = \begin{cases} \theta_1^2 + 1, & \text{if } h=0 \\ 1 & \text{if } h=1 \\ 0 & \text{if } h \geq 2 \end{cases}$

(c): Let's check SSS condition one by one: satisfied?

① $\mu(k)$ is constant ☑

② $\sigma^2(k)$ is constant ☑

$\Rightarrow$ as θ_1 is a constant

③ $Cov(Y_k \cdot Y_{k-h})$ is solely dependent on lag h ☑

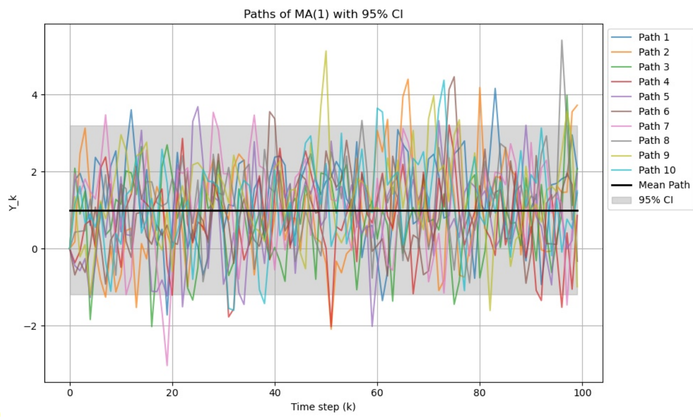
And we prove Y_k to be WSS

Since Y_k is linear combination of gaussian distribution,

$\Rightarrow$ Y_k is also SSS

So, we have Y_k to be SSS proved.

(d): here is simulated python plot



3. Let $\alpha \sim \mathcal{N}(0,1)$ and $\beta \sim \mathcal{N}(0,2)$ be two independent random variables (independent of k). Define the random process $\mathcal{Y} = \{Y_k : k \in \mathbb{N}\}$ as the recursion:

$$Y_k = \alpha + \beta k + k^2 \text{ for all } k \in \mathbb{N}.$$

Find expressions for the mean, the variance, and the autocovariance of the random process as a function of k, as well as the lag h for the covariance.

(1)

$\mu(k) = E[Y_k] = E[\alpha + \beta k + k^2]$

As α, β is mean $= 0$ $= E[\alpha] + kE[\beta] + E[k^2]$

k^2 is deterministic $= 0 + k \cdot 0 + k^2$

$= k^2$

(2)

$\sigma^2(k) = Var(Y_k) = Var(\alpha + \beta k + k^2)$

$= Var(\alpha) + Var(\beta k)$

$= 1 + 2k^2$

(3)

$Y_{k-h} = \alpha + \beta(k-h) + (k-h)^2$

$Cov(Y_k, Y_{k-h}) = Cov(\alpha + \beta k, \alpha + \beta(k-h))$

$= Cov(\alpha, \alpha) + Cov(\beta k, \beta(k-h))$

$= 1 + 2k(k-h)$

(4) Let's check WSS condition:

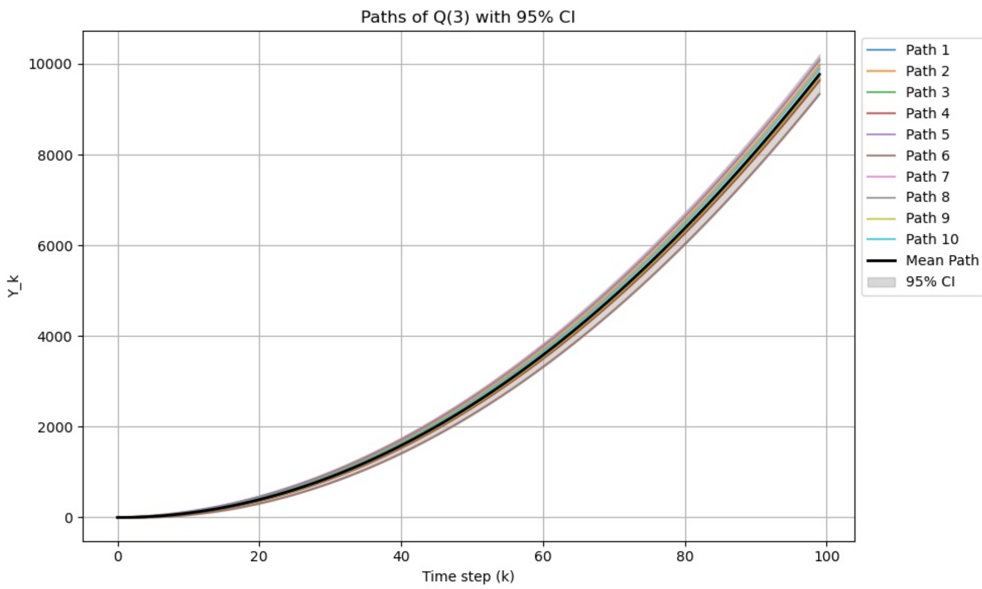
① Constant mean : $\boxed{X}$

② Constant variance $\boxed{X}$

③ Time invariant Cov $\boxed{X}$

So it's not WSS

BTW, here's plot



4. Consider a fair coin toss game where you start with \$10. Every time you flip the coin:

- If it lands heads, you win \$1.
- If it lands tails, you lose \$1.

Let Y_k represent the amount of money you have after k coin flips, with $Y_0 = 10$. For each coin flip, the outcome is independent of previous flips, and the probability of heads or tails is 1/2. Answer the following questions

(a) Find the expected value of Y_k, i.e., the amount of money you will have after k coin flips.

(b) Compute the variance of Y_k after k coin flips.

$P_H = P_T = 50\%$, then $Y_k = Y_{k-1} + P_H \times 1$

(a): $\Rightarrow Y_k = Y_0 + \sum_{i=0}^{k} x_i$, where x_i is R.V. of x_i to be head at time i

$E[x_i] = \frac{1}{2} \times (1) + \frac{1}{2} \times (-1) = 0$

$$E[Y_k] = E\left[Y_0 + \sum_{i=0}^{k} x_i\right]$$

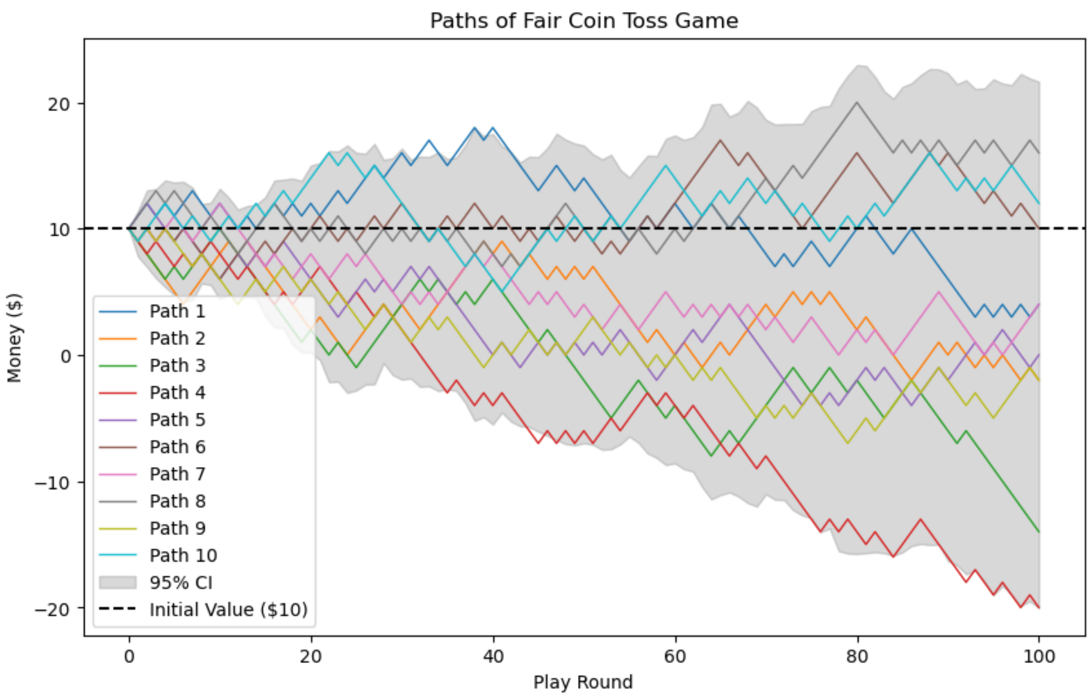
$$= 10 + E\left[\sum_{i=0}^{k} x_i\right] = 10 + \sum_{i=0}^{k} E[x_i] = 10$$

(b): $Var(x_i) = E[x_i^2] - E^2[x_i]$

while $E[x_i^2] = \frac{1}{2} \times (1)^2 + \frac{1}{2}(-1)^2 = 1$

So $Var(Y_k) = Var\left(\sum_{i=0}^{k} x_i\right) = \sum_{i=0}^{k} Var(x_i) = k$

(c): don't gamble!



5.

Example 5: (Asymptotic) Autocovariance of AR(1)

Consider the AR(1) stochastic process, defined by the recursion:

$$Y_k = \phi Y_{k-1} + \epsilon_k \quad \text{with} \quad |\phi| < 1,$$

5. Modify the proof in Example 5 to consider the correlation between Y_k and Y_{k+h} (with a positive lag h, rather than a negative lag). Follow these steps:

   (a) Start by expressing Y_{k+h} as a function of ϕ, Y_k, and the noise terms ϵ at different time steps.

   (b) Compute the expectation $\mathbb{E}[Y_k Y_{k+h}]$ and expand the equation into a sum of expectations.

   (c) Apply the statistical properties of the noise terms (i.e., their zero mean and lack of correlation) to simplify the expression.

   (d) Compare your expression for the autocovariance to the one obtained in Example 5, noting that the final expression will be similar but not identical.

ex5:

$$E[Y_k Y_{k-h}] = \phi^h E[Y_{k-h}^2]$$

$$Cov(Y_k, Y_{-h}) = \phi^h \frac{\sigma_\varepsilon^2}{1-\phi^2}$$

(a): $Y_{k+h} = \phi^{(h)} Y_k + \sum_{i=1}^{h} \phi^{h-i} \varepsilon_{k+i}$

(b): $E[Y_k Y_{k+h}] = E\left[\phi^h Y_k^2 + Y_k \sum_{i=1}^{h} \phi^{h-i} \varepsilon_{k+i}\right]$

$= \phi^{(h)} E[Y_k^2] + \sum_{i=1}^{h} \phi^{h-i} E[Y_k \varepsilon_{k+i}]$

(c): Because ε_{k+i} is white noise

$\Rightarrow$ $E[\varepsilon_{k+i}] = 0$

ε_{k+i} is not correlated w/ Y_k

So for $\forall i \in \mathbb{Z}$, $E[Y_k \varepsilon_{k+i}] = 0$

$\Rightarrow$ we have $E[Y_k Y_{k+h}] = \phi^h E[Y_k^2]$

And $Cov(Y_k, Y_{+h}) = \phi^h \frac{\sigma_\varepsilon^2}{1-\phi^2}$

(d): Though we have same formula of the auto regression,

$\Rightarrow$ past lag should be abs value as the auto covariance should be solely dependent on the lag length, instead of the direction

correlation:

$Cor[Y_k, Y_{k+h}] = \phi^h$

Same as Example 5

$\Rightarrow \phi^h \frac{\sigma_\varepsilon^2}{1-\phi^2}$

Note: I've posted on Ed, querying about this

6.

6. Consider the AR(2) process defined in Example 7. Derive an expression for the probability density function:

$$f_{Y_k|\mathcal{F}_{k-1}}(Y_k \mid Y_{k-1}=y_{k-1}, Y_{k-2}=y_{k-2}, Y_{k-3}=y_{k-3}, \dots).$$

Since AR(2) is second order Markov Process, we know only last two values are correlated

$$\Rightarrow f_{Y_k|F_{k-1}}(Y_k \mid Y_{k-1}=y_{k-1}, Y_{k-2}=y_{k-2}, Y_{k-3}=y_{k-3}, \dots)$$

$$= f_{Y_k|F_{k-1}}(Y_k \mid Y_{k-1}=y_{k-1}, Y_{k-2}=y_{k-2})$$

Given $Y_k = \phi_0 Y_{k-1} + \phi_1 Y_{k-2} + \varepsilon_{k-1}$, $\varepsilon_{k-1} \sim N(0, \sigma^2)$

$$\Rightarrow Y_k \mid Y_{k-1}, Y_{k-2} \sim N(\mu_k, \sigma^2), \quad \mu_k = \phi_0 Y_{k-1} + \phi_1 Y_{k-2}$$

and here $|\phi_0| + |\phi_1| < 1$ ↳ from ex 7

Since, $f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ for $X \sim N(\mu, \sigma^2)$

↓ plug in μ_k

So $f_{Y_k|F_{k-1}}(Y_k \mid Y_{k-1}=y_{k-1}, Y_{k-2}=y_{k-2})$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}\left[y_k - (\phi_0 y_{k-1} + \phi_1 y_{k-2})^2\right]}$$

7.

7. Consider the AR(3) process defined by the recursion:

$$Y_{k+1} = \phi_0 Y_k + \phi_1 Y_{k-1} + \phi_2 Y_{k-2} + \epsilon_k,$$

where $\epsilon_k \overset{\text{iid}}{\sim} \mathcal{N}(0, \sigma^2)$ is a white noise process. This process is autoregressive of order 3, meaning it depends on the previous three observations.

Answer the questions below:

(a) What is the order of this higher-order Markov process? Justify your answer.

(b) What is the conditional distribution of Y_{k+1} given all the previous observations?

(c) Rewrite the AR(3) process as a vector-valued first-order Markov process. Define the AR(3) process as a vector-valued recursion using a transition matrix and provide an explicit expression for the covariance matrix of the noise vector.

(a):

It's a 3 order Markov process (*)

As future state Y_{k+1} depends only on Y_k, Y_{k-1}, Y_{k-2}

$\Rightarrow P(Y_{k+1} \mid Y_k, Y_{k-1}, Y_{k-2}, \ldots, Y_0) = P(Y_{k+1} \mid Y_k, Y_{k-1}, Y_{k-2})$

Justified (*)

(b): Because ϕ_0, ϕ_1, ϕ_2 are constant, Then Y_k is linear combination of guassian distribution

$\Rightarrow Y_{k+1} \sim N_0(\phi_0 Y_k + \phi_1 Y_{k-1} + \phi_2 Y_{k-2}, \sigma^2)$

(c): let $\vec{X}_k = \begin{pmatrix} Y_k \\ Y_{k-1} \\ Y_{k-2} \end{pmatrix}$

Then, AR(3) is:

$$\vec{X}_{k+1} = A \vec{X}_k + \vec{W}_k$$

where $A = \begin{bmatrix} \phi_0 & \phi_1 & \phi_2 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ $W_k = \begin{bmatrix} \varepsilon_k \\ 0 \\ 0 \end{bmatrix}$

Since $\varepsilon_k \sim N(0, \sigma^2)$ is uncorrelated with past values

$\Rightarrow Q = E[W_k W_k^T] = \sigma^2 \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

8.

8. Consider a two-dimensional VAR(1) process defined by the recursion:

$$\mathbf{Y}_k = A\mathbf{Y}_{k-1} + \boldsymbol{\epsilon}_k,$$

where:

$$\mathbf{Y}_k = \begin{bmatrix} Y_k^1 \\ Y_k^2 \end{bmatrix}, \quad A = \begin{bmatrix} \phi_{11} & 0 \\ \phi_{21} & \phi_{22} \end{bmatrix}, \quad \boldsymbol{\epsilon}_k \sim \mathcal{N}(\mathbf{0}, \mathbb{I}_2),$$

and $\boldsymbol{\epsilon}_k = \begin{bmatrix} \epsilon_k^1 \\ \epsilon_k^2 \end{bmatrix}$ with independent noise terms ϵ_k^1 and ϵ_k^2, each drawn from $\mathcal{N}(0,1)$.

Answer the following questions:

(a) Does the autocorrelation function of $\mathcal{Y}^1 = \{Y_k^1 : k \in \mathbb{N}\}$ depend on ϕ_{21}? Explain your answer.

(b) Does the autocorrelation of $\mathcal{Y}^2 = \{Y_k^2 : k \in \mathbb{N}\}$ depend on ϕ_{21}? Explain your answer.

(c) Derive an expression for the autocorrelation function for $\mathcal{Y}^1$.

(d) What is the cross-correlation between $\mathcal{Y}^1$ and $\mathcal{Y}^2$ at lag 0 and lag 1?

(a): $\begin{bmatrix} Y_k^1 \\ Y_k^2 \end{bmatrix} = \begin{bmatrix} \phi_{11} & 0 \\ \phi_{21} & \phi_{22} \end{bmatrix} \begin{bmatrix} Y_{k-1}^1 \\ Y_{k-1}^2 \end{bmatrix} + \begin{bmatrix} \varepsilon_k^1 \\ \varepsilon_k^2 \end{bmatrix}$

$\Rightarrow \begin{cases} Y_k^1 = \phi_{11} Y_{k-1}^1 + \varepsilon_k^1 \\ Y_k^2 = \phi_{21} Y_{k-1}^1 + \phi_{22} Y_{k-1}^2 + \varepsilon_k^2 \end{cases}$

Since $\mathcal{Y}^1$ is a standard AR(1)

solely depends on ϕ_{11}

Then we say:

$ACF(\vec{Y}^1, h) = \phi_{11}^h, \quad h \geq 0$

$\Rightarrow \mathcal{Y}^1$ doesn't depend on ϕ_{21}

(b): Yes, as Y_k^1 is AR(1)

Y_k^2 is based on Y_{k-1}^1 via ϕ_{21}

(c): Y^1 is AR(1),

So $Cov(Y_k^1, Y_{k-h}^1) = \phi_{11}^{|h|} \gamma_{Y_1}(0) = \phi_{11}^{|h|} \frac{\sigma_{\varepsilon_1}^2}{1-\phi_{11}^2} = \frac{\phi_{11}^{|h|}}{1-\phi_{11}^2}$

Autocorrelation is $\frac{\gamma_{Y_1}(h)}{\gamma_{Y_1}(0)} = \phi_{11}^{|h|}$

(d): skip

--

(e) Is the conditional distribution $F_{\mathbf{Y}_k|\mathcal{F}_{k-1}}$ a Gaussian distribution? If so, what are the mean and the covariance matrix of this conditional distribution?

(f) Is VAR(1) a Markov process? Justify your answer based on the definition of a vector-valued Markov process.

(g) If $\phi_{21} \neq 0$, does $\mathcal{Y}^1$ Granger-cause $\mathcal{Y}^2$? Does $\mathcal{Y}^2$ Granger-cause $\mathcal{Y}^1$?

(h) If $\phi_{21} = 0$, how would this impact Granger causality between the two processes?

(e): Yes. As we have a 1st order VAR(1),

$\Rightarrow \vec{Y}_k$ is a deterministic ter given $A\vec{Y}_{k-1}$

and white noise $\vec{\varepsilon}_k$

Then it should be a linear combination of guassain distribution

Conditional Mean:

$E[\vec{Y}_k | F_{k-1}] = A\vec{Y}_{k-1} = \begin{bmatrix} \phi_{11} Y_{k-1}^1 \\ \phi_{21} Y_{k-1}^1 + \phi_{22} Y_{k-1}^2 \end{bmatrix}$

$Cov(Y_k | F_{k-1}) = Cov[\varepsilon_k] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

(f): Yes, because $\vec{Y}_k$ is solely dependent on $\vec{Y}_{k-1}$.

$\Rightarrow P(\vec{Y}_k | \vec{Y}_{k-1}, \vec{Y}_{k-2}, \cdots, \vec{Y}_0) = P(\vec{Y}_k | \vec{Y}_{k-1})$

then it's a 1st order Markov press

(g): ① Y^1 Granger-Causes Y^2

② Y^2 doesn't Granger-Causes Y^1

As note in (a), dependence of Y_{k-1}^1 & Y_{k-1}^2 shows

(h): $\Rightarrow \begin{cases} Y_k^1 = \phi_{11} Y_{k-1}^1 + \epsilon_k^1 \\ Y_k^2 = \phi_{22} Y_{k-1}^2 + \epsilon_k^2 \end{cases}$

Since there are no crossdependence,

$\Rightarrow$ No Granger-casuality between Y^1 or Y^2