Multivariate Time Series Analysis Book Complementary

Adhitya Arif Wibowo (AWibowo6@gatech.edu)

21 April 2022

Contents

Preface **2**

How To Use The Document **2**

Prerequisites **2**

R **2**

Book Related Library 3

Reference By Course Lecture **3**

Module 4 - Lesson 1 - Slide 4 3

Module 4 - Lesson 2 - Slide 3 3

Module 4 - Lesson 2 - Slide 4 3

Module 4 - Lesson 2 - Slide 6 and 7 3

Module 4 - Lesson 3 - Slide 10 and 11 3

Module 4 - Lesson 4 - Slide 5, 6, 7, and 8 3

Module 4 - Lesson 4 - Slide 9, 10, and 11 4

Module 4 - Lesson 4 - Slide 9, 10, and 11 4

Module 4 - Lesson 6 - Slide 4, 5, 6 4

Module 4 - Lesson 6 - Slide 7 4

Module 4 - Lesson 8 4

Module 4 - Lesson 10 - Slide 3 4

Module 4 - Lesson 10 - Slide 6 4

Module 4 - Lesson 12 4

Reference By Book 5

1.1 INTRODUCTION 5

1.2.1 Stationarity 5

1. Mean (1^{st} Moment) 5

2. Covariance (2^{nd} Moment) 5

1.2.2 Linearity 6

1.4 SAMPLE CCM 8

2.1 INTRODUCTION 13

2.4 VAR(p) MODELS 13

2.4.1 A VAR(1) Representation 14

Example 2.1 15

Preface

This document is created based on *Multivariate Time Series Analysis : With R and Financial Applications, 1st Edition, 2014* by *Ruey S.Tsay* as one of the textbook suggested to be used in *ISYE 6402 - Time Series Analysis* course taught in Georgia Tech.

The scope of the document only discuss materials in the book that are also discussed in the lecturer.

How To Use The Document

The course related documents reference will be the documents used in Spring 2022 semester for the course. As for the book reference will be based on the 1^{st} edition published in 2014 of the book mentioned in Preface section above.

One should have both the book and the course related documents opened to refer to section, example, figures, module, lesson, slide etc., that are discussed in this document.

Prerequisites

1. Basic probability and statistics, refer to *Probability, Statistics, and Stochastic Processes* by *Peter Olofsson* [https://g.co/kgs/tHPF95].
2. Basic programming in R or any language is expected to follow along with the code in this document, refer to *R for Everyone: Advanced Analytics and Graphics* by *Jared P. Lander* [https://g.co/kgs/qjgw2B].

R

The programming language that is used in the course is R, this section will deal with the most common setup throughout the document.

Book Related Library

Below is the library related with the book *Multivariate Time Series Analysis : With R and Financial Applications, 1st Edition, 2014* by *Ruey S.Tsay*:

```
install.packages("MTS")
```

Reference By Course Lecture

Module 4 - Lesson 1 - Slide 4

Multivariate Time Series

1. 1.1 INTRODUCTION

Module 4 - Lesson 2 - Slide 3

Multivariate Time Series

1. 1.1 INTRODUCTION

Module 4 - Lesson 2 - Slide 4

Stationary Multivariate Time Series

1. 1.2.1 Stationarity

Module 4 - Lesson 2 - Slide 6 and 7

Contemporaneous, Cross Covariance & Correlation Estimators

1. 1.4 SAMPLE CCM

Module 4 - Lesson 3 - Slide 10 and 11

Simulation Study: White Noise

1. 1.4 SAMPLE CCM

Module 4 - Lesson 4 - Slide 5, 6, 7, and 8

VAR Model Formulation, Stable VAR(p) process.

1. 2.1 INTRODUCTION
2. 2.4 VAR(p) MODELS
3. 2.4.1 A VAR(1) Representation
4. 1.2.2 Linearity
5. Example 2.1

Module 4 - Lesson 4 - Slide 9, 10, and 11

VAR vs Multivariate Regression

1. [2.5.1 Least-Squares Methods]

Module 4 - Lesson 4 - Slide 9, 10, and 11

Exogenous VAR : VARX(p)

1. [6.3.1 VARX Models]

Module 4 - Lesson 6 - Slide 4, 5, 6

OLS Estimation

1. [2.5.1.2 Ordinary Least-Squares Estimate]

Module 4 - Lesson 6 - Slide 7

GLS Estimation

1. [2.5.1.1 Generalized Least Squares Estimate]

Module 4 - Lesson 8

VAR(p) Order Selection

1. [2.6 Order Selection]

Module 4 - Lesson 10 - Slide 3

VAR(p) models and Granger Causality

1. [2.2.1 Model Structure and Granger Causality]

Module 4 - Lesson 10 - Slide 6

Forecasting VAR(p) models

1. [2.9 FORECASTING]

Module 4 - Lesson 12

Modelling Nonstationary Multivariate Time Series

1. [Unit-Root Nonstationary Processes]

Reference By Book

1.1 INTRODUCTION

A *multivariate series* of k-dimensional (k-variable) equally spaced time series observed at time index t:

$$\vec{z}_t = \begin{bmatrix} z_{1,t} \\ z_{2,t} \\ ... \\ z_{k,t} \end{bmatrix}$$

where we have the row $i = 1, 2, ..., k$ as the dimension (variable) index, and $t = 1, 2, ..., T$ as the time index.

In one of the book example we have *2*-dimensional time series, $z_{1,t}$ as quarterly U.S. real gross domestic product (GDP) and $z_{2,t}$ as quarterly U.S. civilian unemployment rate from 1948 to 2011. For this particular example then we have:

$$\vec{z}_t = \begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix}$$

or if we were to think of using labels instead of index, we have:

$$\vec{z}_t = \begin{bmatrix} z_{GDP,t} \\ z_{UnEmp,t} \end{bmatrix}$$

1.2.1 Stationarity

We should have the first two moments of $\vec{z}_t$ to be time invariant.

1. Mean (1^{st} Moment)

$E[\vec{z}_t] = \vec{\mu}$, a k-dimensional **constant** vector.

$$E[\vec{z}_t] = E\left[\begin{bmatrix} z_{1,t} \\ z_{2,t} \\ ... \\ z_{k,t} \end{bmatrix}\right] = \begin{bmatrix} E[z_{1,t}] \\ E[z_{2,t}] \\ ... \\ E[z_{k,t}] \end{bmatrix} = \begin{bmatrix} \mu_1 \\ \mu_2 \\ ... \\ \mu_k \end{bmatrix} = \vec{\mu} \tag{1.5a}$$

2. Covariance (2^{nd} Moment)

This section comes from the book section **1.3 CROSS-COVARIANCE AND CORRELATIONMA-TRICES**, I move the discussion earlier to match the lecture slide flow.

The **between-dependence** or **lag** l **cross-covariance matrix** Γ_l should be a function of l, not the time index t:

$$
\begin{aligned}
\Gamma_l = Cov[\vec{z}_t, \vec{z}_{t-l}] &= E[(\vec{z}_t - \vec{\mu})(\vec{z}_{t-l} - \vec{\mu})^T] \qquad (1.5b)\\
&= E\left[\begin{bmatrix} z_{1,t} - \mu_1 \\ z_{2,t} - \mu_2 \\ ... \\ z_{k,t} - \mu_k \end{bmatrix} \begin{bmatrix} (z_{1,t-l} - \mu_1) & (z_{2,t-l} - \mu_2) & ... & (z_{k,t-l} - \mu_k) \end{bmatrix}\right]\\
&= E\left[\begin{bmatrix} (z_{1,t} - \mu_1)(z_{1,t-l} - \mu_1) & (z_{1,t} - \mu_1)(z_{2,t-l} - \mu_2) & ... & (z_{1,t} - \mu_1)(z_{k,t-l} - \mu_k) \\ (z_{2,t} - \mu_2)(z_{1,t-l} - \mu_1) & (z_{2,t} - \mu_2)(z_{2,t-l} - \mu_2) & ... & (z_{2,t} - \mu_2)(z_{k,t-l} - \mu_k) \\ ... & ... & ... & ... \\ (z_{k,t} - \mu_k)(z_{1,t-l} - \mu_1) & (z_{k,t} - \mu_k)(z_{2,t-l} - \mu_2) & ... & (z_{k,t} - \mu_k)(z_{k,t-l} - \mu_k) \end{bmatrix}\right]\\
&= \begin{bmatrix} E[(z_{1,t} - \mu_1)(z_{1,t-l} - \mu_1)] & E[(z_{1,t} - \mu_1)(z_{2,t-l} - \mu_2)] & ... & E[(z_{1,t} - \mu_1)(z_{k,t-l} - \mu_k)] \\ E[(z_{2,t} - \mu_2)(z_{1,t-l} - \mu_1)] & E[(z_{2,t} - \mu_2)(z_{2,t-l} - \mu_2)] & ... & E[(z_{2,t} - \mu_2)(z_{k,t-l} - \mu_k)] \\ ... & ... & ... & ... \\ E[(z_{k,t} - \mu_k)(z_{1,t-l} - \mu_1)] & E[(z_{k,t} - \mu_k)(z_{2,t-l} - \mu_2)] & ... & E[(z_{k,t} - \mu_k)(z_{k,t-l} - \mu_k)] \end{bmatrix} \qquad (1.5)
\end{aligned}
$$

where each $(i,j)th$ element of Γ_l is $Cov[z_{i,t} z_{j,t-l}]$ the covariance between $z_{i,t}$ and $z_{j,t-l}$ denoted $\gamma_{l,i,j}$, thus we have $\Gamma_l = [\gamma_{l,i,j}]$. Note that I use a little bit indexing concencus l, i, j as oppose to l, ij used in the book as I wanted to make clear the indexing hierarchy $lag, row, column$. Also another thing to notice is that here I put the equation tag the same that is equation 1.5 even though the form is a bit different than the book, just for the reader to realize the equivalence of the equation.

For $i = j$, then $\gamma_{l,i,i}$ is the lag l **within-dependence** or **auto-covariance** of dimension (variable) i.

If $l = 0$, Γ_0 is the **contemporaneous covariance** or simply **covariance matrix** of $\vec{z}_t$.

Notice that in the lecture slide the index of lag is k instead of l, thus in the lecture we have Γ_k instead of Γ_l

1.2.2 Linearity

A k-dimensional time series $\vec{z}_t$ is linear if it can be represented as an infinite-order of Moving-Average in the form:

$$
\vec{z}_t = \vec{\mu} + \sum_{s=0}^{\infty} \psi_s \vec{a}_{t-s} \qquad (1.1)
$$

$$
\vec{z}_t = \begin{bmatrix} \mu_1 \\ \mu_2 \\ ... \\ \mu_3 \end{bmatrix} + \sum_{s=0}^{\infty} \begin{bmatrix} \psi_{s,1,1} & \psi_{s,1,2} & ... & \psi_{s,1,k} \\ \psi_{s,2,1} & \psi_{s,2,2} & ... & \psi_{s,2,k} \\ ... & ... & ... & ... \\ \psi_{s,k,1} & \psi_{s,k,2} & ... & \psi_{s,k,k} \end{bmatrix} \begin{bmatrix} a_{1,t-s} \\ a_{2,t-s} \\ ... \\ a_{k,t-s} \end{bmatrix}
$$

where $\vec{\mu}$ is a k-dimensional constant vector, $\psi_0 = I_k$,the $k \times k$ identity matrix, $\psi_s, s > 0$ are $k \times k$ constant matrices, and $\{\vec{a}_t\}$ is a sequence of independent and identically distributed random vectors with mean zero and a positive-definite covariance matrix Σ_a, that is $\vec{a}_t \sim WN(\vec{0}, \Sigma_a)$ where $\vec{0}$ is zero vector of length k.

Where we have:

$$E[\vec{z}_t] = E\left[\begin{bmatrix}\mu_1\\ \mu_2\\ ...\\ \mu_3\end{bmatrix} + \sum_{s=0}^{\infty}\begin{bmatrix}\psi_{s,1,1} & \psi_{s,1,2} & ... & \psi_{s,1,k}\\ \psi_{s,2,1} & \psi_{s,2,2} & ... & \psi_{s,2,k}\\ ... & ... & ... & ...\\ \psi_{s,k,1} & \psi_{s,k,2} & ... & \psi_{s,k,k}\end{bmatrix}\begin{bmatrix}a_{1,t-s}\\ a_{2,t-s}\\ ...\\ a_{k,t-s}\end{bmatrix}\right]$$

$$E[\vec{z}_t] = \begin{bmatrix}\mu_1\\ \mu_2\\ ...\\ \mu_3\end{bmatrix} + \sum_{s=0}^{\infty}\begin{bmatrix}\psi_{s,1,1} & \psi_{s,1,2} & ... & \psi_{s,1,k}\\ \psi_{s,2,1} & \psi_{s,2,2} & ... & \psi_{s,2,k}\\ ... & ... & ... & ...\\ \psi_{s,k,1} & \psi_{s,k,2} & ... & \psi_{s,k,k}\end{bmatrix}\begin{bmatrix}E[a_{1,t-s}]\\ E[a_{2,t-s}]\\ ...\\ E[a_{k,t-s}]\end{bmatrix}$$

$$E[\vec{z}_t] = \begin{bmatrix}\mu_1\\ \mu_2\\ ...\\ \mu_3\end{bmatrix} + \sum_{s=0}^{\infty}\begin{bmatrix}\psi_{s,1,1} & \psi_{s,1,2} & ... & \psi_{s,1,k}\\ \psi_{s,2,1} & \psi_{s,2,2} & ... & \psi_{s,2,k}\\ ... & ... & ... & ...\\ \psi_{s,k,1} & \psi_{s,k,2} & ... & \psi_{s,k,k}\end{bmatrix}\begin{bmatrix}0\\ 0\\ ...\\ 0\end{bmatrix},\ since\ \vec{a}_t \sim WN(\vec{0}, \Sigma_a)$$

$$E[\vec{z}_t] = \begin{bmatrix}\mu_1\\ \mu_2\\ ...\\ \mu_3\end{bmatrix}$$

$$E[\vec{z}_t] = \vec{\mu} \tag{1.3a}$$

The fact that $\vec{a}_t \sim WN(\vec{0}, \Sigma_a)$ has implication:

$$Cov[\vec{a}_i, \vec{a}_j] = \begin{cases}\Sigma_a,\ for\ i = j\\ \Sigma_0,\ a\ k \times k\ matrix\ with\ all\ element\ \Sigma_{0,i,j} = 0,\ i,j = 1,2,...,k\end{cases} \tag{1.3b}$$

Substituting equation 1.1 to equation 1.5b, we have:

$$\begin{aligned}
\Gamma_l =& Cov[\vec{z}_t, \vec{z}_{t-l}] = E[(\vec{z}_t - \vec{\mu})(\vec{z}_{t-l} - \vec{\mu})^T]\\
\Gamma_l =& E[(\vec{\mu} + \sum_{r=0}^{\infty}\psi_r\vec{a}_{t-r} - \vec{\mu})(\vec{\mu} + \sum_{s=0}^{\infty}\psi_s\vec{a}_{(t-l)-s} - \vec{\mu})^T]\\
\Gamma_l =& E[(\sum_{r=0}^{\infty}\psi_r\vec{a}_{t-r})(\sum_{s=0}^{\infty}\psi_s\vec{a}_{(t-l)-s})^T]\\
\Gamma_l =& E[(\sum_{r=0}^{\infty}\psi_r\vec{a}_{t-r})(\sum_{s=0}^{\infty}(\psi_s\vec{a}_{(t-l)-s})^T)]\\
\Gamma_l =& E[(\sum_{r=0}^{\infty}\psi_r\vec{a}_{t-r})(\sum_{s=0}^{\infty}\vec{a}^T_{(t-l)-s}\psi_s^T)]\\
\Gamma_l =& \sum_{r=0}^{\infty}\sum_{s=0}^{\infty}\psi_r E[\vec{a}_{t-r}\vec{a}^T_{(t-l)-s}]\psi_s^T\\
& by\ equation\ 1.3b\ E[\vec{a}_{t-r}\vec{a}^T_{(t-l)-s}]\ has\ values\ only\ when\\
& t - r = (t-l) - s \iff s = r - l \iff r = s + l,\ thus:\\
\Gamma_l =& \sum_{r=l}^{\infty}\psi_r E[\vec{a}_{t-r}\vec{a}^T_{t-r}]\psi^T_{r-l} = \sum_{s=0}^{\infty}\psi_{s+l}E[\vec{a}_{t-(s+l)}\vec{a}^T_{t-(s+l)}]\psi_s^T\\
\Gamma_l =& \sum_{r=l}^{\infty}\psi_r\Sigma_a\psi^T_{r-l} = \sum_{s=0}^{\infty}\psi_{s+l}\Sigma_a\psi_s^T
\end{aligned} \tag{1.3c}$$

As summary we have the equivalence of the book's equation 1.3 below:

$$E[\vec{z}_t] = \vec{\mu}, \quad \Gamma_l = \sum_{r=l}^{\infty} \psi_r \Sigma_a \psi_{r-l}^T = \sum_{s=0}^{\infty} \psi_{s+l} \Sigma_a \psi_s^T \tag{1.3}$$

1.4 SAMPLE CCM

Things to be noticed is that the difference of denominator between the book and the lecture slide for the **contemporaneous covariance** or simply **covariance matrix** estimator $\hat{\Gamma}_0$ and the **between-dependence** or **lag** l **cross-covariance matrix** estimator $\hat{\Gamma}_l$.

In the lecture slide the denominator is T to get the desirable property that for each $k \geq 1$, the k-dimensional sample covariance matrix is non-negative definite while in the book use the unbiased denominator $T - 1$.

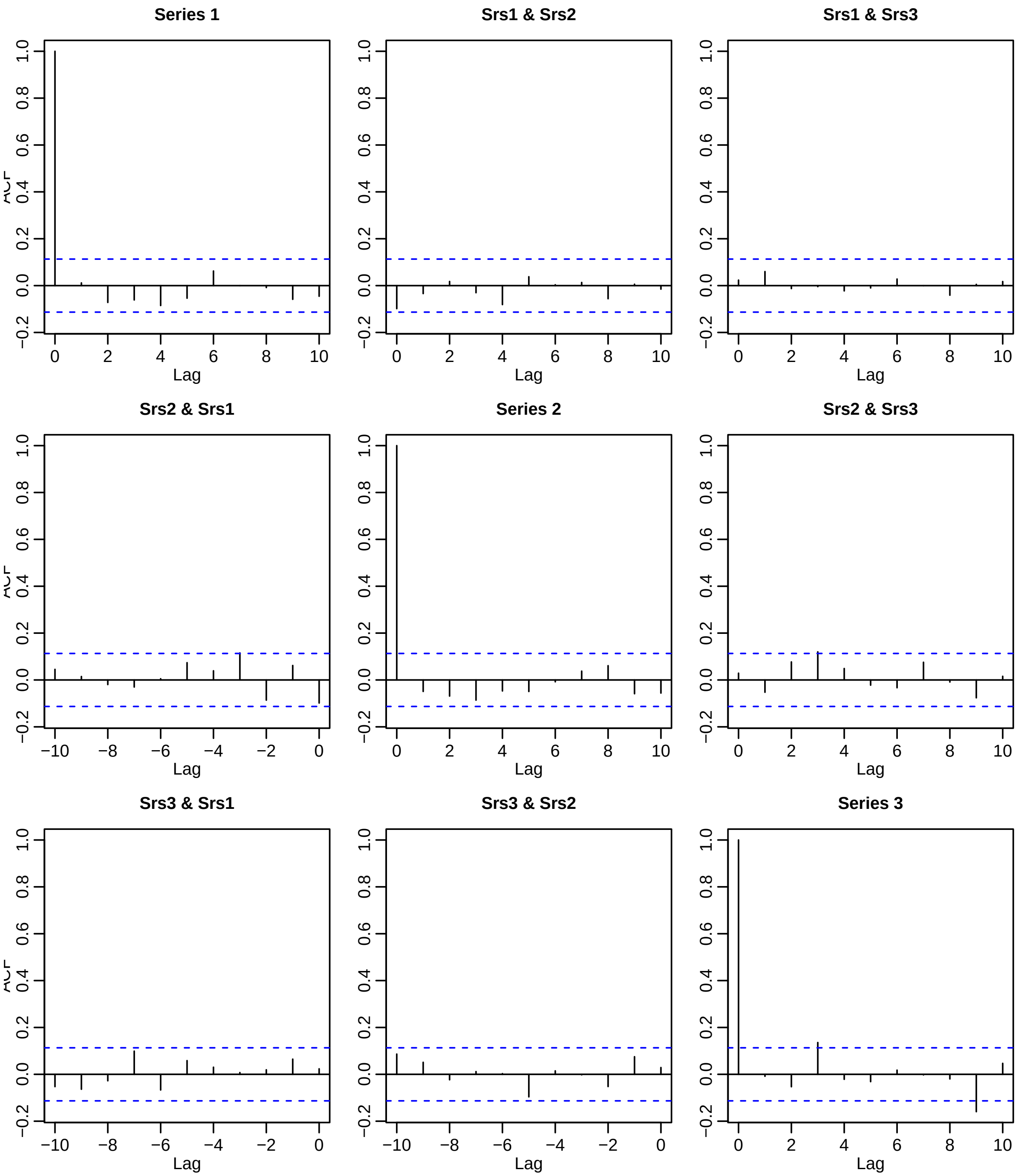
The calculation in lecture slide is the one used in `acf` R command, while the one used in the book is the one used in `ccm` R command under the book `MTS` package in the Book Related Library section. We will have an illustration on the difference by using R code below:

First we create a simulation of multivariate time series:

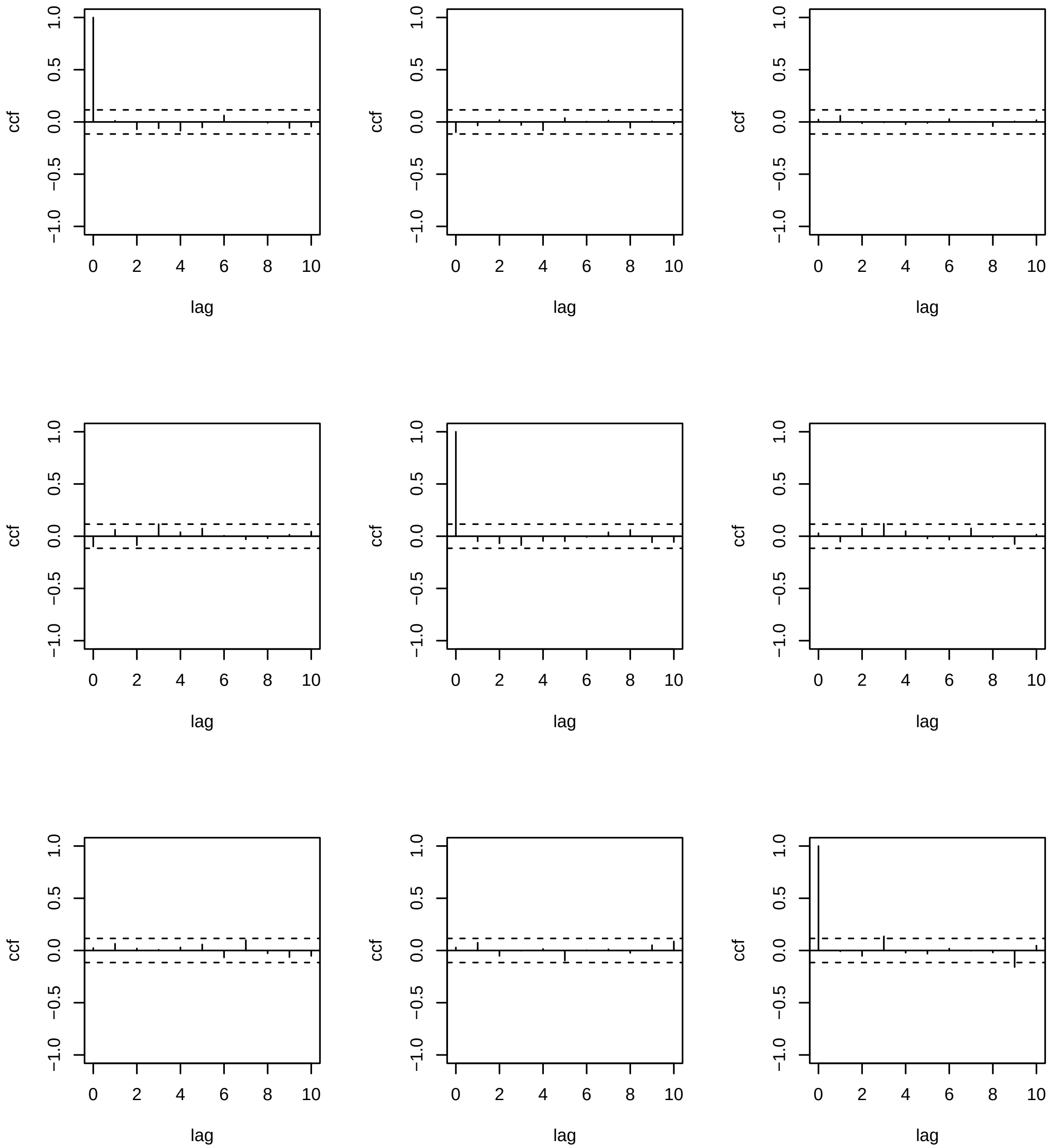
```
sig <- diag( 3 )
set.seed( 42 )
x <- mvtnorm::rmvnorm( 300, rep( 0, 3 ), sig )
```

Then we do the sample cross-correlation matrix calculation:

```
acf.object.test <- acf( x, lag.max = 10 )
```




```r
ccm.object.test <- MTS::ccm( x, lags = 10, level = FALSE )
```



Notice that the **acf** command makes the lag index negative for plots under the diagonal while the **ccm** command keep it as it is.

Below is an example on how to get the sample cross-correlation matrix at a particular lag for each **acf** and **ccm** command output respectively:

```
extract.sccm.from.acf <- function ( acf.object, lag ) {
    return( acf.object$acf[ lag + 1,, ] )
}
```

```r
extract.sccm.from.ccm <- function ( ccm.object, lag ) {
    content <- ccm.object$ccm[ , lag + 1 ]
    return( matrix(
        content,
        nrow = sqrt( length( content ) )
    ) )
}
```

Getting cross-correlation matrix at lag 3:

```r
extract.sccm.from.acf( acf.object.test, 3 )
```

```
##              [,1]         [,2]         [,3]
## [1,] -0.061187503 -0.030266134 -0.004237558
## [2,]  0.115279318 -0.086012154  0.119866959
## [3,]  0.007424076 -0.002100335  0.135626816
```

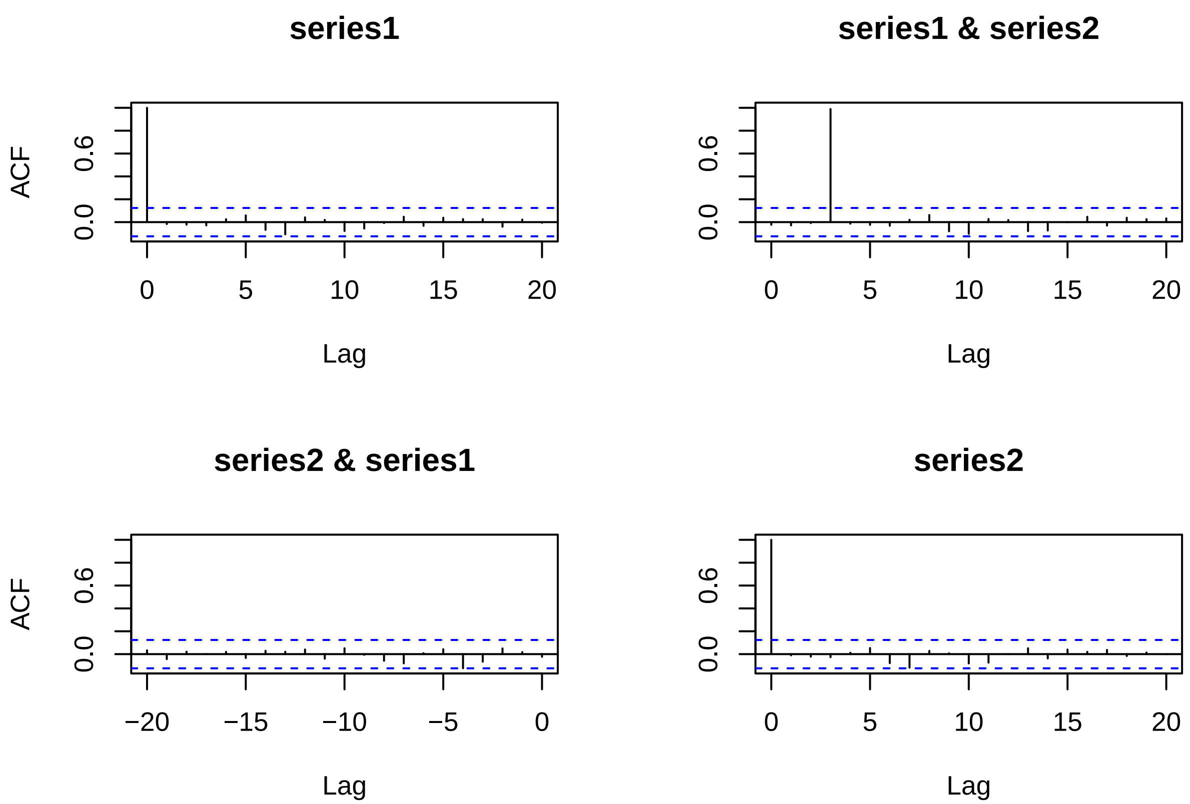
```r
extract.sccm.from.ccm( ccm.object.test, 3 )
```

```
##              [,1]         [,2]         [,3]
## [1,] -0.060983544 -0.030165247 -0.004223433
## [2,]  0.114895053 -0.085725446  0.119467402
## [3,]  0.007399329 -0.002093334  0.135174726
```

As we've explained before, the difference in the denominator choices are reflected on the slightly (since we have quite large sample of data, 300 data points) difference value in the each command's output matrix elements.

The output matrices above are in correspondence position with the plots. To make things clear, let's look at the lecture material for **Module 4 - Lesson 3 - Slide 10 and 11**. We have below slightly edited code to show this point:

```r
set.seed( 42 )
e <- rnorm(300)
series1 <- ts(e[1:250])
series2 <- ts(e[4:253])
acf.object.test2 <- acf(ts.union(series1,series2))
```




```r
extract.sccm.from.acf( acf.object.test2, 3 )
```

```
##             [,1]        [,2]
## [1,] -0.02868473  0.98962742
## [2,] -0.06604173 -0.02655358
```

As we can see, the only significant value at lag 3 is the one on the upper right corner of the matrix as well as the plot.

Another thing with this simulation example is that it is a very good opportunity to discuss mathematical reasoning as why only the lag 3 have the significant value for lag $l > 0$. Recall that `rnorm` command simulate independent normally distributed random variables $e_i \sim IIDN(0,1)$ for $i = 1, 2, ..., 300$ with mean $\mu = 0$ and standard deviation $\sigma = 1$, where we have:

$$\begin{aligned} Cov[e_i, e_j] &= E[(e_i - E[e_i])(e_j - E[e_j])] \\ Cov[e_i, e_j] &= E[(e_i - 0)(e_j - 0)] \\ Cov[e_i, e_j] &= E[e_i e_j] = \begin{cases} E[e_i]E[e_j] = 0 \times 0 = 0, \ if \ i \neq j \\ E[e_i e_i] = E[e_i^2] = Var[e_i] + E[e_i]^2 = 1 + 0^2 = 1, \ if \ i = j \end{cases} \end{aligned} \tag{1.4a}$$

For our simulated data example, we have:

$$series1 = z_{1,t} = \{e_1, e_2, ..., e_{250}\},\ z_{1,1} = e_1, z_{1,2} = e_2, z_{1,3} = e_3, z_{1,4} = e_4, ..., z_{1,250} = e_{250}$$
$$series2 = z_{2,t} = \{e_4, e_5, ..., e_{253}\},\ z_{2,1} = e_4, z_{2,2} = e_5, z_{2,3} = e_6, z_{2,4} = e_7, ..., z_{2,250} = e_{253}$$
$$where\ we\ have\ z_{1,t} = z_{2,t-3} = e_t\ by\ equation\ 1.4a \implies l = 3\ as\ the\ only\ possibility\ for\ Cov[e_i, e_j] \neq 0$$
$$next\ from\ equation\ 1.5\ with\ l = 3\ and\ k = 2,\ we\ have\ the\ theoretical\ cross\text{-}correlation\ matrix:$$
$$\Gamma_3 = Cov[\vec{z}_t, \vec{z}_{t-3}] = \begin{bmatrix} E[(z_{1,t}-\mu_1)(z_{1,t-3}-\mu_1)] & E[(z_{1,t}-\mu_1)(z_{2,t-3}-\mu_2)] \\ E[(z_{2,t}-\mu_2)(z_{1,t-3}-\mu_1)] & E[(z_{2,t}-\mu_2)(z_{2,t-3}-\mu_2)] \end{bmatrix}$$
$$\Gamma_3 = Cov[\vec{z}_t, \vec{z}_{t-3}] = \begin{bmatrix} E[(z_{1,t})(z_{1,t-3})] & E[(z_{1,t})(z_{2,t-3})] \\ E[(z_{2,t})(z_{1,t-3})] & E[(z_{2,t})(z_{2,t-3})] \end{bmatrix},\ since\ e_i \sim IIDN(0,1),\ thus\ \mu_1 = \mu_2 = 0$$
$$\Gamma_3 = Cov[\vec{z}_t, \vec{z}_{t-3}] = \begin{bmatrix} E[e_t e_{t-3}] & E[e_t^2] \\ E[e_{t+3} e_{t-3}] & E[e_{t+3} e_t] \end{bmatrix}$$
$$\Gamma_3 = Cov[\vec{z}_t, \vec{z}_{t-3}] = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix},\ by\ equation\ 1.4a$$

which agrees with the sample cross-correlation matrix.

2.1 INTRODUCTION

The k-dimensional multivariate time series $\vec{z}_t$ follows a VAR model of order p, VAR(p), if

$$\vec{z}_t = \vec{\phi}_0 + \sum_{s=1}^{p} \phi_s \vec{z}_{t-s} + \vec{a}_t \tag{2.1}$$
$$\vec{z}_t = \begin{bmatrix} \phi_{0,1} \\ \phi_{0,2} \\ ... \\ \phi_{0,k} \end{bmatrix} + \sum_{s=1}^{p} \begin{bmatrix} \phi_{s,1,1} & \phi_{s,1,2} & ... & \phi_{s,1,k} \\ \phi_{s,2,1} & \phi_{s,2,2} & ... & \phi_{s,2,k} \\ ... & ... & ... & ... \\ \phi_{s,k,1} & \phi_{s,k,2} & ... & \phi_{s,k,k} \end{bmatrix} \vec{z}_{t-s} + \begin{bmatrix} a_{1,t} \\ a_{2,t} \\ ... \\ a_{k,t} \end{bmatrix}$$

where $\vec{\phi}_0$ is a k-dimensional constant vector and ϕ_s are $k \times k$ matrices for $s > 0$, that is $\phi_s = [\phi_s, i, j],\ s = 1, 2, ..., p\ and\ i, j = 1, 2, ..., k$, with $\phi_p \neq O$ where O is a $k \times k$ matrix which all elements are 0, that is $O_{i,j} = 0,\ i, j = 1, 2, ..., k$, and $\vec{a}_t$ is a sequence of independent and identically distributed (IID) random vectors with mean zero and covariance matrix Σ_a, which is positive-definite, that is $\vec{a}_t \sim WN(\vec{0}, \Sigma_a)$ where $\vec{0}$ is zero vector of length k.

Notice the equivalence of equation 2.1 with the lecture slide:

$$\vec{z}_t = Y_t,\ \ \vec{\phi}_0 = c,\ \ \vec{a}_t = \epsilon_t,\ \ \phi_s = \Pi_s\ for\ s = 1, 2, ..., p$$

I've also adopt the indexing nomenclature $s, i, j = lag\ index,\ row\ index,\ column\ index$ in $\phi_s = [\phi_s, i, j],\ s = 1, 2, ..., p\ and\ i, j = 1, 2, ..., k$ instead of π^s_{ij} used in lecture slide to avoid the perception of lag index s being a power of the matrix element, the adopted indexing nomenclature also slightly different from the one used in the book $\phi_s = [\phi_s, ij]$ to further clarify separation between row and column index i, j.

2.4 VAR(p) MODELS

First, let's transform equation 2.1 to its lag operator notation equivalence:

$$
\begin{aligned}
\vec{z}_t &= \vec{\phi}_0 + \sum_{s=1}^{p} \phi_s \vec{z}_{t-s} + \vec{a}_t \\
\vec{z}_t - \sum_{s=1}^{p} \phi_s \vec{z}_{t-s} &= \vec{\phi}_0 + \vec{a}_t \\
I_k \vec{z}_t - \sum_{s=1}^{p} \phi_s B^s \vec{z}_t &= \vec{\phi}_0 + \vec{a}_t \\
(I_k - \sum_{s=1}^{p} \phi_s B^s) \vec{z}_t &= \vec{\phi}_0 + \vec{a}_t \qquad (2.21a) \\
\phi(B) \vec{z}_t &= \vec{\phi}_0 + \vec{a}_t \qquad (2.21)
\end{aligned}
$$

If we take expectation of equation 2.21a, we have:

$$
\begin{aligned}
E[(I_k - \sum_{s=1}^{p} \phi_s B^s) \vec{z}_t] &= E[\vec{\phi}_0 + \vec{a}_t] \\
(I_k - \sum_{s=1}^{p} \phi_s B^s) E[\vec{z}_t] &= \vec{\phi}_0 + E[\vec{a}_t] \\
(I_k - \sum_{s=1}^{p} \phi_s B^s) \vec{\mu} &= \vec{\phi}_0 + \vec{0} \\
(I_k - \sum_{s=1}^{p} \phi_s B^s) \vec{\mu} &= \vec{\phi}_0 \\
\phi(B) \vec{\mu} &= \vec{\phi}_0 \qquad (2.21b)
\end{aligned}
$$

Now substituting equation 2.21b to equation 2.21:

$$
\begin{aligned}
\phi(B) \vec{z}_t &= \vec{\phi}_0 + \vec{a}_t \\
\phi(B) \vec{z}_t &= \phi(B) \vec{\mu} + \vec{a}_t \\
\phi(B) \vec{z}_t - \phi(B) \vec{\mu} &= \vec{a}_t \\
\phi(B)(\vec{z}_t - \vec{\mu}) &= \vec{a}_t \\
\phi(B) \vec{\tilde{z}}_t &= \vec{a}_t \qquad (2.22)
\end{aligned}
$$

Where

$$
\vec{\tilde{z}}_t = \begin{bmatrix} \tilde{z}_{1,t} \\ \tilde{z}_{2,t} \\ \dots \\ \tilde{z}_{k,t} \end{bmatrix} = \begin{bmatrix} z_{1,t} - \mu_1 \\ z_{2,t} - \mu_2 \\ \dots \\ z_{k,t} - \mu_k \end{bmatrix} = \vec{z}_t - \vec{\mu} \qquad (2.22a)
$$

2.4.1 A VAR(1) Representation

Define a pk-dimensional time series vectors:

$$Z_t = (\vec{\tilde{z}}_t^{\,T},\ \vec{\tilde{z}}_{t-1}^{\,T}, ...,\ \vec{\tilde{z}}_{t-(p-1)}^{\,T})^T,\quad \vec{b}_t = (\vec{a}_t^{\,T}, \vec{0}^{\,T})^{\,T}$$

where each $\vec{\tilde{z}}_s,\ s = t,\ t-1,\ ...,\ t-(p-1)$ defined in equation 2.22a, and $\vec{0}$ being a $k(p-1)$-dimensional zero vector, then the VAR(p) model in equation 2.22 can be written as:

$$\begin{bmatrix} \tilde{z}_{1,t} \\ \tilde{z}_{2,t} \\ ... \\ \tilde{z}_{k,t} \\ \tilde{z}_{1,t-1} \\ \tilde{z}_{2,t-1} \\ ... \\ \tilde{z}_{k,t-1} \\ ... \\ \tilde{z}_{1,t-(p-2)} \\ \tilde{z}_{2,t-(p-2)} \\ ... \\ \tilde{z}_{k,t-(p-2)} \\ \tilde{z}_{1,t-(p-1)} \\ \tilde{z}_{2,t-(p-1)} \\ ... \\ \tilde{z}_{k,t-(p-1)} \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & ... & \phi_{p-1} & \phi_p \\ I_k & 0_k & ... & 0_k & 0_k \\ 0_k & I_k & ... & 0_k & 0_k \\ ... & ... & ... & ... & ... \\ O_k & 0_k & ... & I_k & 0_k \end{bmatrix} \begin{bmatrix} \tilde{z}_{1,t-1} \\ \tilde{z}_{2,t-1} \\ ... \\ \tilde{z}_{k,t-1} \\ \tilde{z}_{1,t-2} \\ \tilde{z}_{2,t-2} \\ ... \\ \tilde{z}_{k,t-2} \\ ... \\ \tilde{z}_{1,t-(p-1)} \\ \tilde{z}_{2,t-(p-1)} \\ ... \\ \tilde{z}_{k,t-(p-1)} \\ \tilde{z}_{1,t-p} \\ \tilde{z}_{2,t-p} \\ ... \\ \tilde{z}_{k,t-p} \end{bmatrix} + \begin{bmatrix} a_{1,t} \\ a_{2,t} \\ ... \\ a_{k,t} \\ 0 \\ 0 \\ ... \\ 0 \\ ... \\ 0 \\ 0 \\ ... \\ 0 \\ 0 \\ 0 \\ ... \\ 0 \end{bmatrix}$$

$$Z_t = \begin{bmatrix} \phi_1 & \phi_2 & ... & \phi_{p-1} & \phi_p \\ I_k & 0_k & ... & 0_k & 0_k \\ 0_k & I_k & ... & 0_k & 0_k \\ ... & ... & ... & ... & ... \\ O_k & 0_k & ... & I_k & 0_k \end{bmatrix} Z_{t-1} + \vec{b}_t$$

$$Z_t = \Phi Z_{t-1} + \vec{b}_t \tag{2.23}$$

where it is understood that I_k and 0_k are the $k \times k$ identity and zero matrix, respectively and ϕ_s defined in equation 2.1 for $s = 1, 2, ..., p$. The $pk \times pk$ matrix Φ in equation 2.23 is called the **companion matrix** of the matrix polynomial $\phi(B)$ in equation 2.22.

VAR(p) is stable if magnitude of the roots of $\phi(B)$ are all > 1 or the eigenvalues of the companion matrix Φ all have magnitude < 1.

Example 2.1

Consider the bi-variate VAR(1) model:

$$\begin{bmatrix} z_{1,t} \\ z_{2,t} \end{bmatrix} = \begin{bmatrix} 5 \\ 3 \end{bmatrix} + \begin{bmatrix} 0.2 & 0.3 \\ -0.6 & 1.1 \end{bmatrix} \begin{bmatrix} z_{1,t-1} \\ z_{2,t-1} \end{bmatrix} \begin{bmatrix} a_{1,t} \\ a_{2,t} \end{bmatrix},\quad \Sigma_a = \begin{bmatrix} 1.0 & 0.8 \\ 0.8 & 2.0 \end{bmatrix}$$

we have $\phi_1 = \begin{bmatrix} 0.2 & 0.3 \\ -0.6 & 1.1 \end{bmatrix}$ and since $p = 1$ and $k = 2$ our companion matrix Φ will have dimension $pk \times pk = (1)(2) \times (1)(2) = 2 \times 2$ thus in this case, $\Phi = \phi_1$.

To check if for stability we need to check the root of $det(I_2 - \phi_1 B)$ or the eigenvalues of the companion matrix Φ.

The first method:

$$\begin{aligned}
det(I_2 - \phi_1 B) &= 0 \\
det(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.2 & 0.3 \\ -0.6 & 1.1 \end{bmatrix} B) &= 0 \\
det(\begin{bmatrix} 1 - 0.2B & -0.3B \\ 0.6B & 1 - 1.1B \end{bmatrix}) &= 0 \\
1 - 1.1B - 0.2B + 0.22B^2 + 0.18B^2 &= 0 \\
1 - 1.3B + 0.4B^2 &= 0 \\
(1 - 0.5B)(1 - 0.8B) &= 0 \\
B_1 = 2,\ B_2 &= 1.25
\end{aligned}$$

since all the root's magnitude of $\phi(B)$ are > 1, the process is stable.

The second method

$$\begin{aligned}
det(\lambda I_2 - \Phi) &= 0 \\
det(\lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 0.2 & 0.3 \\ -0.6 & 1.1 \end{bmatrix}) &= 0 \\
det(\begin{bmatrix} \lambda - 0.2 & -0.3 \\ 0.6 & \lambda - 1.1 \end{bmatrix}) &= 0 \\
\lambda^2 - 1.3\lambda + 0.4 &= 0 \\
\lambda_1 = 0.5,\ \lambda_2 &= 0.8
\end{aligned}$$

both the eigenvalues of the companion matrix Φ have magnitude < 1, thus the process is stable.