



**ATHENS UNIVERSITY
OF ECONOMICS AND BUSINESS**

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

«Testing the Goodness of Fit of a Vector
Autoregressive Model»

By

Maria Andrea Kleideri

A THESIS

Submitted to the Department of Statistics
of the Athens University of Economics and Business
in partial fulfilment of the requirements for
the degree of Master of Science in Statistics

Athens, Greece
October 2008







ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ ΜΕΤΑΠΤΥΧΙΑΚΟ ΠΡΟΓΡΑΜΜΑ

«Έλεγχος Καλής Προσαρμογής ενός Διανυσματικού
Αυτοπαλίνδρομου Μοντέλου»

Μαρία Ανδρέα Κλειδέρη

ΔΙΑΤΡΙΒΗ

Που υποβλήθηκε στο Τμήμα Στατιστικής
του Οικονομικού Πανεπιστημίου Αθηνών
ως μέρος των απαιτήσεων για την απόκτηση
Μεταπτυχιακού Διπλώματος Ειδίκευσης στη Στατιστική

Αθήνα
Οκτώβριος 2008



ACKNOWLEDGEMENTS

Θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα της διπλωματικής μου εργασίας Επίκουρο Καθηγητή κ. Ευάγγελο Ιωαννίδη για την καθοδήγηση, την συμβολή και το ενδιαφέρον που επέδειξε καθ' όλη την διάρκεια της εκπόνησης της εργασίας αυτής.



ABSTRACT

Maria Kleideri

Testing the Goodness of Fit of a Vector Autoregressive Model

October 2008

In this thesis we study by simulation a Portmanteau test statistic for the goodness of fit of a vector autoregressive model (VAR). Assuming we have a multivariate process, we fit a VAR model of order p on the data and we check the whiteness of the fitted residuals of the model by using this test. According to the theory already developed on the test, this statistic has an asymptotical chi-squared distribution. The issue addressed in this thesis is to what extent this chi-squared distribution holds in finite samples situations. We, therefore, simulate the distribution of the statistic for a series of models, for various sample sizes and orders fitted and by examining the percentage of rejecting each model at different lags, we may compare this simulated distribution to the one predicted by asymptotic theory. From these results we can have a better estimation of whether we may use the statistic. We conclude that the Portmanteau test can be really effective at large samples of the process and at equivalently large lags as the order of the fitted model increases. However, the asymptotic theory may be a very bad approximation to the true distribution when the order fitted is large and the model is close to non-stationarity.



ΠΕΡΙΛΗΨΗ

Μαρία Κλειδέρη

Έλεγχος Καλής Προσαρμογής ενός Διανυσματικού Αυτοπαλίνδρομου Μοντέλου

Οκτώβριος 2008

Σε αυτήν την εργασία μελετάμε με προσομοίωση έναν στατιστικό έλεγχο Portmanteau για την καλή προσαρμογή ενός διανυσματικού αυτοπαλίνδρομου μοντέλου. Έστω ότι έχουμε μία πολυμεταβλητή χρονοσειρά, εφαρμόζουμε ένα αυτοπαλίνδρομο μοντέλο τάξης p στα δεδομένα και ελέγχουμε αν τα προσαρμοσμένα κατάλοιπα του μοντέλου είναι λευκός θόρυβος χρησιμοποιώντας αυτόν τον έλεγχο. Σύμφωνα με την θεωρία που έχει ήδη αναπτυχθεί γύρω από τον έλεγχο, αυτό το στατιστικό ακολουθεί ασυμπτωτικά χ^2 κατανομή. Το θέμα που εξετάζεται σε αυτήν την εργασία είναι σε ποιο βαθμό αυτή η κατανομή ισχύει σε περιπτώσεις πεπερασμένων δειγμάτων. Για τον σκοπό αυτό, προσομοιώνουμε την κατανομή του στατιστικού για μια σειρά από μοντέλα, για διάφορα δείγματα και διάφορες τάξεις και εξετάζοντας το ποσοστό απόρριψης του κάθε μοντέλου σε διαφορετικές χρονικές υστερήσεις, μπορούμε να συγκρίνουμε την προσομοιωμένη κατανομή με την προβλεπόμενη από την ασυμπτωτική θεωρία. Από τα αποτελέσματα της μελέτης, μπορούμε να έχουμε μία καλύτερη εκτίμηση του πότε μπορεί να χρησιμοποιηθεί ο έλεγχος αυτός. Συμπεραίνουμε ότι το Portmanteau μπορεί να είναι ιδιαίτερα αποτελεσματικό για μεγάλα μεγέθη χρονοσειράς και ανάλογα μεγάλες χρονικές υστερήσεις όσο αυξάνει η τάξη του προσαρμοσμένου μοντέλου. Ωστόσο, η ασυμπτωτική θεωρία μπορεί να είναι μία πολύ κακή προσέγγιση της πραγματικής κατανομής όταν η τάξη του μοντέλου είναι μεγάλη και το μοντέλο τείνει να μην είναι στάσιμο.





TABLE OF CONTENTS

	Page
Acknowledgements	I
Abstract	II
Περίληψη	III
Chapter 1: Introduction	1-2
Chapter 2: Basics of Multivariate Time Series	3-8
2.1 Introduction	3
2.2 Mean Vector and Covariance Matrix	3
2.3 Spectral Characteristics for a Stationary Vector Process	5
2.4 Linear Model Representation of a Stationary Vector Process	7
Chapter 3: Introduction to Vector Autoregressive Models	9-14
3.1 Introduction	9
3.2 Definition of a Vector Autoregressive Model	9
3.3 Yule – Walker Equations for Covariance Matrices of a VAR Process	10
3.4 Least Squares Estimations of VAR(p) Models	11
Chapter 4: Portmanteau Test	15-21
4.1 Introduction	15
4.2 The Portmanteau Test for White Noise Processes	15
4.3 The Portmanteau Test for VAR(p) Models	18
4.4 Comments on the Use of Portmanteau Test	19
Chapter 5: Simulation Study	21-34
5.1 Introduction	21
5.2 Generation of Models	21
5.3 Design of Simulation	22
5.4 Applying the Portmanteau Test on Specific Models	24
5.4.1 White Noise ($p=0$)	24
5.4.2 Bivariate Models of Order $p=1$	29
5.4.3 A Bivariate Model of Order $p=2$ (Model E)	32
Chapter 6: Conclusion	35-36
Appendix 1: Simulation Figures and Results	37
Appendix 2: Algorithms in R	47
References	51





LIST OF TABLES

Table	Page
Table 5.1: Percentages of Rejection of Whiteness for White Noise ($n=100$)	26
Table 5.2: Percentages of Rejection of Whiteness for White Noise ($n=200$)	27
Table 5.3: Percentages of Rejection of Whiteness for White Noise ($n=500$)	28
Table 5.4: Percentages of Rejection of Whiteness for VAR(1) Models	30
Table 5.5: Percentages of Rejection of Whiteness for VAR(2) Model E	33
Table A.1: Percentages of Rejection of Whiteness for Model A	38
Table A.2: Percentages of Rejection of Whiteness for Model B	40
Table A.3: Percentages of Rejection of Whiteness for Model C	42
Table A.4: Percentages of Rejection of Whiteness for Model D	44





LIST OF FIGURES

Figure	Page
Fig 5.1: A sample of the bivariate White Noise Process	25
Fig 5.2: Percentages of Rejection of Whiteness for White Noise ($n=100$)	26
Fig 5.3: Percentages of Rejection of Whiteness for White Noise ($n=200$)	27
Fig 5.4: Percentages of Rejection of Whiteness for White Noise ($n=500$)	28
Fig A.1: The spectra and the coherency of the bivariate VAR(1) Model A	37
Fig A.2: A sample of the bivariate VAR(1) Model A	37
Fig A.3: The spectra and the coherency of the bivariate VAR(1) Model B	39
Fig A.4: A sample of the bivariate VAR(1) Model B	39
Fig A.5: The spectra and the coherency of the bivariate VAR(1) Model C	41
Fig A.6: A sample of the bivariate VAR(1) Model C	41
Fig A.7: The spectra and the coherency of the bivariate VAR(1) Model D	43
Fig A.8: A sample of the bivariate VAR(1) Model D	43
Fig A.9: The spectra and the coherency of the bivariate VAR(2) Model E	45
Fig A.10: A sample of the bivariate VAR(2) Model E	45





CHAPTER 1

Introduction

The vector autoregressive models (VAR) are commonly used to describe a multivariate process. In this thesis we will study these models and in particular examine by simulation a test for the goodness of fit for a VAR model on a multivariate time series. There are many tests suggested to check the adequacy of a fitted model. One of them is a Portmanteau test statistic.

Let $X_t = \Phi_1 X_{t-1} + \dots + \Phi_p X_{t-p} + \varepsilon_t$ be a VAR model of order p where $\varepsilon_t \sim \text{WN}(\mathbf{0}, \Sigma)$. Let also $\hat{\varepsilon}_t := X_t - \hat{\Phi}_1 X_{t-1} - \dots - \hat{\Phi}_p X_{t-p}$ be the estimated residuals of the above fitted model where $\hat{\Phi}_i$ $i=1, \dots, p$ are the least squares estimators of the parameters. If the initial model is a VAR(p), then $\hat{\varepsilon}_t$ should be close to ε_t , that is behave as white noise process.

To examine the whiteness of the fitted residuals this Portmanteau test uses a statistic based on the estimated auto-covariances and cross-covariances of $\hat{\varepsilon}_t$ for lags $u = 1$ to some h . Note that under the null hypothesis of whiteness the corresponding true quantities are all equal to zero. It has been proved that this statistic follows a chi-squared distribution (Hosking (1980) and Ahn (1988)) and therefore we can test the whiteness by comparing the value of the statistic to the corresponding critical value. If the hypothesis is rejected we conclude that the fitted model is not appropriate for the given time series and a model of different order is needed.

In the literature (see Chitturi (1974), Hosking (1980), Ahn (1988), Lütkepohl(1993)) it has been noted that there are finite sample situations where the chi-squared approximation to the distribution of the test statistic might not be reliable. For example, a small sample of the time series or small values of lags may result to misleadingly regarding the asymptotic distributional properties.



Moreover, in this thesis we will particularly study extreme cases of models that are close to the border of non-stationarity. The stationary condition states that all the roots of the polynomial $\det\{\Phi(z)\} = 0$ must be greater than one in absolute value, where $\Phi(z)$ is the matrix of coefficients in the VAR model. Any departures can affect the accuracy of the test. The processes used in the thesis are generated from bivariate autoregressive models that have been constructed based on their spectral characteristics and by choosing roots that are close enough to the unit circle in order to check the adequacy of the asymptotic distribution for such extreme cases.

In chapter 2, we present the basic properties and the spectral characteristics of the multivariate processes. In chapter 3, we introduce the vector autoregressive time series along with their properties. We, also, explain the least squares method that will be used to estimate the fitted residuals.

In chapter 4, we study the Portmanteau test statistic. Apart from its definition and its distributional properties, we make a further analysis on what it is expected according to former articles and studies on the particular statistic.

In chapter 5, we present the simulation study made on six different models. At first, we explain the generation of the initial models and the procedure followed at the study. We present the results for each model and we make the appropriate judgments regarding the use and the accuracy of the Portmanteau test.



CHAPTER 2

Basics of Multivariate Time Series

2.1 Introduction

In this chapter we introduce the stationary multivariate time series by giving the basic definitions and properties and we also present the linear representation of a stationary process according to Wold's theorem.

2.2 Mean Vector and Covariance Matrix

Consider we have more than one time series to study. Instead of examining each one separately we may analyse them all together as a vector of time series. In that way we can also check the possible dependence between the time series and how the one affects the other.

Let $\mathbf{X}_t = (X_{t1}, X_{t2}, \dots, X_{tm})'$ be a vector of m time series observed at times t . The mean and covariance of each X_{ti} are $\mu_{ti} := E[X_{ti}]$ and $\gamma_{ij}(t+u, t) := E[(X_{t+u,i} - \mu_{ti}) \cdot (X_{tj} - \mu_{tj})]$ for $i, j = 1, \dots, m$ and $u = 0, \pm 1, \pm 2, \dots$ accordingly. Then, the *mean* vector of $\{\mathbf{X}_t\}$ is defined as $\boldsymbol{\mu}_t := E[\mathbf{X}_t] = (\mu_{t1}, \mu_{t2}, \dots, \mu_{tm})'$ and the *covariance* matrices

$$\Gamma(t+u, t) := \begin{bmatrix} \gamma_{11}(t+u, t) & \cdot & \cdot & \cdot & \gamma_{1m}(t+u, t) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \gamma_{m1}(t+u, t) & \cdot & \cdot & \cdot & \gamma_{mm}(t+u, t) \end{bmatrix} \quad \text{where } \gamma_{ij}(t+u, t) := \text{cov}(X_{t+u,i}, X_{tj}).$$

The matrix can be also expressed as $\Gamma(t+u, t) := E[(\mathbf{X}_{t+u} - \boldsymbol{\mu}_{t+u}) \cdot (\mathbf{X}_t - \boldsymbol{\mu}_t)']$.

Definition 2.2.1

The multivariate time series $\{\mathbf{X}_t\}$ is called (*weakly*) *stationary* if (i) $\boldsymbol{\mu}_X(t)$ is independent of t and (ii) $\Gamma_X(t+u, t)$ is independent of t for each u .



In the stationary case the mean vector and the covariance matrices can be written

$$\boldsymbol{\mu} := E[\mathbf{X}_t] = (\mu_1, \mu_2, \dots, \mu_m)' \quad \text{and} \quad \Gamma(u) := E[(\mathbf{X}_{t+u} - \boldsymbol{\mu}) \cdot (\mathbf{X}_t - \boldsymbol{\mu})'] \Rightarrow$$

$$\Gamma(u) = \begin{bmatrix} \gamma_{11}(u) & \cdot & \cdot & \cdot & \gamma_{1m}(u) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \gamma_{m1}(u) & \cdot & \cdot & \cdot & \gamma_{mm}(u) \end{bmatrix} \quad \text{at lag } u.$$

Since $\{X_t\}$ is stationary, then for each i , $\{X_{it}\}$ is also stationary with auto-covariance function $\gamma_{ii}(u)$. The function $\gamma_{ij}(\cdot)$, $i \neq j$, is called the *cross-covariance* function of two series $\{X_{it}\}$, $\{X_{jt}\}$ and in general is not equal to $\gamma_{ji}(\cdot)$.

The covariance matrix $\Gamma(u)$ has the following properties:

- (i) $\Gamma(u) = \Gamma'(-u)$
- (ii) $|\gamma_{ij}(u)| \leq (\gamma_{ii}(0) \cdot \gamma_{jj}(0))^{1/2}$, $i, j = 1, \dots, m$
- (iii) $\gamma_{ii}(\cdot)$ is an auto-covariance function, $i = 1, \dots, m$
- (iv) $\sum_{j,k=1}^n a_j' \Gamma(j-k) a_k \geq 0$ for all $n \in \{1, 2, \dots\}$ and $a_1, \dots, a_n \in \mathbb{R}^m$.

We may also define the correlation matrix function $R(\cdot)$ as follows:

$$R(u) := \begin{bmatrix} \rho_{11}(u) & \cdot & \cdot & \cdot & \rho_{1m}(u) \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot \\ \rho_{m1}(u) & \cdot & \cdot & \cdot & \rho_{mm}(u) \end{bmatrix} \quad \text{where } \rho_{ij}(u) = \gamma_{ij}(u) / [\gamma_{ii}(0)\gamma_{jj}(0)]^{1/2}.$$

The correlation matrix has corresponding properties to the covariance matrix and additionally we have $\rho_{ii}(0) = 1$ for all $i = 1, \dots, m$.

Definition 2.2.2

The m -variate time series $\{\mathbf{Z}_t\}$ is called *white noise with mean $\mathbf{0}$ and covariance matrix $\boldsymbol{\Sigma}$* if $\{\mathbf{Z}_t\}$ is stationary with mean vector $\mathbf{0}$ and covariance matrix function $\Gamma(u) = \{ \boldsymbol{\Sigma} \text{ if } u = 0 \text{ and } 0, \text{ otherwise} \}$ and is written $\{\mathbf{Z}_t\} \sim \text{WN}(\mathbf{0}, \boldsymbol{\Sigma})$.



Definition 2.2.3

The m -variate time series $\{\mathbf{Z}_t\}$ is called *iid noise with mean $\mathbf{0}$ and covariance matrix Σ* if the random vectors $\{\mathbf{Z}_t\}$ are independent and identically distributed with mean $\mathbf{0}$ and covariance matrix Σ and is written $\{\mathbf{Z}_t\} \sim \text{iid}(\mathbf{0}, \Sigma)$.

The mean vector and covariance matrix of a multivariate time series can be estimated as follows. An unbiased estimator of the mean vector $\boldsymbol{\mu}$ of an m -variate time series $\{\mathbf{X}_t\}$ may be the vector of sample means based on the observations $\mathbf{X}_1, \dots, \mathbf{X}_n$, that is $\bar{\mathbf{X}}_n = \frac{1}{n} \sum_{i=1}^n \mathbf{X}_i$. The covariance matrix $\Gamma(u)$ at lag u can be estimated by the sample covariance matrix as:

$$\hat{\Gamma}(u) = \begin{cases} n^{-1} \sum_{t=1}^{n-u} (\mathbf{X}_{t-u} - \bar{\mathbf{X}}_n)(\mathbf{X}_t - \bar{\mathbf{X}}_n)' & , \text{ for } 0 \leq u \leq n- \\ \hat{\Gamma}'(-u) & , \text{ for } -n+1 \leq u < 0 \end{cases}$$

2.3 Spectral Characteristics for a Stationary Vector ProcessDefinition 2.3.1

The *spectral density* matrix of a stationary process $\{\mathbf{X}_t\}$ is:

$$f(\lambda) = \frac{1}{2\pi} g(e^{-i\lambda}) = \frac{1}{2\pi} \sum_{u=-\infty}^{\infty} \Gamma(u) e^{-iu\lambda}, \quad -\pi \leq \lambda < \pi, \text{ where } g(\cdot) \text{ is the covariance -}$$

generating function defined as: $g(z) = \sum_{u=-\infty}^{\infty} \Gamma(u) z^u$ provided that $\sum_{u=-\infty}^{\infty} |\gamma_{ij}(u)| < \infty$.

The (i, j) -th element of $f(\lambda)$ is $f_{ij}(\lambda) = \frac{1}{2\pi} \sum_{u=-\infty}^{\infty} \gamma_{ij}(u) e^{-iu\lambda}$, $i, j = 1, \dots, m$.

For $i=j$, the $f_{ii}(\lambda)$ is the auto-spectral density function of the time series X_{it} . For $i \neq j$, the $f_{ij}(\lambda)$ is the cross-spectral density function of X_{it} and X_{jt} .

The function $f(\lambda)$ has the following properties:

- (i) $f_{ii}(\lambda)$ is real-valued and non-negative
- (ii) $f_{ij}(\lambda)$ is complex-valued, since $\gamma_{ij}(u) \neq \gamma_{ij}(-u)$



- (iii) $f_{ij}(\lambda) = f_{ji}(-\lambda)$
- (iv) $f(\lambda)$ is Hermitian matrix, that is $f(\lambda) = f(-\lambda)'$
- (v) $f(\lambda)$ is non-negative matrix, that is $\mathbf{b}' f(\lambda) \mathbf{b} \geq 0$ for any m -dimensional vector \mathbf{b} .

We may notice that $\Gamma(u) = \int_{-\pi}^{\pi} e^{iu\lambda} f(\lambda) d\lambda$ for $u = 0, \pm 1, \pm 2, \dots$, so that

$$\gamma_{ij}(u) = \int_{-\pi}^{\pi} e^{iu\lambda} f_{ij}(\lambda) d\lambda.$$

The time series $\{X_t\}$ can be expressed by the following spectral representation:

$$X_t = \int_{-\pi}^{\pi} e^{-i\lambda t} d\mathbf{Z}(\lambda) \text{ where } \mathbf{Z}(\lambda) = (Z_1(\lambda), \dots, Z_m(\lambda))' \text{ is an } m\text{-dimensional complex-}$$

valued continuous-parameter process defined on the continuous interval $[-\pi, \pi]$ with

the property that $E[d\mathbf{Z}(\lambda_1)d\bar{\mathbf{Z}}(\lambda_2)'] = 0$ if $\lambda_1 \neq \lambda_2$ while $E[d\mathbf{Z}(\lambda)d\bar{\mathbf{Z}}(\lambda)'] = f(\lambda)d\lambda$.

That is, $f(\lambda)d\lambda$ represents the covariance matrix of $d\mathbf{Z}(\lambda)$.

Definition 2.3.2

The (*squared*) *coherency* spectrum of a pair of series X_{it} and X_{jt} is defined as

$$K_{ij}^2(\lambda) = |f_{ij}(\lambda)|^2 / \{f_{ii}(\lambda)f_{jj}(\lambda)\}.$$

The coherency $K_{ij}(\lambda)$ at frequency λ can be interpreted as the correlation coefficient between the random components $dZ_i(\lambda)$ and $dZ_j(\lambda)$ in the spectral representations of X_{it} and X_{jt} . We could say that $K_{ij}(\lambda)$ measures the extent to which the two processes X_{it} and X_{jt} are linearly related in terms of the degree of linear association of their random components at different frequencies λ .

The spectral matrix and the coherency of a process can be used to construct a multivariate time series model VAR(p) as we will see in a following chapter.



2.4 Linear Model Representation of a Stationary Vector Process

We will present a linear model representation of an m -variate stationary process based on a generalization of Wold's Theorem for multivariate processes. Firstly, we shall give some useful definitions.

Definition 2.4.1

A doubly infinite sequence of real numbers $\{a_i\}$, $i=0,\pm 1,\dots$ is *absolutely summable* if $\lim_{n \rightarrow \infty} \sum_{i=-n}^n |a_i|$ exists and is finite.

A sequence of $m \times m$ matrices $\{\mathbf{A}_i\}$, $i=0,\pm 1,\dots$ is *absolutely summable* if the components of each matrix are absolutely summable.

Definition 2.4.2

The m -variate time series $\{\mathbf{X}_t\}$ is a *linear process* if it has the following representation: $\mathbf{X}_t = \sum_{j=-\infty}^{\infty} \Psi_j \mathbf{Z}_{t-j}$, $\{\mathbf{Z}_t\} \sim \text{WN}(\mathbf{0}, \Sigma)$ where $\{\Psi_j\}$ is a sequence of $m \times m$ matrices whose components are absolutely summable.

The above form is also mentioned as *linear filter* relating the processes $\{\mathbf{X}_t\}$ and $\{\mathbf{Z}_t\}$.

Proposition 2.4.3

The linear process is stationary with mean $\mathbf{0}$ and covariance function $\Gamma(u) = \sum_{j=-\infty}^{\infty} \Psi_{j+u} \Sigma \Psi_j'$, $u = 0, \pm 1, \pm 2, \dots$

Proposition 2.4.4

A linear filter is *causal* when the matrices $\Psi_j = 0$ for $j < 0$, so that $\mathbf{X}_t = \sum_{j=-\infty}^{\infty} \Psi_j \mathbf{Z}_{t-j}$ is expressible in terms of only present and past values of the input process $\{\mathbf{Z}_t\}$.



Theorem 2.4.5 (Wold's Theorem)

Let $\{X_t\}$ be a purely non deterministic stationary process with (constant) mean vector μ . Then $X_t - \mu$ can always be represented as the output of a causal linear filter with white noise input.

A proof of the above theorem can be found in Brockwell and Davis (1987).

According to 2.4.5, the m -variate X_t can be written as a representation of an infinite vector moving average (MA) process, that is

$$X_t = \mu + \sum_{j=0}^{\infty} \Psi_j \varepsilon_{t-j} = \mu + \Psi(B) \varepsilon_t, \quad \text{where } \Psi_0 = \mathbf{I}.$$

The $\Psi(B) = \sum_{j=0}^{\infty} \Psi_j B^j$ is an $m \times m$ matrix in the backshift operator B , such that

$$B^j \varepsilon_t = \varepsilon_{t-j} \quad \text{and} \quad \Psi_{ij}(B) = \sum_{m=0}^{\infty} c_{ij}(m) B^m. \quad \text{The coefficients } \Psi_j \text{ are not necessarily}$$

absolutely summable but do satisfy the weaker condition $\sum_{j=0}^{\infty} \|\Psi_j\|^2 < \infty$.

The vector $\{\varepsilon_t\}$ is a white noise process with $\varepsilon_t = (\varepsilon_{1t}, \dots, \varepsilon_{mt})'$ such that $E(\varepsilon_t) = 0$, $E(\varepsilon_t \varepsilon_t') = \Sigma$ and $E(\varepsilon_t, \varepsilon_{t+k}) = 0$ for $k \neq 0$. According to the above Wold's representation, $\{\varepsilon_t\}$ is defined as the error of the best one-step ahead linear predictor of $\{X_t\}$ based on the infinite past X_{t-1}, X_{t-2}, \dots , with $\varepsilon_t = X_t - \hat{X}_{t-1}$, where \hat{X}_{t-1} is the best linear predictor of $\{X_t\}$ based on X_{t-1}, X_{t-2}, \dots . Then, the ε_t are mutually uncorrelated by construction, since ε_t is uncorrelated with X_{t-j} for all $j \leq 1$ and they also have a constant covariance matrix by stationarity of the process $\{X_t\}$. The best one-step ahead linear

predictor can be expressed as $\hat{X}_{t-1} = \mu + \sum_{j=1}^{\infty} \Psi_j (X_{t-j} - \hat{X}_{t-j-1}) = \mu + \sum_{j=1}^{\infty} \Psi_j \varepsilon_{t-j}$.



CHAPTER 3

Introduction to Vector Autoregressive Models

3.1 Introduction

In this chapter we study the vector autoregressive (VAR) models by giving their definition, some basic properties and we present the least squares method for estimating the parameters of a VAR(p) model.

3.2 Definition of a Vector Autoregressive Model

Definition 3.2.1

A vector autoregressive VAR(p) model is defined as a stationary vector time series $\{\mathbf{X}_t\}$ which satisfies the equation:

$$(\mathbf{X}_t - \boldsymbol{\mu}) - \sum_{j=1}^p \boldsymbol{\Phi}_j (\mathbf{X}_{t-j} - \boldsymbol{\mu}) = \boldsymbol{\varepsilon}_t,$$

where $\boldsymbol{\varepsilon}_t \sim \text{WN}(\mathbf{0}, \boldsymbol{\Sigma})$ and $\boldsymbol{\Phi}_j$, $j = 1, \dots, p$ are $m \times m$ matrices of coefficients.

An equivalent form is $\boldsymbol{\Phi}(B)(\mathbf{X}_t - \boldsymbol{\mu}) = \boldsymbol{\varepsilon}_t$, where $\boldsymbol{\Phi}(B) = \mathbf{I} - \boldsymbol{\Phi}_1 B - \dots - \boldsymbol{\Phi}_p B^p$.

In order for a VAR(p) process to be stationary and causal the following condition must be satisfied. This condition is usually mentioned as the *stationary condition*.

Proposition 3.2.2

The process $\{\mathbf{X}_t\}$ is stationary and causal if all the roots of $\det\{\boldsymbol{\Phi}(B)\} = 0$ are greater than one in absolute value.

Under this condition the VAR(p) model can be written as

$$\mathbf{X}_t = \boldsymbol{\mu} + \sum_{j=0}^{\infty} \boldsymbol{\Psi}_j \boldsymbol{\varepsilon}_{t-j} = \boldsymbol{\mu} + \boldsymbol{\Psi}(B) \boldsymbol{\varepsilon}_t, \quad \text{where} \quad \boldsymbol{\Psi}(B) = \boldsymbol{\Phi}(B)^{-1} = \sum_{j=0}^{\infty} \boldsymbol{\Psi}_j B^j \quad \text{and}$$



$\sum_{j=0}^{\infty} \|\Psi_j\|^2 < \infty$. The coefficients Ψ_j can be given by the relation $\Phi(B)\Psi(B) = I$.

It is mentioned that the latter form is an infinite vector moving average (MA) process.

For example, assume we have a VAR(1) model, $X_t = \Phi X_{t-1} + \varepsilon_t$. In order for the process to be stationary we have to check the roots of $\det\{I - \Phi B\} = 0$. From this we take $\det\{\lambda I - \Phi\} = 0$ with $\lambda = 1/B$. In that way, the process will be stationary if and only if the eigenvalues of Φ be less than one in absolute value.

The case of VAR(1) can be easily extended for models VAR(p). Firstly, we should notice that any VAR(p) model $(X_t - \mu) - \sum_{j=1}^p \Phi_j (X_{t-j} - \mu) = \varepsilon_t$ can always be expressed in the form of a np - dimensional vector VAR(1) in terms of the vector $X_t = (X'_t, \dots, X'_{t-p+1})'$ as $X_t = \Phi X_{t-1} + e_t$ with $e_t = (\varepsilon'_t, 0', \dots, 0)'$ and Φ equal to the $np \times np$ companion matrix associated with the VAR(p) operator

$$\Phi(B), \text{ that is, } \Phi = \begin{bmatrix} \Phi_1 & \Phi_2 & \dots & \dots & \Phi_p \\ I & 0 & \dots & \dots & 0 \\ 0 & I & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & \dots & I & 0 \end{bmatrix}.$$

Then, because $\det\{\Phi(B)\} = \det\{I - \Phi B\} = 0$, the stationarity condition on the roots of $\det\{\Phi(B)\} = 0$ in the VAR(p) model is equivalent to the condition in the corresponding VAR(1) representation that all eigenvalues of the $np \times np$ companion matrix Φ be less than one in absolute value.

3.3 Yule – Walker Equations for Covariance Matrices of a VAR Process

As follows from above any lagged value X_{t-u} of the stationary VAR(p)

process $X_t = \sum_{j=1}^p \Phi_j X_{t-j} + \varepsilon_t$ can be expressed in terms of past values



$\boldsymbol{\varepsilon}_{t-u}, \boldsymbol{\varepsilon}_{t-u-1}, \dots$ in the infinite MA form $\mathbf{X}_{t-u} = \boldsymbol{\Psi}(B)\boldsymbol{\varepsilon}_{t-u} = \sum_{j=0}^{\infty} \boldsymbol{\Psi}_j \boldsymbol{\varepsilon}_{t-u-j}$. From this it follows that $E[\mathbf{X}_{t-u} \boldsymbol{\varepsilon}'_t] = 0$ for $u > 0$ and $E[\mathbf{X}_t \boldsymbol{\varepsilon}'_t] = \boldsymbol{\Sigma}$.

Now, we can write the Yule – Walker equations satisfied by the covariance matrices $\Gamma(u)$ for $u = 1, 2, \dots$ that are defined as $\Gamma(u) = E[\mathbf{X}_{t-u} \mathbf{X}'_t] \Rightarrow$
 $\Gamma(u) = E[\mathbf{X}_{t-u} [\sum_{j=1}^p \boldsymbol{\Phi}_j \mathbf{X}_{t-j} + \boldsymbol{\varepsilon}_t]'] = \sum_{j=1}^p \Gamma(u-j) \boldsymbol{\Phi}'_j$ and $\Gamma(0) = \sum_{j=1}^p \Gamma(-j) \boldsymbol{\Phi}'_j + \boldsymbol{\Sigma}$.

These equations can be used to compute the covariance matrices $\Gamma(u)$ for all values of u in terms of the VAR parameter matrices $\boldsymbol{\Phi}_j$ and $\boldsymbol{\Sigma}$. On the other hand, the matrices $\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \dots, \boldsymbol{\Phi}_p$ and $\boldsymbol{\Sigma}$ can be determined from the covariance matrices $\Gamma(0), \Gamma(1), \dots, \Gamma(p)$ by solving the system of matrix Yule – Walker equations $\sum_{j=1}^p \Gamma(u-j) \boldsymbol{\Phi}'_j = \Gamma(u)$ for $u = 1, 2, \dots, p$. The latter equations can also be written in matrix form as $\Gamma_p \boldsymbol{\Phi} = \Gamma_{(p)}$ with solution $\boldsymbol{\Phi} = \Gamma_p^{-1} \Gamma_{(p)}$ where $\boldsymbol{\Phi} = (\boldsymbol{\Phi}_1, \boldsymbol{\Phi}_2, \dots, \boldsymbol{\Phi}_p)'$, $\Gamma_{(p)} = (\Gamma(1)', \Gamma(2)', \dots, \Gamma(p)')$ and Γ_p a $np \times np$ matrix with (i,j) -th block of elements equal to $\Gamma(i-j)$. Having computed the matrix $\boldsymbol{\Phi}$ we may now find $\boldsymbol{\Sigma}$ as

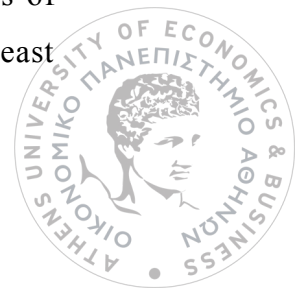
$$\boldsymbol{\Sigma} = \Gamma(0) - \sum_{j=1}^p \Gamma(-j) \boldsymbol{\Phi}'_j = \Gamma(0) - \Gamma'_{(p)} \boldsymbol{\Phi} = \Gamma(0) - \Gamma'_{(p)} \Gamma_p^{-1} \Gamma_{(p)} = \Gamma(0) - \boldsymbol{\Phi}' \Gamma_p \boldsymbol{\Phi}.$$

The spectral matrix of a VAR(p) model can be written as:

$$f(\lambda) = \frac{1}{2\pi} (\boldsymbol{\Phi}'(\bar{z}) \boldsymbol{\Sigma}^{-1} \boldsymbol{\Phi}(z))^{-1} \text{ where } z = e^{-i\lambda}, \lambda \in [-\pi, \pi].$$

3.4 Least Squares Estimations of VAR(p) Models

Let $\mathbf{X}_t = (X_{1t}, \dots, X_{mt})'$ be an m -dimensional vector series that follows the VAR(p) model $\mathbf{X}_t - \boldsymbol{\mu} = \sum_{j=1}^p \boldsymbol{\Phi}_j (\mathbf{X}_{t-j} - \boldsymbol{\mu}) + \boldsymbol{\varepsilon}_t$, where $\{\boldsymbol{\varepsilon}_t\}$ are vector white noise with zero mean and covariance matrix $\boldsymbol{\Sigma} = \text{Cov}(\boldsymbol{\varepsilon}_t)$. In the estimation of VAR(p) models, it is preferable to use least squares estimates of the VAR parameters rather than Yule – Walker estimates. Although these two types of estimates have the same asymptotic properties in the stationary case, the least



squares estimates have better sampling behavior for situations near the non-stationary boundary. In general, the Yule – Walker estimators have larger biases than the least squares estimators due to the impact from “end effects” in the use of the sample covariance matrices $\hat{\Gamma}(u)$. For more see Shaman and Stine (1989) and Pope (1990).

We shall study first the case of $p = 1$. Assume we have the VAR(1) model $X_t - \mu = \Phi_1(X_{t-1} - \mu) + \varepsilon_t$ for $t = 2, 3, \dots, n$. The model can be written equivalently as $X_t = \delta + \Phi_1 X_{t-1} + \varepsilon_t$ with $\delta = (I - \Phi_1)\mu$. Letting $Y = [X_2, X_3, \dots, X_n]'$ and $X = [\mathbf{1}_N, (X_1, X_2, \dots, X_{n-1})]'$, where $\mathbf{1}_N$ denotes an $N \times 1$ column vector of ones with $N = n-1$, we can have the form of the multivariate linear model with $N = n-1$ vector observations. That is $Y = XB + \varepsilon$ with $B = (\delta, \Phi_1)'$ and $\varepsilon = [\varepsilon_2, \varepsilon_3, \dots, \varepsilon_n]'$.

Having written the model in a multivariate linear form we can easily estimate the parameters of the linear model using the usual least squares method. The least squares estimator of $B^* = \Phi_1'$ is given by

$$\hat{\Phi}_1' = (\tilde{X}'\tilde{X})^{-1}\tilde{X}'\tilde{Y} = \left[\sum_{t=2}^n (X_{t-1} - \bar{X}_{(1)})(X_{t-1} - \bar{X}_{(1)})' \right]^{-1} \sum_{t=2}^n (X_{t-1} - \bar{X}_{(1)})(X_t - \bar{X}_{(0)})',$$

where $\tilde{Y} = [(X_2 - \bar{X}_{(0)}), \dots, (X_n - \bar{X}_{(0)})]'$,

$$\bar{X}_{(0)} = N^{-1} \sum_{t=2}^n X_t,$$

$$\bar{X}_{(1)} = N^{-1} \sum_{t=2}^n X_{t-1} \quad \text{and} \quad \tilde{X} = [(X_1 - \bar{X}_{(1)}), \dots, (X_{n-1} - \bar{X}_{(1)})]'$$

The estimator of Σ is: $\hat{\Sigma} = [1/(N - (m+1))]S_1$, where $S_1 = \sum_{t=2}^n \hat{\varepsilon}_t \hat{\varepsilon}_t'$ is the residual sum of squares matrix, with $\hat{\varepsilon}_t = (X_t - \bar{X}_{(0)}) - \hat{\Phi}_1(X_{t-1} - \bar{X}_{(1)})$ the residual vectors.



Theorem 3.4.1

Let $\hat{\phi} = \text{vec}(\hat{\Phi}'_1)$ and $\phi = \text{vec}(\Phi'_1)$. The asymptotic distribution of $N^{1/2}(\hat{\phi} - \phi)$ is multivariate normal $N(0, \Sigma \otimes \Gamma(0)^{-1})$ as $N = n-1 \rightarrow \infty$, where $\Gamma(0) = E[(X_{t-1} - \mu)(X_{t-1} - \mu)']$.

From this it follows that for large N the distribution of $\hat{\phi}$ is approximately $N(\phi, N^{-1}(\Sigma \otimes \Gamma(0)^{-1}))$.

We can easily see that $N(\tilde{X}'\tilde{X})^{-1} \approx \hat{\Gamma}(0)^{-1} \rightarrow \Gamma(0)^{-1}$ as $N \rightarrow \infty$, so that $(\tilde{X}'\tilde{X})^{-1} \approx N^{-1}\Gamma(0)^{-1}$ for large N . In that way the approximate distribution of $\hat{\phi}$ can be estimated as $N(\phi, \hat{\Sigma} \otimes (\tilde{X}'\tilde{X})^{-1})$.

From above, we notice that the LSE of Φ_1 in the stationary VAR(1) model has approximately the same form of distributional properties as the LSE in the standard multivariate linear model. Now, if we write $\Phi_1' = (\phi_1, \phi_2, \dots, \phi_m)$, then the model can be written in the form $X_{it} - \mu = (X_{t-1} - \mu)' \phi_i + \varepsilon_{it}$, meaning that the ϕ_i are the autoregressive coefficients in the (auto)regression for the i -th variable X_{it} . Therefore, $\text{cov}(\hat{\phi}_i) = \hat{\sigma}_{ii}(\tilde{X}'\tilde{X})^{-1}$, $\text{cov}(\hat{\phi}_i, \hat{\phi}_j) = \hat{\sigma}_{ij}(\tilde{X}'\tilde{X})^{-1}$ and $\text{cov}(\hat{\phi}_i, \hat{\phi}_j) = \hat{\rho}_{ij}((\tilde{X}'\tilde{X})^{-1})$ for $i, j = 1, \dots, m$.

The estimates $\hat{\phi}$ and the associated estimates of individual variances and correlations between parameter estimates can be used to make various inferences about the elements of the autoregression matrix Φ_1 , for example hypothesis testing, like $H_0: \Phi_1 = 0$ against $H_1: \Phi_1 \neq 0$. Such a test can be checked by using the likelihood ratio test statistic

$$M_1 = -(N - m - 1 - 1/2) \log[\det(S_1) / \det(S_0)] \text{ where } S_0 = \sum_{t=2}^n (X_t - \bar{X}_{(0)})(X_t - \bar{X}_{(0)})'$$

is the residual sum of squares matrix under H_0 . For large N , M_1 the test statistic is approximately distributed as $\chi^2_{m^2}$ and for large values of M_1 we reject the null hypothesis.



The least squares method can be generalized for VAR(p) models in a similar way as in VAR(1). Let $X_t - \mu = \sum_{j=1}^p \Phi_j (X_{t-j} - \mu) + \varepsilon_t = \Phi'_{(p)} \tilde{X}_t + \varepsilon_t$, where $\Phi'_{(p)} = (\Phi_1, \dots, \Phi_p)$ and $\tilde{X}_t = [(X_{t-1} - \mu)', \dots, (X_{t-p} - \mu)']'$. As we did in the case of $p=1$, we can write the model in the equal form $X_t = \delta + \sum_{j=1}^p \Phi_j X_{t-j} + \varepsilon_t = B' X_t^* + \varepsilon_t$, where $X_t^* = [1, X'_{t-1}, \dots, X'_{t-p}]'$ and $B' = (\delta, \Phi_1, \dots, \Phi_p)$. Moreover, considering we have $N=n-p$ observations we can define the following matrices in order to rewrite the VAR(p) model in a multivariate linear model form. Firstly, we have the $N \times m$ data matrix $Y = [X_{p+1}, X_{p+2}, \dots, X_n]'$, the $N \times (pm+1)$ matrix X whose typical row is $X'_{t,p} = [1, X'_{t-1}, \dots, X'_{t-p}]$, $t = p+1, \dots, n$ and $\varepsilon = [\varepsilon_{p+1}, \dots, \varepsilon_n]'$. Then, the multivariate linear model VAR(p) is $Y = XB + \varepsilon$.

The LS estimator of $B_* = \Phi_{(p)}$ when fitting a VAR(p) model is $\hat{\Phi}_{(p)} = (\tilde{X}'\tilde{X})^{-1} \tilde{X}'\tilde{Y}$, where $\tilde{X}'\tilde{X} = \sum_{t=p+1}^n \tilde{X}_{t,p} \tilde{X}'_{t,p}$, $\tilde{X}'\tilde{Y} = \sum_{t=p+1}^n \tilde{X}_{t,p} (X_t - \bar{X}_{(0)})'$ and $\tilde{Y} = [(X_{p+1} - \bar{X}_{(0)}), \dots, (X_n - \bar{X}_{(0)})]'$. Additionally, the estimator of Σ is $\hat{\Sigma} = (N - (mp + 1))^{-1} S_p$ with $S_m = \sum_{t=p+1}^n \hat{\varepsilon}_t \hat{\varepsilon}'_t$ being the residual sum of squares matrix.

Following similar assumptions and calculations as in the VAR(1) case we get the approximate distribution of $\hat{\phi} = \text{vec}(\hat{\Phi}_{(p)}) \sim N(\phi, \hat{\Sigma} \otimes (\tilde{X}'\tilde{X})^{-1})$ for large N and we can make any testing about the elements of matrix $\hat{\Phi}_{(p)}$.



CHAPTER 4

Portmanteau Test

4.1 Introduction

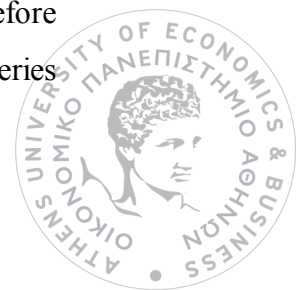
In this chapter we introduce the test statistic Portmanteau which is suggested for testing the goodness of fit of a VAR(p) model by examining the whiteness of its residuals.

The Portmanteau test was firstly introduced by Box and Pierce (1970) for univariate ARMA models and later Chitturi (1974) extended its use to multiple AR processes. In his article it is shown that the residual autocorrelations can be approximately represented as a singular linear transform of the corresponding WN autocorrelations and that they possess a multivariate singular normal distribution. Therefore, he proposed a simple, approximate χ^2 statistic to test the fit. The use of the Portmanteau statistic for multivariate ARMA models was introduced by Hosking (1980).

In reality, Portmanteau can be seen as an additional test of the order chosen for the fitted model. Before reaching this point there are many criteria that can be used regarding the appropriate choice of the order p of each model such as AIC, BIC and FPE. These model selection criteria are computed for various models fitted by maximum likelihood to the series and the fitted model with the minimum value of the given criterion is chosen. For more details on their properties see Quinn (1980) and Reinsel (1997). The idea of using such criteria based on fitting VAR models of various orders can also be applied as a testing procedure for the whiteness of a process (Pukkila and Krishnaiah 1988).

4.2 The Portmanteau Test for White Noise Processes

In most cases it is assumed that the model residuals of a VAR(p) are white noise process. For that reason it is useful to make a further study on these processes before moving on with the testing. As it is mentioned on chapter 2, an m -variate time series



$\{\mathbf{Z}_t\}$ is called *white noise with mean $\mathbf{0}$ and covariance matrix Σ* if $\{\mathbf{Z}_t\}$ is stationary with mean vector $\mathbf{0}$ and covariance matrix function $\Gamma(u) = \{\Sigma_Z$ if $u = 0$ and 0 , otherwise $\}$ and is written $\{\mathbf{Z}_t\} \sim \text{WN}(\mathbf{0}, \Sigma_Z)$.

Assuming we have a process $\{\mathbf{Z}_t\} \sim \text{WN}(\mathbf{0}, \Sigma_Z)$ we may define $\mathbf{Z} := (\mathbf{Z}_1, \dots, \mathbf{Z}_n)$. Then, the autocovariance matrices of $\{\mathbf{Z}_t\}$, $t = 1, \dots, n$ can be estimated as

$$\mathbf{C}_u := \hat{\Gamma}_Z(u) := \frac{1}{n} \sum_{t=u+1}^n \mathbf{Z}_t \mathbf{Z}'_{t-u} = \frac{1}{n} \mathbf{Z} \mathbf{F}_u \mathbf{Z}' \text{ for } u = 0, 1, \dots, h, \text{ for some } h < n.$$

The $n \times n$ matrix \mathbf{F}_u is defined in the obvious way. For instance, for $u=2$:

$$\mathbf{F}_u := \begin{bmatrix} 0 & 0 & \dots & 0 & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & & 0 & 0 & 0 \\ \vdots & & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 1 & 0 & \dots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \dots & 1 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

Having estimated the matrices \mathbf{C}_u we can also estimate the autocorrelation matrices of $\{\mathbf{Z}_t\}$ as: $\mathbf{R}_u := \mathbf{D}^{-1} \mathbf{C}_u \mathbf{D}^{-1}$ for $u = 0, 1, \dots, h < n$, where \mathbf{D} is an $m \times m$ diagonal matrix with elements the square roots of the diagonal elements of \mathbf{C}_0 .

Theorem 4.2.1

Let $\mathbf{c}^{(h)} := \text{vec}(\mathbf{C}_1, \dots, \mathbf{C}_h)$ and $\mathbf{r}^{(h)} := \text{vec}(\mathbf{R}_1, \dots, \mathbf{R}_h)$. Then, under the assumption that $\{\mathbf{Z}_t\} \sim \text{WN}(\mathbf{0}, \Sigma_Z)$, we may write the asymptotic distributions of the autocovariance and autocorrelation matrices, for $h \geq 1$, that is:

$$\sqrt{n} \mathbf{c}^{(h)} \xrightarrow{d} \text{N}(0, \mathbf{I}_h \otimes \Sigma_Z \otimes \Sigma_Z) \quad \text{and} \quad \sqrt{n} \mathbf{r}^{(h)} \xrightarrow{d} \text{N}(0, \mathbf{I}_h \otimes \mathbf{R}_0 \otimes \mathbf{R}_0).$$

The matrix \mathbf{C}_0 ($u=0$) is a consistent estimator of Σ_Z .

Based on this theorem, it may be noted that $\sqrt{n} \text{vec}(\mathbf{C}_i)$, $\sqrt{n} \text{vec}(\mathbf{C}_j)$ have the same asymptotic distribution, i.e. $\text{N}(0, \Sigma_Z \otimes \Sigma_Z)$ and also for $i \neq j$ are asymptotically independent. A similar conclusion can be reached for $\sqrt{n} \text{vec}(\mathbf{R}_i)$, $\sqrt{n} \text{vec}(\mathbf{R}_j)$.

The asymptotic properties of \mathbf{C}_u can be extended even for processes $\{\mathbf{Z}_t\}$ that have a

non-zero mean, let $\bar{\mathbf{Z}} = \frac{1}{n} \sum_{t=1}^n \mathbf{Z}_t$, where then $\mathbf{C}_u := \frac{1}{n} \sum_{t=u+1}^n (\mathbf{Z}_t - \bar{\mathbf{Z}})(\mathbf{Z}_{t-u} - \bar{\mathbf{Z}})'$.



Since $\{\mathbf{Z}_t\} \sim \text{WN}(\mathbf{0}, \Sigma_Z)$ the auto-correlations of $\{\mathbf{Z}_t\}$ are zero for each t . In a matrix form this may be written as $\mathbf{R}^{(h)} := (\mathbf{R}_1, \dots, \mathbf{R}_h) = \mathbf{0}$. Therefore, the null hypothesis that is implemented here is $H_0: \mathbf{R}^{(h)} = \mathbf{0}$ against the alternative $H_1: \mathbf{R}^{(h)} \neq \mathbf{0}$. The Portmanteau test statistic is suggested for checking the above hypothesis regarding the autocorrelation matrices of $\{\mathbf{Z}_t\}$. The test statistic is defined as:

$$P_h := n \sum_{u=1}^h \text{tr}(\hat{\mathbf{R}}_u' \hat{\mathbf{R}}_0^{-1} \hat{\mathbf{R}}_u \hat{\mathbf{R}}_0^{-1}).$$

There are some equivalent forms of the Portmanteau statistic that might be useful. It can be shown that P_h can be written as a function of the covariance matrices of $\{\mathbf{Z}_t\}$. From the above definition we have:

$$P_h := n \sum_{u=1}^h \text{tr}(\hat{\mathbf{R}}_u' \hat{\mathbf{R}}_0^{-1} \hat{\mathbf{R}}_u \hat{\mathbf{R}}_0^{-1} \hat{\mathbf{D}}^{-1} \hat{\mathbf{D}}) = n \sum_{u=1}^h \text{tr}(\hat{\mathbf{D}} \hat{\mathbf{R}}_u' \hat{\mathbf{D}} \hat{\mathbf{D}}^{-1} \hat{\mathbf{R}}_0^{-1} \hat{\mathbf{D}}^{-1} \hat{\mathbf{D}} \hat{\mathbf{R}}_u \hat{\mathbf{D}} \hat{\mathbf{D}}^{-1} \hat{\mathbf{R}}_0^{-1} \hat{\mathbf{D}}^{-1}) \Rightarrow$$

$$P_h := n \sum_{u=1}^h \text{tr}(\hat{\mathbf{C}}_u' \hat{\mathbf{C}}_0^{-1} \hat{\mathbf{C}}_u \hat{\mathbf{C}}_0^{-1})$$

Following the last expression we can use the estimated autocovariance matrices instead to compute the statistic.

In order to use the Portmanteau statistic defined above, we would like to know its asymptotical distributional properties. For this reason, we give below some definitions and propositions that are important in defining the distribution of the statistic.

Definition 4.2.2

An $m \times m$ matrix \mathbf{A} is called *idempotent* if $\mathbf{A} \cdot \mathbf{A} = \mathbf{A}^2 = \mathbf{A}$. For example, $\mathbf{A} = \mathbf{I}_m$.

Proposition 4.2.4

Let $\mathbf{Y} \sim \text{N}(\boldsymbol{\mu}_Y, \Sigma_Y)$. Then, $(\mathbf{Y} - \boldsymbol{\mu}_Y)' \cdot \mathbf{A} \cdot (\mathbf{Y} - \boldsymbol{\mu}_Y) \sim \chi^2_r$, with $r = \text{rank}(\mathbf{A})$ if and only if $\mathbf{A} \cdot \Sigma_Y \cdot \mathbf{A} = \mathbf{A}$.

For more details on these propositions see Reinsel (1997).



From the last expression of the Portmanteau statistic we can get $P_h := n \sum_{u=1}^h \text{tr}(\hat{C}_u' \hat{C}_0^{-1} \hat{C}_u \hat{C}_0^{-1}) = n \text{vec}(\hat{C}_h)' (\mathbf{I}_h \otimes \hat{C}_0^{-1} \otimes \hat{C}_0^{-1}) \text{vec}(\hat{C}_h)$. Having in mind the asymptotic properties of theorem 4.2.1 and by using proposition 4.2.4 it can be shown that $n \text{vec}(\hat{C}_h)' (\mathbf{I}_h \otimes \hat{C}_0^{-1} \otimes \hat{C}_0^{-1}) \text{vec}(\hat{C}_h)$ asymptotically follows a chi-squared distribution with $m^2 h$ degrees of freedom. For more, see Hosking (1980) and Lütkepohl (1993).

Theorem 4.2.5

Let $\{\mathbf{Z}_t\}$ be an identically distributed standard white noise process. Then, for large n and h , we have approximately $P_h \sim \chi^2$ with degrees of freedom $\nu = m^2 h$.

We have to mention that it is very important to choose high values for the sample size n and especially the lag h since the chi-squared distribution for the Portmanteau test may not be accurate for small values. We shall further discuss these parameters in a following section. However, in cases of small samples a modified form of the statistic can be suggested to avoid misleading, such as:

$\bar{P}_h := n^2 \sum_{u=1}^h (n-u)^{-1} \text{tr}(\hat{C}_u' \hat{C}_0^{-1} \hat{C}_u \hat{C}_0^{-1})$. The modification made is an adjustment for the number of terms in the computation of autocovariance matrices \hat{C}_u . In fact, for $n \rightarrow \infty$, $n / (n^2 / (n-u)) \rightarrow 1$ and so the modified statistic has the same asymptotic distribution as P_h .

4.3 The Portmanteau Test for VAR(p) Models

The process $\{\mathbf{Z}_t\}$ discussed above can represent the residuals of a VAR(p) fitted model. Assuming that the residuals are a White Noise process the above results are valid for them too. In that way, the estimated residuals from the model, usually estimated by least squares method, can be used to check the goodness of fit of the model VAR(p). In fact, we check the whiteness of the estimated residuals and from this test we reach to conclusions regarding the adequacy of the fitted model. The idea



of using the residuals computed from a VAR model for testing the goodness of fit was firstly introduced by Box and Pierce (1970) for the univariate case.

Let $\mathbf{X}_t = (X_{1t}, \dots, X_{mt})'$ be an m -dimensional vector series that follows the VAR(p) model $\mathbf{X}_t = \sum_{j=1}^p \Phi_j \mathbf{X}_{t-j} + \boldsymbol{\varepsilon}_t$, where $\{\boldsymbol{\varepsilon}_t\}$ are the white noise residuals of the model with zero mean and covariance matrix $\boldsymbol{\Sigma} = \text{Cov}(\boldsymbol{\varepsilon}_t)$. We assume here that the mean vector of $\{\mathbf{X}_t\}$ is zero. The least squares estimators of the residuals $\{\boldsymbol{\varepsilon}_t\}$, $t = 1, \dots, n$ are given from the relation: $\hat{\boldsymbol{\varepsilon}}_t = \mathbf{X}_t - \sum_{j=1}^p \hat{\Phi}_j \mathbf{X}_{t-j}$.

In order to check the adequacy of the fitted model, that is check the whiteness of its residuals, we use the same statistic as in the case of white noise processes discussed above. Again the statistic checks the hypothesis $H_0: \mathbf{R}^{(h)} = \mathbf{0}$ against the alternative $H_1: \mathbf{R}^{(h)} \neq \mathbf{0}$, where $\mathbf{R}^{(h)} := (\mathbf{R}_1, \dots, \mathbf{R}_h)$ are the autocorrelation matrices of the estimated residuals of the model VAR(p).

Theorem 4.3.1

Let $\{\mathbf{X}_t\}$ be a stationary, stable, m -dimensional VAR(p) process with identically distributed standard white noise process $\{\mathbf{Z}_t\}$ and let the coefficients be estimated by multivariate least squares. Then, for large n and h , the Portmanteau statistic has an approximate χ^2 distribution with degrees of freedom $\nu = m^2(h-p)$ (Ahn 1988).

Having computed the test statistic P_h for the estimated residuals we need to compare its value to the critical value of χ^2 distribution for the corresponding degrees of freedom. If the computed value of P_h is not larger than the critical value at an $\alpha=5\%$ significance level, then the null hypothesis cannot be rejected, that is we cannot reject the hypothesis of whiteness of the residuals for the fitted model VAR(p).

4.4 Comments on the Use of Portmanteau Test

As we have seen above, the Portmanteau test statistic is a very useful tool that can be used to check the whiteness of the residuals for a VAR(p) model.



The test is based on simple computations that require the knowledge of the estimated residuals. However, there are some things that must be considered before using the statistic in order to have accurate results.

The definition of the Portmanteau statistic depends on the choice of lag h , the maximum lag at which we compute the autocovariance and autocorrelation matrices, as is clearly mentioned at all relevant studies (Chitturi (1974), Hosking (1980), Li and McLeod (1981) and Ahn (1988)). For the asymptotic distribution to be valid, the lag h should be taken sufficiently large, so that the elements of Ψ_j in the infinite MA representation for the VAR model are practically zero for $j > h$.

We should also mention that the stationarity assumption implies that the root of $|\Phi_B| = 0$ must lie outside the unit circle. In that way, if we choose roots that are close to the unit circle we expect that the matrices Φ_B will converge slowly to zero as u increases leading to inconveniences of the statistic. So, we need to be careful with the choice of roots.

Moreover, h must be relative small to the sample size n . According to Hosking (1980), the parameter $h = O(n^{1/2})$. Otherwise, the observed values of P_h will be lower than those predicted by chi-squared distribution and consequently the test will not be sufficiently sensitive to deviations from the model. In fact, if h is chosen close enough to the value of n , we get small values of P_h for the majority of cases and we end up with no rejecting the model even for different values of order p for the fitted model. As u increases we will not reject the model for any value of p .

To sum up, it is really important to pay attention to the choice of the parameters p , h , n in order to be able to trust the result of the Portmanteau test and have a clear view regarding the goodness of fit for the VAR(p) model. In general, as the lag h increases within certain bounds we expect that P_h will get more reliable and accurate for any values of p and n .

All things considered, we will next use the Portmanteau test for various sample sizes n , orders fitted p and lags u in order to check in which cases the test coincides with the above discussion, where it is more convenient to use it and whether the results expected from the theory can be justified.



CHAPTER 5

Simulation Study

5.1 Introduction

In this chapter we will simulate the distribution of the Portmanteau test statistic for various models with the purpose of checking the appropriateness of its chi-squared approximation. In order to have more accurate results we make simulations of 1000 samples from each VAR(p) model and by fitting a new VAR(p') at our data, we estimate the value of the statistic for a variety of lags u and order p' of the fitted model. We then compute the percentage of rejection of the null hypothesis for each combination of lag u and order p' . By comparing the results to the nominal level we conclude in which cases and under what conditions the Portmanteau test can be trusted. For the simulation study and the needed algorithms we use the statistical package R.

5.2 Generation of Models

The Portmanteau statistic will be applied on various models in order to examine its accuracy and reliance at different cases. The models used will be generated based on their spectral matrices. The spectral matrix of a time series has already been defined in chapter 2. We will now see how this characteristic can be used to construct a VAR(p) model.

Let $\{X_t\}$ be an m -variate autoregressive process of order p . This can be expressed in the form $\Phi(B)X_t = \varepsilon_t$ where $\Phi(B)$ is an $m \times m$ polynomial matrix with the parameters of the process.

Let $\Psi(z) = \Sigma^{-1/2} \Phi(z)$. Then, $f(\lambda) = \frac{1}{2\pi} (\Psi'(\bar{z})\Psi(z))^{-1}$. For $m=2$, we can get

$$f(\lambda) = \frac{1}{2\pi(\psi_{11}\psi_{22} - \psi_{12}\psi_{21})^2} \begin{bmatrix} |\psi_{22}|^2 + |\psi_{12}|^2 & -(\bar{\psi}_{11}\psi_{12} + \psi_{22}\bar{\psi}_{21}) \\ -(\psi_{11}\bar{\psi}_{12} + \bar{\psi}_{22}\psi_{21}) & |\psi_{11}|^2 + |\psi_{21}|^2 \end{bmatrix}$$

where $\psi_{ij} = \psi_{ij}(e^{-i\lambda})$ and $\psi_{11}(z)\psi_{22}(z) - \psi_{12}(z)\psi_{21}(z) \neq 0$ for $|z| \leq 1$.



It can be proved (Ioannidis 2007) that there exist polynomials χ_1, χ_2 of maximum order p such that:

$$|\chi_1(z)|^2 = |\psi_{22}(z)|^2 + |\psi_{12}(z)|^2 \quad \text{and} \quad |\chi_2(z)|^2 = |\psi_{11}(z)|^2 + |\psi_{21}(z)|^2 \quad \text{for } |z|=1.$$

We can also define the polynomials d, c of maximum order $2p$ as:

$$d(z) = \psi_{11}(z)\psi_{22}(z) - \psi_{12}(z)\psi_{21}(z) \quad \text{and} \quad c(z) = -z^p(\psi_{11}(\bar{z})\psi_{12}(z) + \psi_{22}(z)\psi_{21}(\bar{z})).$$

There is a relationship that combines the four polynomials and is important for defining a new model. For $|z|=1$: $|\chi_1(z)|^2|\chi_2(z)|^2 = |d(z)|^2 + |c(z)|^2$ (1).

In terms of the polynomials defined above the elements of the spectral matrix may be given from:

$$f_{ii}(\lambda) = \frac{1}{2\pi} \frac{|\chi_i(e^{-i\lambda})|^2}{|d(e^{-i\lambda})|^2} \quad \text{and} \quad f_{12}(\lambda) = \frac{1}{2\pi} \frac{e^{i\lambda p} c(e^{-i\lambda})}{|d(e^{-i\lambda})|^2} \quad (2).$$

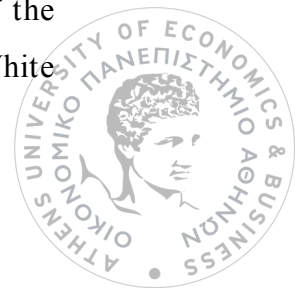
Having given the necessary definitions we can see now how we generate a new model based on the above quantities. In particular, it can be proved (Ioannidis 2007) that for any given χ_1, χ_2 and d, c defined as above that satisfy the relation (1) and $d(z) \neq 0$ for $|z| \leq 1$ there exists a causal VAR($\Phi(B), \Sigma$) model such that its spectral matrix is given from (2).

In that way we can generate any model by giving the polynomials. In fact, we may choose different values of roots z of the polynomials and for each value construct a new model.

5.3 Design of Simulation

The aim of the simulation study is to use the Portmanteau statistic for a variety of cases in order to obtain a more global and clear view of its accuracy and in which cases the test keeps the nominal level. We have seen in chapter 4 that under certain conditions this test cannot be reliable due to many factors, such as the number of lags h or the size of sample n , that influence its asymptotic properties. The most important thing in the use of the Portmanteau test is that can approximately follow a chi-squared distribution and this is something that must not be violated to avoid misleading.

A major role to the effectiveness of the simulation plays the choice of the initial model. We will start the study by simulating under a bivariate White



Noise process where it is expected that the distributional approximation will have the least problems. We then consider a series of models which approach the border of non-stationarity. According to the literature (see section 4.4) it is in such cases that we expect increasing problems and we are interested in checking whether the Portmanteau keeps the nominal level. The models chosen are bivariate VAR(1) with a root of $\det\{|\Phi_z|\}$ approaching 1. We finally consider a bivariate VAR(2) model with a root again close to 1 but with $\text{angle} \neq 0$.

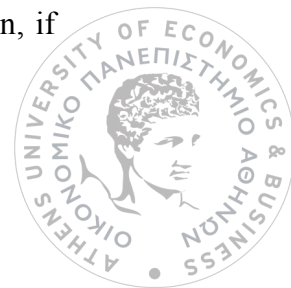
The schedule of the simulation can be easily described in a few lines. For a VAR(p) model given by $\Phi_1, \Phi_2, \dots, \Phi_p$ and Σ (generated as in section 5.2) we simulate a sample series by iterating the equation $X_t = \sum_{j=1}^p \Phi_j X_{t-j} + \varepsilon_t$ and by using a pre-sample of 100.

Moreover, we fit a new VAR(p') model to the time series $\{X_t\} = (X_{t1}, X_{t2})'$ for $t = p' + 1, \dots, n$ and for each combination of order fitted $p' = 1, \dots, 5$ and lag $h = p'+1, \dots, 10$ we compute the Portmanteau statistic as defined in section 4.3.

In particular, we use the form $P_h := (n - p') \sum_{u=1}^h \text{tr}(\hat{C}_u' \hat{C}_0^{-1} \hat{C}_u \hat{C}_0^{-1})$, where the term $(n - p')$ is needed for reducing the bias. We also compute P_h for some extreme values of order fitted, $p' = 8, 10, 15$ for lags $h \leq 20$.

Each value P_h is been compared to the critical value Q of the corresponding χ^2 at $\alpha = 5\%$ significance level and if it is found to be greater than Q we reject the null hypothesis $H_0: R_h = \mathbf{0}$, that is $\{\hat{\varepsilon}_t\}$ is not a white noise process. For example, assume we have fitted a VAR model of order $p' = 2$ and the value of the statistic for $h = 5$ is P_5 . Then, $Q = 21.02601$ (the critical value of $\chi^2(12)$ at $\alpha = 5\%$) and if $P_5 > Q$ we reject H_0 . That means that the fitted model is not accurate and we might need to fit a different model, of different order p' to our initial data.

In fact, we can estimate a percentage of rejection for each value h at a total of M samples. This percentage is given as $\text{Per} = \# (\text{Reject the } H_0) / M$. Let define the variable $D_i = \{1, H_0 \text{ rejected on } i\text{-sample and } 0, \text{ otherwise}\}$. Then, if



$P(D_i=1)=a$, D_i follows a Binomial distribution $B(M,a)$ and we can create a 95% confidence interval for $\text{Per} = \sum_{i=1}^M D_i / M$. For $M = 1000$, $\text{Per} \in (3.6\%, 6.4\%)$ with 95% probability, if $E[D_i]=0.05$ as it should under the null hypothesis. Therefore, we can check whether the computed percentages for each u are close to the predicted values from the above CI.

The simulation study discussed above will be made for three different sample sizes of the time series, $n = 100, 200, 500$ and for various orders p' and lags h .

5.4 Applying the Portmanteau Test on Specific Models

The study considers the six models. As we have mentioned, we will first check a white noise process. Moreover, we will see four models of order $p = 1$ based on pre-selected roots of $\det\{|\Phi_z|\}$, i.e. $z = 0.50, 0.90, 0.95, 0.999$ accordingly and finally one VAR(2) with one root $z_1 = 0.95$ and angle $= 0.50$. It is expected that the test will have difficulties in keeping the nominal level as the root is getting more close to one and as the order of the fitted model increases. We hope that these inconveniences will decrease as the sample size of the time series increases, as it is predicted by the asymptotic theory.

5.4.1 White Noise ($p=0$)

Let $\{X_t\} = (X_{1t}, X_{2t})$ be a bivariate white noise process with zero mean and covariance matrix $\Sigma = I$. We can see a sample of the process in figure 5.1.

We want to fit a VAR(p') model at the time series $\{X_t\}$ and check the goodness of fit with the use of the Portmanteau test. Since the process is white noise, we expect that any model will give white noise residuals, so that the percentages of rejection of the whiteness would be close to 5%.



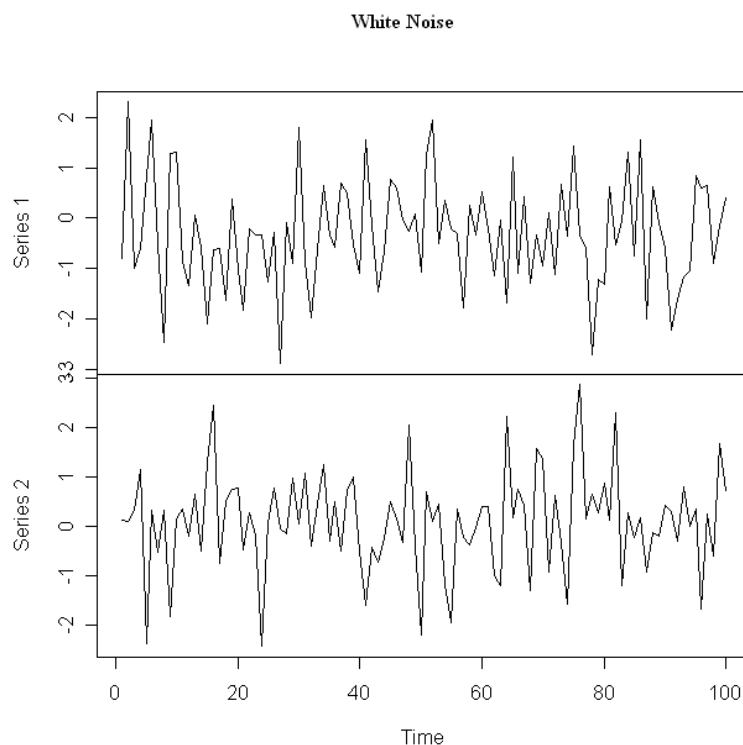


Figure 5.1: A sample of the bivariate White Noise Process

We choose to fit models of order $p' = 0, 1, 2, 3, 4, 5$ and we follow the procedure analysed in the former unit. We will take three different sample sizes for the time series, $n = 100, 200, 500$ and we will compute the Portmanteau statistic for lags $h = 1, 2, \dots, 10$. The percentages of rejecting H_0 at $M = 1000$ samples can be seen in tables 5.1 to 5.3 and their graphical representations at figures 5.2 to 5.4, having a reference line at the nominal level $\alpha = 5\%$. We may reach to the following conclusions:

- Almost all percentages are into the bounds of the predicted confidence interval (3.6%, 6.4%) with small deviances, as it is clearly noted at the figures too, especially for size sample $n=500$.
- As the lag h increases, the percentage decreases for any fixed sample size n and for any fitted order p' leading to great deviances from the nominal level. In that way, the test becomes more and more conservative. For example, at lag $h=10$ and $p'=0$, the percentage is 2.5%

for $n=100$, 3.5% for $n=200$ and 5.2% at $n=500$ with only the last one really close to $\alpha=5\%$.

- As the fitted order p' increases, the test fails to keep the nominal level for small lags h . For example, at $p'=4$ and $h=5$, the percentage is 7.3% for $n=100$ and decreases for larger lags.
- Things tend to be improved as the sample size increases. For $n=500$, all percentages are close to the nominal level $\alpha=5\%$.

WN	n = 100					
	p' = 0	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 1	0,049					
h = 2	0,054	0,051				
h = 3	0,046	0,038	0,054			
h = 4	0,041	0,037	0,036	0,057		
h = 5	0,038	0,032	0,031	0,043	0,073	
h = 6	0,039	0,024	0,027	0,042	0,055	0,108
h = 7	0,037	0,027	0,021	0,036	0,045	0,063
h = 8	0,032	0,024	0,022	0,029	0,035	0,044
h = 9	0,027	0,029	0,019	0,021	0,033	0,036
h = 10	0,025	0,018	0,025	0,013	0,020	0,027

Table 5.1: Percentages of Rejection of Whiteness for White Noise ($n=100$)

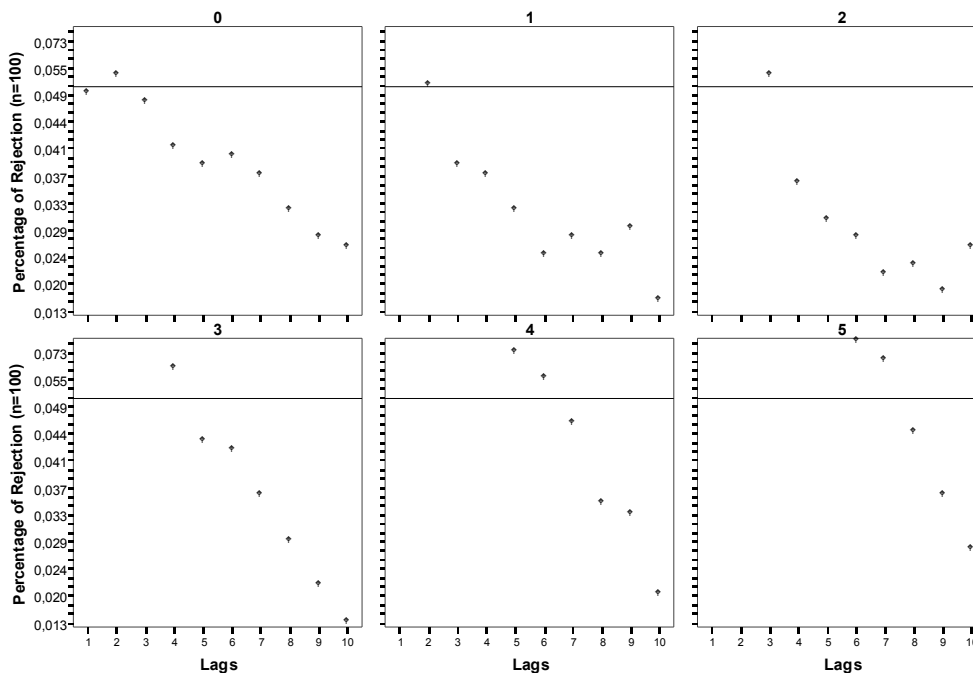


Figure 5.2: Percentages of Rejection of Whiteness for White Noise ($n=100$)

WN	n = 200					
	p' = 0	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 1	0,046					
h = 2	0,037	0,053				
h = 3	0,028	0,039	0,056			
h = 4	0,033	0,051	0,059	0,052		
h = 5	0,025	0,043	0,045	0,041	0,057	
h = 6	0,035	0,035	0,042	0,032	0,039	0,080
h = 7	0,031	0,045	0,032	0,034	0,037	0,062
h = 8	0,032	0,042	0,026	0,032	0,038	0,053
h = 9	0,032	0,039	0,029	0,027	0,039	0,040
h = 10	0,035	0,038	0,024	0,023	0,031	0,039

Table 5.2: Percentages of Rejection of Whiteness for White Noise ($n=200$)

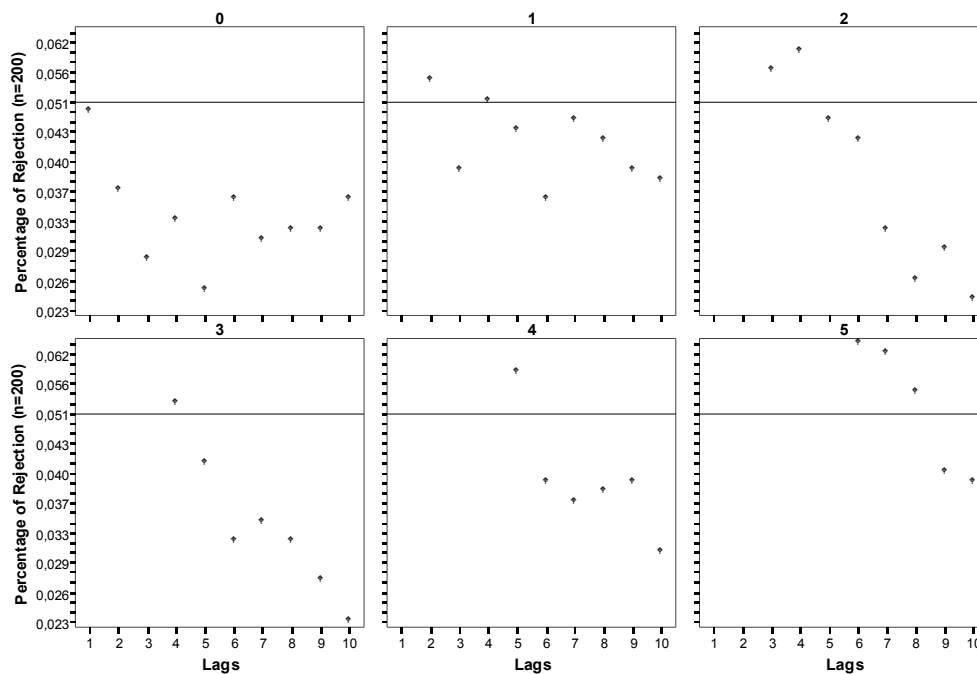


Figure 5.3: Percentages of Rejection of Whiteness for White Noise ($n=200$)



WN	n = 500					
	p' = 0	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 1	0,043					
h = 2	0,052	0,058				
h = 3	0,052	0,055	0,032			
h = 4	0,052	0,054	0,049	0,065		
h = 5	0,059	0,048	0,050	0,049	0,052	
h = 6	0,057	0,053	0,048	0,043	0,053	0,052
h = 7	0,057	0,049	0,036	0,046	0,049	0,056
h = 8	0,051	0,041	0,048	0,037	0,045	0,053
h = 9	0,048	0,040	0,042	0,048	0,047	0,048
h = 10	0,052	0,041	0,044	0,048	0,049	0,054

Table 5.3: Percentages of Rejection of Whiteness for White Noise ($n=500$)

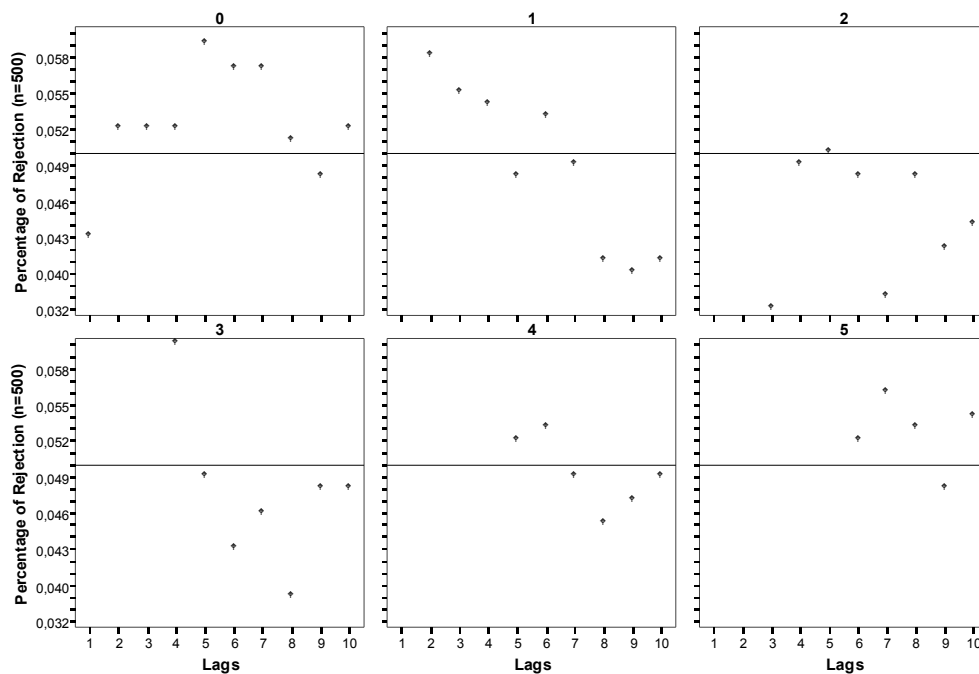


Figure 5.4: Percentages of Rejection of Whiteness for White Noise ($n=500$)



5.4.2 Bivariate Models of Order $p=1$

We present four VAR(1) models on which we apply the Portmanteau test. Each model is been generated using the method discussed in section 5.2. The percentages of rejecting the whiteness of the fitted residuals at each model are been computed again for three different sample sizes at each combination of fitted orders $p' = 1, 2, 3, 4, 5$ and lags $h = 2, \dots, 10$ and for the extreme values $p' = 8, 10, 15$ with lags $h \leq 20$. The parameters of each VAR(1) model are given below.

$$\text{Model A: Root at } 0.50 \rightarrow \Phi_1 = \begin{bmatrix} 0.250 & 0.500 \\ 0.125 & 0.250 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 0.500 & 0.250 \\ 0.250 & 2.125 \end{bmatrix}$$

$$\text{Model B: Root at } 0.90 \rightarrow \Phi_1 = \begin{bmatrix} 0.450 & 0.500 \\ 0.405 & 0.450 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 0.500 & 0.450 \\ 0.450 & 2.405 \end{bmatrix}$$

$$\text{Model C: Root at } 0.95 \rightarrow \Phi_1 = \begin{bmatrix} 0.47500 & 0.50000 \\ 0.45125 & 0.47500 \end{bmatrix} \quad \Sigma = \begin{bmatrix} 0.50000 & 0.47500 \\ 0.47500 & 2.45125 \end{bmatrix}$$

$$\text{Model D: Root at } 0.999 \rightarrow \Phi_1 = \begin{bmatrix} 0.4996285 & 0.4999907 \\ 0.4993861 & 0.4994720 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 0.4999907 & 0.4994720 \\ 0.4994720 & 2.4989170 \end{bmatrix}$$

Some results taken from the simulation made on these models can be seen at the following table 5.4. The complete results along with some figures are given at the Appendix (see tables A.1 to A.4 for models A, B, C and D accordingly).



$n=100$		A	B	C	D
$p'=1$	$h=2$	0,052	0,110	0,115	0,132
	$h=6$	0,039	0,031	0,040	0,046
	$h=10$	0,021	0,017	0,033	0,033
$p'=5$	$h=6$	0,124	0,177	0,182	0,188
	$h=8$	0,052	0,077	0,059	0,088
	$h=10$	0,022	0,038	0,032	0,032
$p'=10$	$h=11$	0,620	0,666	0,718	0,732
	$h=14$	0,137	0,145	0,177	0,183
	$h=17$	0,034	0,052	0,044	0,058
	$h=20$	0,010	0,014	0,012	0,016
$p'=15$	$h=16$	0,999	0,998	0,996	1,000
	$h=18$	0,851	0,856	0,863	0,872
	$h=20$	0,566	0,571	0,593	0,598
$n=200$		A	B	C	D
$p'=1$	$h=2$	0,055	0,096	0,115	0,123
	$h=6$	0,041	0,040	0,051	0,078
	$h=10$	0,050	0,037	0,049	0,064
$p'=5$	$h=6$	0,074	0,144	0,143	0,163
	$h=8$	0,045	0,063	0,066	0,084
	$h=10$	0,044	0,036	0,051	0,048
$p'=10$	$h=11$	0,200	0,318	0,361	0,354
	$h=14$	0,070	0,088	0,115	0,120
	$h=17$	0,039	0,049	0,053	0,062
	$h=20$	0,026	0,031	0,030	0,035
$p'=15$	$h=16$	0,709	0,762	0,782	0,775
	$h=18$	0,292	0,299	0,331	0,343
	$h=20$	0,154	0,175	0,196	0,189
$n=500$		A	B	C	D
$p'=1$	$h=2$	0,070	0,122	0,134	0,145
	$h=6$	0,054	0,056	0,069	0,092
	$h=10$	0,052	0,048	0,059	0,083
$p'=5$	$h=6$	0,080	0,133	0,117	0,157
	$h=8$	0,045	0,067	0,083	0,095
	$h=10$	0,045	0,046	0,062	0,069
$p'=10$	$h=11$	0,112	0,170	0,191	0,190
	$h=14$	0,056	0,077	0,076	0,098
	$h=17$	0,046	0,055	0,063	0,071
	$h=20$	0,035	0,034	0,046	0,063
$p'=15$	$h=16$	0,189	0,286	0,324	0,320
	$h=18$	0,102	0,113	0,138	0,141
	$h=20$	0,068	0,085	0,101	0,088

Table 5.4: Percentages of Rejection of Whiteness for VAR(1) Models



By examining table 5.4 and by comparing the four VAR(1) models, we mention the following:

- For all models, as the lag h increases, the percentages are always decreasing and for small values of order fitted p' the test tends to become conservative. For example, in model B, for $n=100$, $p'=1$ and $h=10$, the percentage of rejecting the whiteness is only 1.7%.
- As the order fitted p' increases, the more lags are needed in order for the test to keep the nominal level. For example, in model C, for $n=200$, the test approaches $\alpha=5\%$ at lag $h=6$ for $p'=1$, at lag $h=10$ for $p'=5$, at lag $h=17$ for $p'=10$ and for $p'=15$ is obvious that we would need a much greater lag.
- For extreme values of order fitted, i.e. $p'=10, 15$, the Portmanteau test fails to keep the nominal level for small lags, especially when the sample size is small. The percentages computed are quite large for any model and seems that only large lags can somehow reduce this problem.
- It is worth mentioning that the larger the sample size gets, the less the problems mentioned above appear. For example, in model A, for $p'=10$ and $h=14$, the percentage of rejection is 13.7% at $n=100$, 7% at $n=200$ and 5.6% at $n=500$, almost as the nominal level. This is in line with the validity of asymptotic theory.
- As the root gets closer to unit, the percentages increase, especially when the lags are small. This is something totally expected, based on the theory, since the stationarity of the process, that tends to be violated for roots almost one, is important for the accuracy of the test. It is, however, encouraging that for larger lags the behaviour of the percentages seems more stable across models. For example, for $n=100$ and $p'=10$, at lag $h=11$, we have 62% for model A and 73% for model D and at lag $h=17$, 3.4% and 5.8% for models A and D accordingly.
- For extreme models, the Portmanteau seems to have the same behaviour regardless the sample size, meaning that even larger sample size is needed for asymptotics to be valid. For example, in model D, for $p'=5$ and $h=8$, we have 8.8% at $n=100$, 8.4% at $n=200$ and 9.5% at $n=500$.



5.4.3 A Bivariate Model of Order $p=2$ (Model E)

We finally study a VAR(2) process generated with the same method based on its spectral characteristics with a root at 0.95 and angle at 0.50. The parameters of the model are the following:

$$\Phi_1 = \begin{bmatrix} 1.492085000 & -0.002893146 \\ 1.635310000 & 1.842728688 \end{bmatrix}, \quad \Phi_2 = \begin{bmatrix} -0.736458000 & 0.01044875 \\ -0.805571200 & -1.09454855 \end{bmatrix} \quad \text{and}$$

$$\Sigma = \begin{bmatrix} 0.8160199 & 0.8925997 \\ 0.8925997 & 2.2018263 \end{bmatrix}.$$

The results of the simulation study on this model are shown at table 5.5. The following points can be noticed in this case:

- The percentages of rejecting the whiteness exceed at great part the expected confidence interval (3.6%, 6.4%) regardless the sample size.
- As the lag increases, the percentages are being reduced without though reaching the nominal level.
- For larger values of order fitted, i.e. $p'=8, 10, 15$, the test tends to get improved for really large lags. For example, at $n=100$, for $p'=8$ and $h=17$, the percentage is 5.1% and for $p'=10$ and $h=20$ is 4.9%. However, for $p'=15$, the test fails to keep the nominal level.
- By comparing this model to the VAR(1) models presented at section 5.4.2, we notice that in this case the Portmanteau test is inappropriate almost for all combinations of the parameters. This may be caused by the double root at 0.95.



E	n = 100			
	p' = 2	p' = 3	p' = 4	p' = 5
h = 3	0,559			
h = 4	0,358	0,571		
h = 5	0,250	0,365	0,560	
h = 6	0,188	0,256	0,330	0,593
h = 7	0,149	0,189	0,209	0,368
h = 8	0,140	0,150	0,147	0,244
h = 9	0,127	0,123	0,100	0,178
h = 10	0,119	0,108	0,088	0,126

E	n = 200			
	p' = 2	p' = 3	p' = 4	p' = 5
h = 3	0,537			
h = 4	0,312	0,530		
h = 5	0,225	0,346	0,531	
h = 6	0,186	0,251	0,320	0,540
h = 7	0,149	0,202	0,214	0,346
h = 8	0,135	0,158	0,172	0,225
h = 9	0,120	0,133	0,141	0,193
h = 10	0,125	0,123	0,113	0,158

E	n = 500			
	p' = 2	p' = 3	p' = 4	p' = 5
h = 3	0,539			
h = 4	0,331	0,539		
h = 5	0,238	0,321	0,551	
h = 6	0,197	0,228	0,328	0,538
h = 7	0,165	0,206	0,239	0,347
h = 8	0,150	0,176	0,190	0,245
h = 9	0,125	0,158	0,154	0,194
h = 10	0,117	0,143	0,124	0,157

Model E (0,95 - 0,5)	n = 100			n = 200			n = 500		
	p' = 8	p' = 10	p' = 15	p' = 8	p' = 10	p' = 15	p' = 8	p' = 10	p' = 15
h = 9	0,762			0,629			0,570		
h = 10	0,496			0,392			0,357		
h = 11	0,325	0,904		0,282	0,723		0,262	0,593	
h = 12	0,249	0,700		0,216	0,464		0,192	0,393	
h = 13	0,177	0,507		0,184	0,317		0,173	0,287	
h = 14	0,114	0,350		0,155	0,231		0,131	0,214	
h = 15	0,082	0,248		0,117	0,177		0,109	0,184	
h = 16	0,070	0,176	0,998	0,102	0,137	0,932	0,115	0,164	0,696
h = 17	0,051	0,134	0,993	0,090	0,096	0,726	0,096	0,152	0,482
h = 18	0,036	0,089	0,947	0,077	0,072	0,561	0,087	0,128	0,364
h = 19	0,027	0,071	0,862	0,069	0,068	0,425	0,077	0,110	0,275
h = 20	0,026	0,049	0,751	0,056	0,058	0,343	0,068	0,110	0,238

Table 5.5: Percentages of Rejection of Whiteness for VAR(2) Model E





CHAPTER 6

Conclusion

In this thesis we have studied a Portmanteau test statistic for the goodness of fit of a VAR(p) model. Having in mind the theory already developed on this test, we have simulated samples from a series of models and have checked the adequacy of the test in each case for various sample sizes, orders fitted and lags. This Portmanteau test is based on an asymptotic distributional theory and the aim has been to see in which cases the simulation coincides with the results expected from this theory.

The simulation study made on six different models has shown that for any sample size and any order fitted, the percentages of rejecting the whiteness of the fitted residuals are decreasing as the number of lags increases. In particular, we have seen that the test tends to keep the nominal level 5%, especially for large sample sizes but it becomes more conservative as the lag increases, especially for small orders fitted. On the other hand, when the order fitted is large, more and more lags are needed in order to be in agreement with the asymptotic theory.

What is also worth mentioning is that the Portmanteau becomes more inappropriate as the model gets closer to non-stationarity. In these extreme cases, a significant large sample is needed in order for the simulated distribution to approximate better the true distribution.

All things considered, the Portmanteau test studied in this thesis can be accurate and reliable for large sample sizes and large number of lags. In order to avoid the problems appeared in more extreme models, the distribution of the statistic could be approximated by bootstrap, making the Portmanteau test statistic appropriate for even more cases.





Appendix 1: Simulation Figures and Results

1. Model A

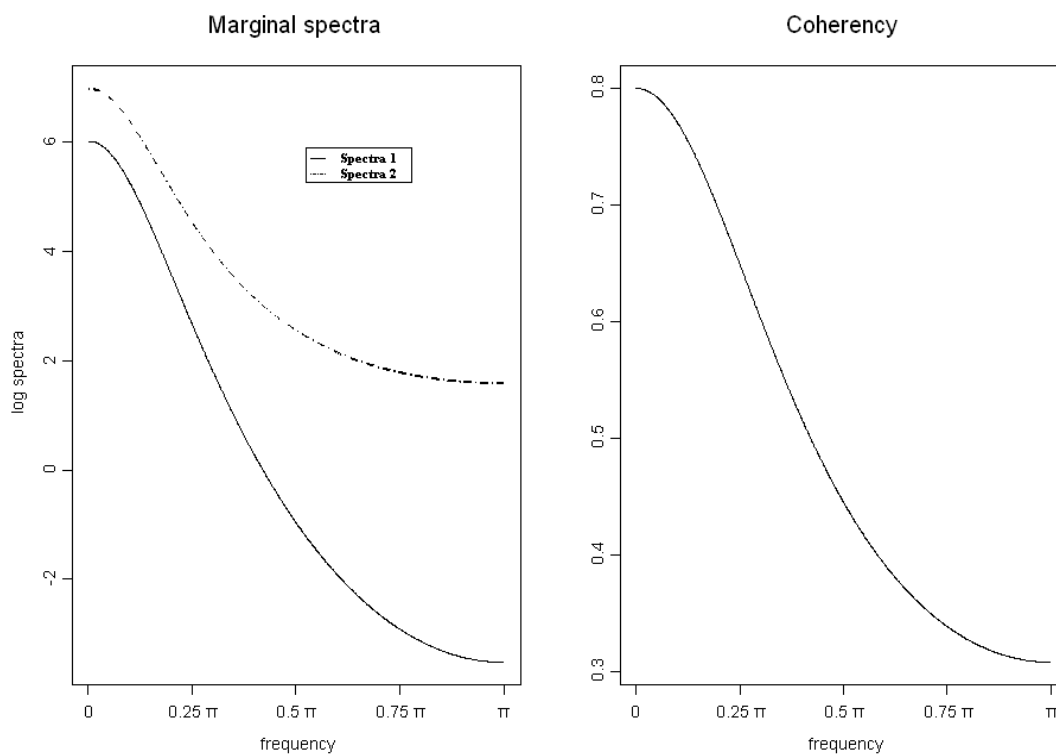


Figure A.1: The spectra and the coherency of the bivariate VAR(1) Model A

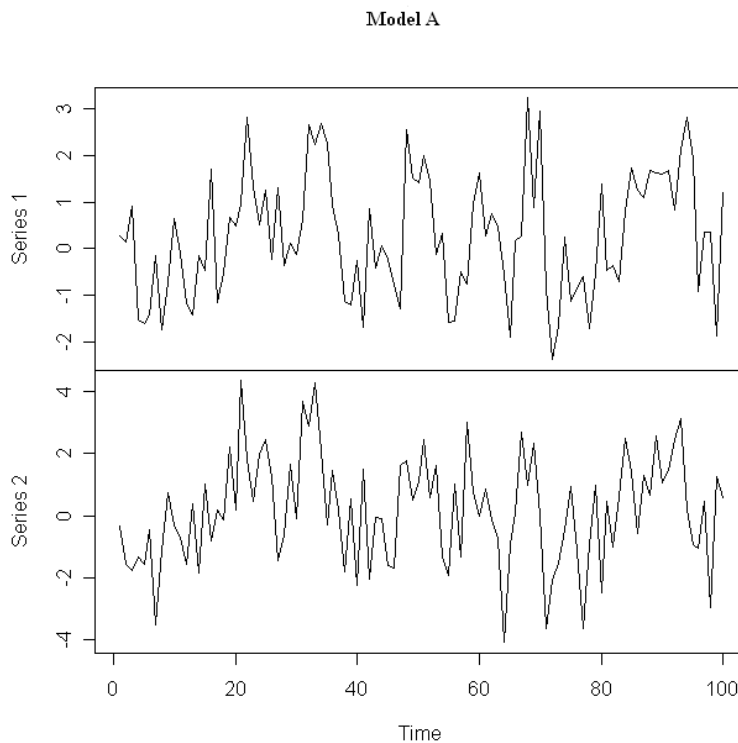


Figure A.2: A sample of the bivariate VAR(1) Model A



A	n = 100				
	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 2	0,052				
h = 3	0,044	0,071			
h = 4	0,039	0,053	0,067		
h = 5	0,039	0,048	0,048	0,092	
h = 6	0,039	0,035	0,044	0,045	0,124
h = 7	0,034	0,032	0,032	0,038	0,078
h = 8	0,026	0,028	0,035	0,019	0,052
h = 9	0,021	0,024	0,020	0,015	0,034
h = 10	0,021	0,022	0,017	0,016	0,022

A	n = 200				
	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 2	0,055				
h = 3	0,049	0,061			
h = 4	0,045	0,051	0,060		
h = 5	0,042	0,052	0,050	0,071	
h = 6	0,041	0,043	0,040	0,062	0,074
h = 7	0,040	0,041	0,036	0,052	0,055
h = 8	0,045	0,039	0,031	0,044	0,045
h = 9	0,046	0,041	0,031	0,038	0,043
h = 10	0,050	0,034	0,021	0,032	0,044

A	n = 500				
	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 2	0,070				
h = 3	0,057	0,059			
h = 4	0,053	0,056	0,049		
h = 5	0,047	0,057	0,054	0,069	
h = 6	0,054	0,046	0,050	0,057	0,080
h = 7	0,049	0,052	0,048	0,045	0,067
h = 8	0,055	0,048	0,054	0,052	0,045
h = 9	0,056	0,047	0,045	0,046	0,050
h = 10	0,052	0,048	0,044	0,057	0,045

Model A (0,50)	n = 100			n = 200			n = 500		
	p' = 8	p' = 10	p' = 15	p' = 8	p' = 10	p' = 15	p' = 8	p' = 10	p' = 15
h = 9	0,311			0,132			0,068		
h = 10	0,156			0,089			0,057		
h = 11	0,096	0,620		0,071	0,200		0,061	0,112	
h = 12	0,059	0,327		0,056	0,129		0,047	0,075	
h = 13	0,051	0,207		0,053	0,089		0,048	0,064	
h = 14	0,031	0,137		0,051	0,070		0,045	0,056	
h = 15	0,021	0,093		0,036	0,057		0,041	0,055	
h = 16	0,013	0,069	0,999	0,028	0,060	0,709	0,048	0,053	0,189
h = 17	0,009	0,034	0,951	0,021	0,039	0,406	0,040	0,046	0,122
h = 18	0,011	0,022	0,851	0,013	0,041	0,292	0,032	0,047	0,102
h = 19	0,006	0,012	0,709	0,010	0,033	0,213	0,036	0,046	0,090
h = 20	0,002	0,010	0,566	0,011	0,026	0,154	0,030	0,035	0,068

Table A.1: Percentages of Rejection of Whiteness for Model A



2. Model B

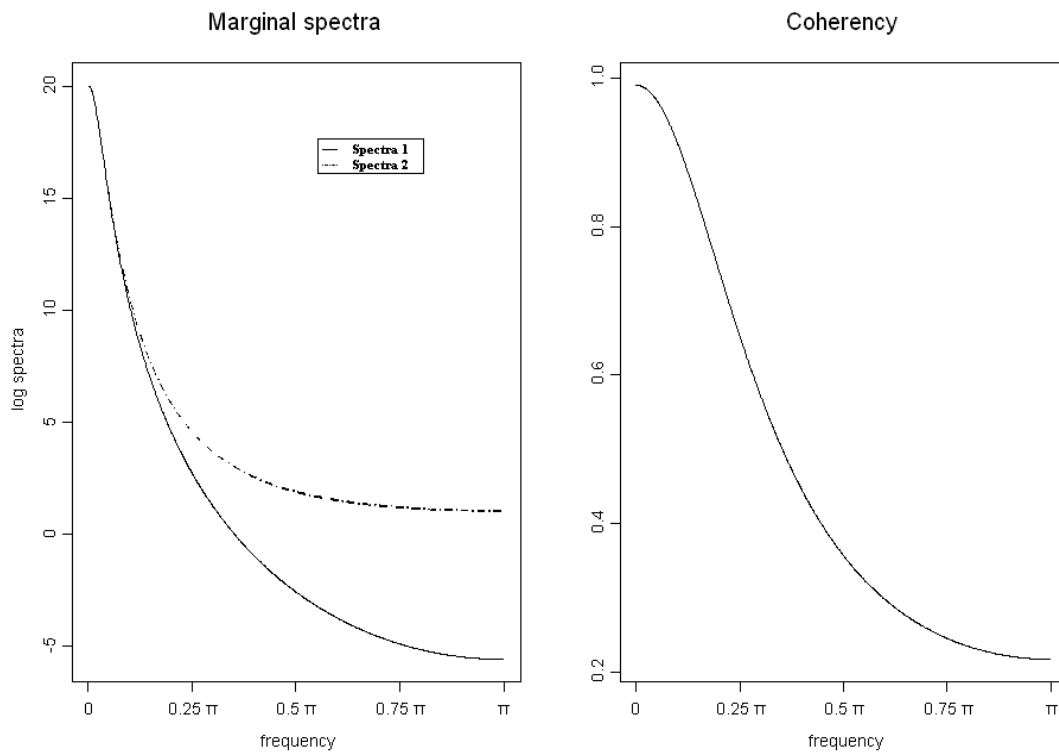


Figure A.3: The spectra and the coherency of the bivariate VAR(1) Model B

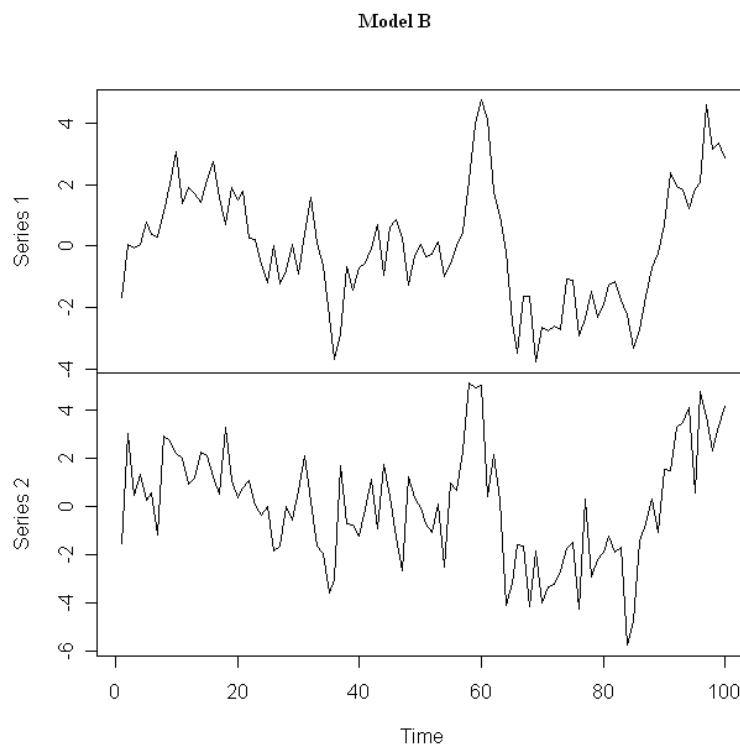


Figure A.4: A sample of the bivariate VAR(1) Model B



B	n = 100				
	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 2	0,110				
h = 3	0,068	0,110			
h = 4	0,051	0,066	0,111		
h = 5	0,044	0,052	0,069	0,156	
h = 6	0,031	0,039	0,042	0,075	0,177
h = 7	0,028	0,030	0,045	0,055	0,101
h = 8	0,026	0,026	0,039	0,036	0,077
h = 9	0,025	0,023	0,033	0,026	0,057
h = 10	0,017	0,022	0,024	0,012	0,038

B	n = 200				
	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 2	0,096				
h = 3	0,066	0,118			
h = 4	0,062	0,079	0,103		
h = 5	0,047	0,064	0,066	0,131	
h = 6	0,040	0,051	0,073	0,078	0,144
h = 7	0,051	0,044	0,044	0,062	0,090
h = 8	0,039	0,042	0,046	0,057	0,063
h = 9	0,039	0,039	0,039	0,045	0,045
h = 10	0,037	0,037	0,036	0,038	0,036

B	n = 500				
	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 2	0,122				
h = 3	0,080	0,087			
h = 4	0,067	0,061	0,102		
h = 5	0,062	0,050	0,077	0,118	
h = 6	0,056	0,047	0,063	0,090	0,133
h = 7	0,050	0,038	0,051	0,071	0,074
h = 8	0,051	0,041	0,051	0,060	0,067
h = 9	0,043	0,038	0,041	0,049	0,056
h = 10	0,048	0,040	0,044	0,050	0,046

Model B (0,90)	n = 100			n = 200			n = 500		
	p' = 8	p' = 10	p' = 15	p' = 8	p' = 10	p' = 15	p' = 8	p' = 10	p' = 15
h = 9	0,407			0,201			0,140		
h = 10	0,211			0,126			0,087		
h = 11	0,111	0,666		0,095	0,318		0,070	0,170	
h = 12	0,082	0,360		0,072	0,176		0,068	0,118	
h = 13	0,045	0,215		0,053	0,103		0,059	0,104	
h = 14	0,028	0,145		0,040	0,088		0,053	0,077	
h = 15	0,020	0,117		0,033	0,076		0,052	0,068	
h = 16	0,019	0,072	0,998	0,025	0,059	0,762	0,051	0,068	0,286
h = 17	0,009	0,052	0,959	0,025	0,049	0,463	0,049	0,055	0,151
h = 18	0,007	0,035	0,856	0,022	0,043	0,299	0,042	0,046	0,113
h = 19	0,004	0,027	0,716	0,012	0,031	0,219	0,037	0,035	0,107
h = 20	0,002	0,014	0,571	0,016	0,031	0,175	0,035	0,034	0,085

Table A.2: Percentages of Rejection of Whiteness for Model B



3. Model C

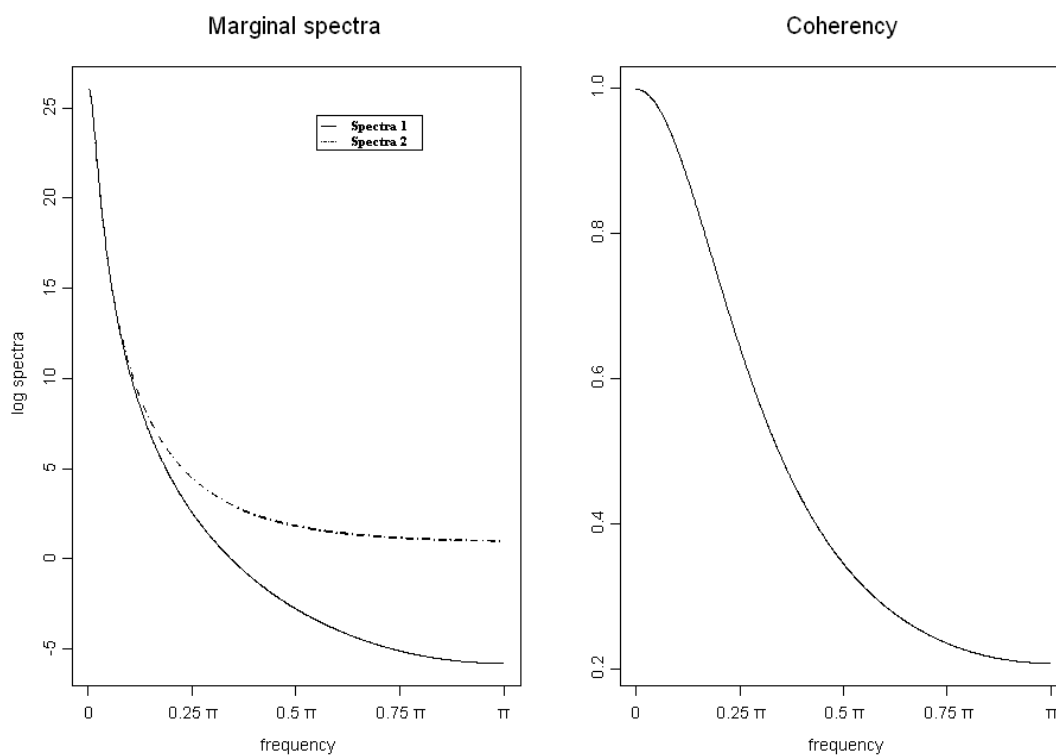


Figure A.5: The spectra and the coherency of the bivariate VAR(1) Model C

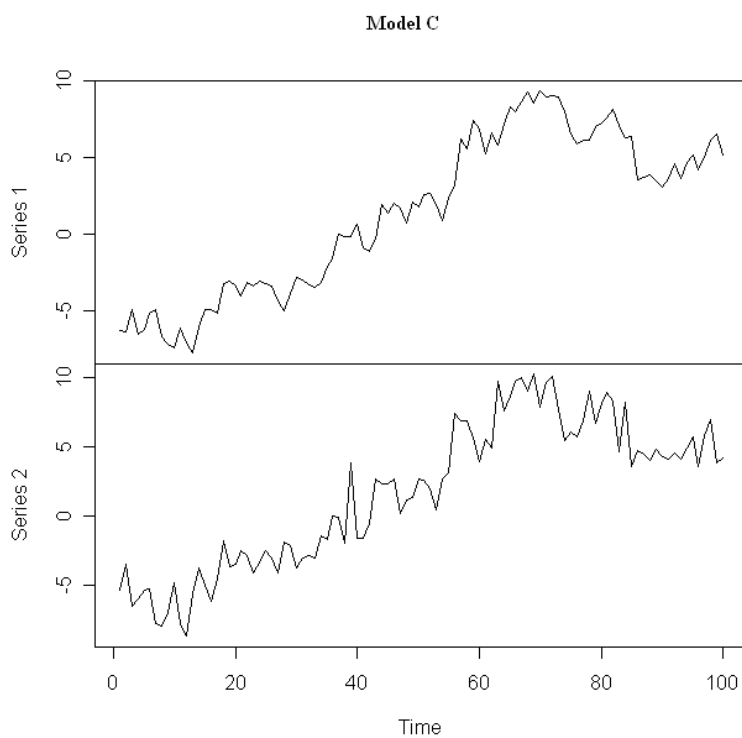


Figure A.6: A sample of the bivariate VAR(1) Model C



C	n = 100				
	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 2	0,115				
h = 3	0,056	0,113			
h = 4	0,049	0,071	0,124		
h = 5	0,050	0,057	0,066	0,173	
h = 6	0,040	0,041	0,043	0,099	0,182
h = 7	0,040	0,031	0,035	0,056	0,103
h = 8	0,037	0,033	0,029	0,039	0,059
h = 9	0,033	0,028	0,021	0,028	0,043
h = 10	0,033	0,026	0,015	0,026	0,032

C	n = 200				
	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 2	0,115				
h = 3	0,087	0,124			
h = 4	0,079	0,095	0,119		
h = 5	0,065	0,079	0,098	0,150	
h = 6	0,051	0,060	0,060	0,098	0,143
h = 7	0,049	0,064	0,058	0,084	0,088
h = 8	0,046	0,056	0,055	0,064	0,066
h = 9	0,049	0,046	0,057	0,056	0,053
h = 10	0,049	0,046	0,046	0,043	0,051

C	n = 500				
	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 2	0,134				
h = 3	0,092	0,127			
h = 4	0,079	0,085	0,116		
h = 5	0,072	0,080	0,083	0,146	
h = 6	0,069	0,072	0,071	0,104	0,117
h = 7	0,066	0,069	0,069	0,081	0,081
h = 8	0,065	0,066	0,051	0,071	0,083
h = 9	0,066	0,050	0,049	0,057	0,078
h = 10	0,059	0,048	0,047	0,050	0,062

Model C (0,95)	n = 100			n = 200			n = 500		
	p' = 8	p' = 10	p' = 15	p' = 8	p' = 10	p' = 15	p' = 8	p' = 10	p' = 15
h = 9	0,429			0,223			0,165		
h = 10	0,236			0,141			0,118		
h = 11	0,151	0,718		0,090	0,361		0,104	0,191	
h = 12	0,105	0,433		0,069	0,206		0,085	0,120	
h = 13	0,073	0,257		0,060	0,152		0,086	0,106	
h = 14	0,043	0,177		0,051	0,115		0,077	0,076	
h = 15	0,027	0,120		0,034	0,093		0,065	0,081	
h = 16	0,018	0,071	0,996	0,030	0,067	0,782	0,065	0,071	0,324
h = 17	0,013	0,044	0,954	0,027	0,053	0,488	0,050	0,063	0,197
h = 18	0,010	0,027	0,863	0,021	0,043	0,331	0,047	0,060	0,138
h = 19	0,005	0,018	0,741	0,010	0,041	0,233	0,042	0,057	0,117
h = 20	0,004	0,012	0,593	0,010	0,030	0,196	0,046	0,046	0,101

Table A.3: Percentages of Rejection of Whiteness for Model C



4. Model D

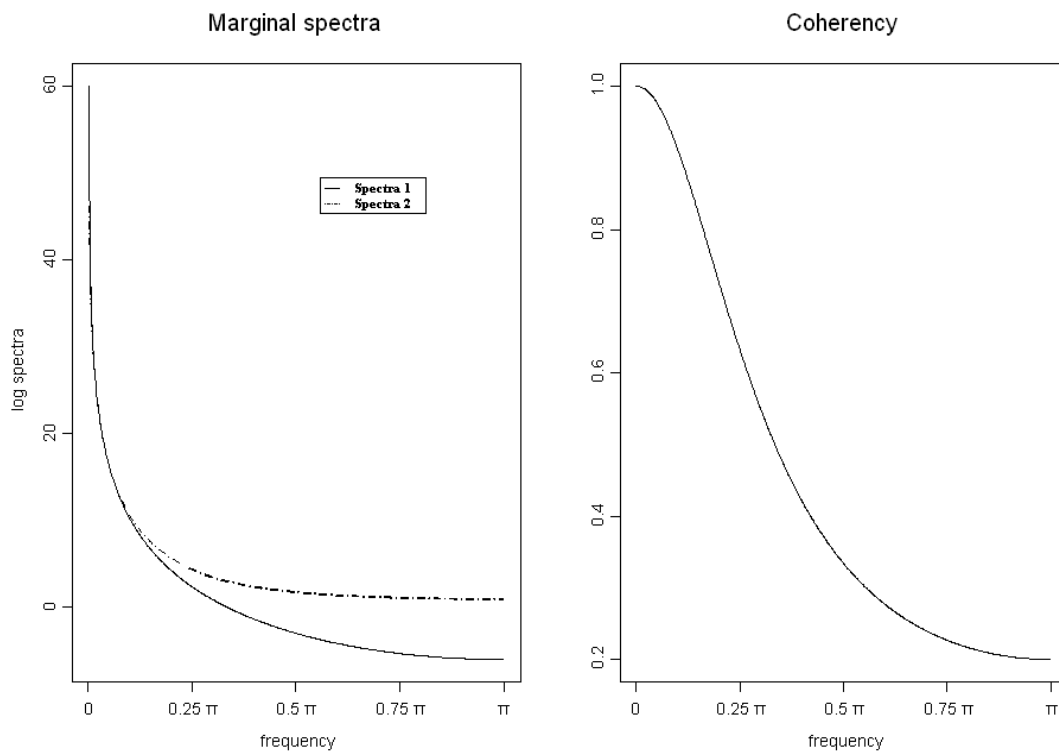


Figure A.7: The spectra and the coherency of the bivariate VAR(1) Model D

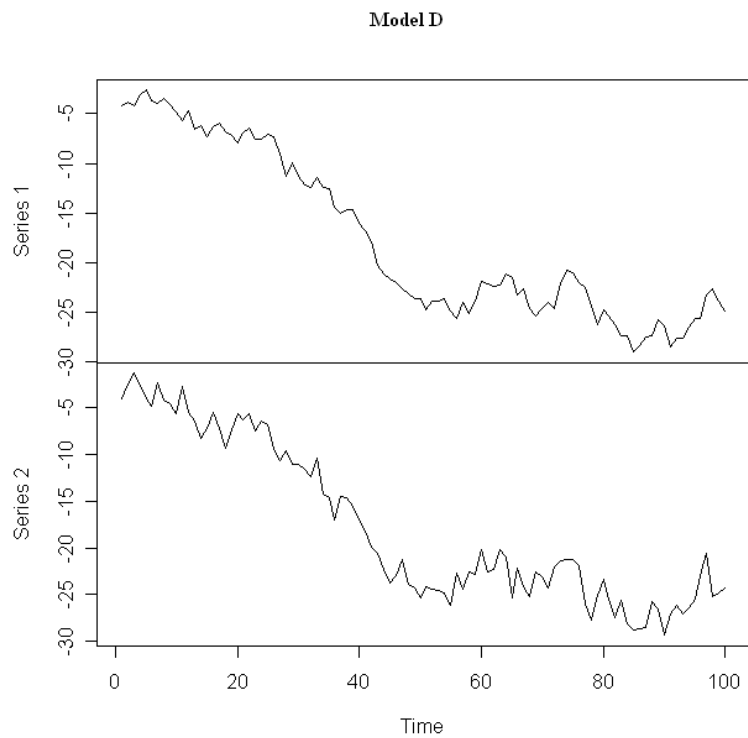


Figure A.8: A sample of the bivariate VAR(1) Model D



D	n = 100				
	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 2	0,132				
h = 3	0,092	0,159			
h = 4	0,076	0,101	0,166		
h = 5	0,063	0,091	0,093	0,180	
h = 6	0,046	0,080	0,062	0,119	0,188
h = 7	0,038	0,065	0,051	0,071	0,128
h = 8	0,038	0,048	0,050	0,043	0,088
h = 9	0,039	0,042	0,044	0,034	0,047
h = 10	0,033	0,039	0,035	0,024	0,032

D	n = 200				
	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 2	0,123				
h = 3	0,106	0,140			
h = 4	0,084	0,090	0,146		
h = 5	0,074	0,076	0,095	0,156	
h = 6	0,078	0,062	0,081	0,100	0,163
h = 7	0,070	0,052	0,068	0,073	0,116
h = 8	0,060	0,052	0,060	0,071	0,084
h = 9	0,061	0,040	0,055	0,064	0,069
h = 10	0,064	0,039	0,048	0,057	0,048

D	n = 500				
	p' = 1	p' = 2	p' = 3	p' = 4	p' = 5
h = 2	0,145				
h = 3	0,116	0,128			
h = 4	0,097	0,104	0,145		
h = 5	0,098	0,092	0,100	0,135	
h = 6	0,092	0,069	0,089	0,089	0,157
h = 7	0,092	0,067	0,075	0,085	0,126
h = 8	0,086	0,058	0,066	0,082	0,095
h = 9	0,088	0,064	0,062	0,077	0,084
h = 10	0,083	0,062	0,064	0,063	0,069

Model D (0,999)	n = 100			n = 200			n = 500		
	p' = 8	p' = 10	p' = 15	p' = 8	p' = 10	p' = 15	p' = 8	p' = 10	p' = 15
h = 9	0,434			0,232			0,155		
h = 10	0,234			0,147			0,108		
h = 11	0,139	0,732		0,123	0,354		0,095	0,190	
h = 12	0,104	0,468		0,084	0,213		0,090	0,136	
h = 13	0,064	0,291		0,071	0,172		0,080	0,107	
h = 14	0,050	0,183		0,057	0,120		0,071	0,098	
h = 15	0,031	0,120		0,037	0,100		0,056	0,088	
h = 16	0,025	0,089	1,000	0,035	0,081	0,775	0,060	0,079	0,320
h = 17	0,020	0,058	0,965	0,030	0,062	0,499	0,064	0,071	0,195
h = 18	0,019	0,041	0,872	0,025	0,053	0,343	0,063	0,075	0,141
h = 19	0,011	0,029	0,757	0,022	0,043	0,251	0,060	0,067	0,109
h = 20	0,006	0,016	0,598	0,015	0,035	0,189	0,054	0,063	0,088

Table A.4: Percentages of Rejection of Whiteness for Model D



5. Model E

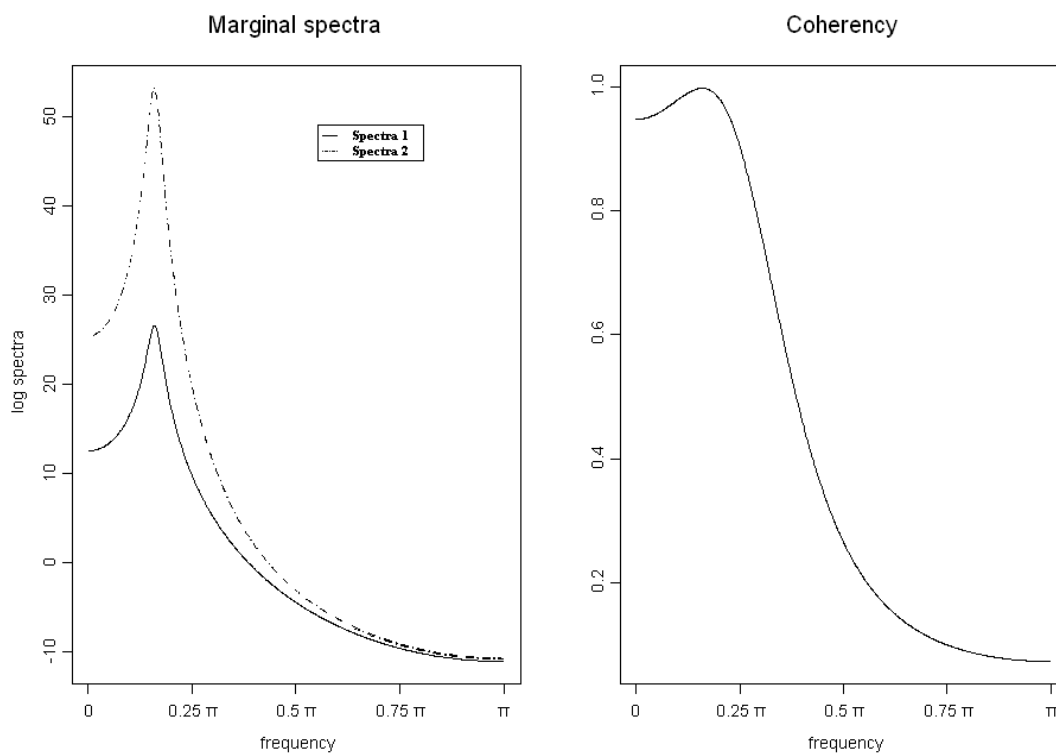


Figure A.9: The spectra and the coherency of the bivariate VAR(2) Model E

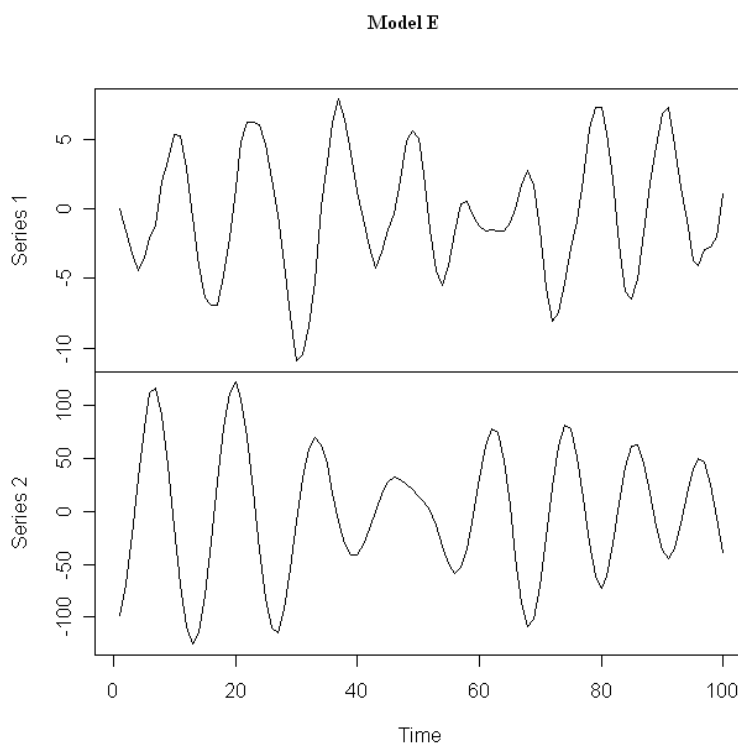


Figure A.10: A sample of the bivariate VAR(2) Model E





Appendix 2: Algorithms in R

◇ Function Multi generates a VAR(p) time series $\{X_t\}$

```
Multi <- function (k,n,p,f1,f2,Sd) {
  l <- eigen(Sd)$values
  V <- eigen(Sd)$vectors
  L <- l*diag(2)
  Snew <- V%*%sqrt(L)%*%t(V)
  h <- mvrnorm(n+k,rep(0,2),diag(2))
  h <- t(h)
  enew <- data.frame(matrix(matrix(0,2,1),2,(n+k)))

  for (i in 1:(n+k)) {
    enew[,i] <- Snew%*%h[,i]
  }

  Xtnew <- data.frame(matrix(matrix(0,2,1),2,(n+k)))

  if (p==0) Xtnew <- enew
  else {
    for (i in 1:p) {
      Xtnew[,i] <- enew[,i]
    }
    for (i in (p+1):(n+k)) {
      Xtnew[,i] <- f1%*%Xtnew[,i-1] + f2%*%Xtnew[,i-2] + enew[,i]
    }
  }

  Xnew <- Xtnew[(k+1):(n+k)]
  Xnew <- as.matrix(Xnew)
  a <- ar.yw(t(Xnew),aic=F,order.max=2)
  return(t(Xnew))
}
```



◇ Function Model fits a new VAR model $\{X_t\}$ and computes the fitted residuals R

```
Model <- function (k,n,p,p.est,f1,f2,Sd) {
  X <- Multi(k,n,p,f1,f2,Sd)
  Y <- matrix(c(X[(p.est+1):n,1],X[(p.est+1):n,2]),n-p.est,2)
  R <- matrix(0,n-p.est,2)
  if (p.est==0) R <- X
  else {
  X1 <- matrix(0,n-p.est,2*(p.est))
    for (j in 1:(p.est)) {
      X1[,2*(j-1)+1] <- X[(p.est+1-j):(n-j),1]
      X1[,2*(j-1)+2] <- X[(p.est+1-j):(n-j),2]
    }
  fit1 <- lm(Y[,1] ~ X1)
  R[,1] <- fit1$residuals
  fit2 <- lm(Y[,2] ~ X1)
  R[,2] <- fit2$residuals
  }
  return(R)
}
```

◇ Function Stat computes the autocovariance matrices for all lags

```
Stat <- function(k,n,p,p.est,f1,f2,Sd,u) {
  res <- Model(k,n,p,p.est,f1,f2,Sd)
  C <- NULL
  i<-0
  while (i<=u) {
    R1 <- res[(1+i):(n-p.est),]
    U1 <- res[1:(n-p.est-i),]
    C1 <- (1/(n-p.est))*t(R1)%*%U1
  }
```



```

        C <- c(C,C1)
        i <- i+1
    }
    C <- matrix(C,2*u+2,2,byrow=T)
    return(C)
}

```

◇ Function Test gives the percentage of rejecting the whiteness

```

Test <- function(k,n,p,p.est,f1,f2,Sd,umax,M) {
    rej <- rep(0,umax)
    perc <- rep(0,umax)
    for (i in 1:M) {
        S <- Stat(k,n,p,p.est,f1,f2,Sd,umax)
        for (u in (p.est + 1):umax) {
            S1 <- solve(S[1:2,])
            h1 <- NULL
            i <- 3
            while (i<=(2*u+1)) {
                S2new <- S[i:(i+1),]
                tr1 <- diag(t(S2new)%*%S1)%*%S2new)%*%S1)
                tr2 <- sum(tr1)
                h1 <- c(h1,tr2)
                i <- i + 2
            }
            P <- (n-p.est)*sum(h1)
            Q <- qchisq(0.95, 4*(u-p.est))
            if (P>Q) rej[u] <- rej[u]+1 else rej[u] <- rej[u]
                perc[u] <- rej[u]/M
        }
    }
    return(perc)
}

```





References

- Ahn, S.K. (1988).** Distribution for Residual Autocovariances in Multivariate Autoregressive Models with Structured Parameterisation, *Biometrika*, 75, 590-593.
- Box, G.E.P. and Pierce, D.A. (1970).** Distribution of Residual Autocorrelations in Autoregressive Integrated Moving Average Time Series Models, *Journal of the American Statistical Association*, 65, 1509-1526.
- Brockwell, P.J. and Davis, R.A. (1987).** Time Series: Theory and Methods, New York, Springer – Verlag.
- Chitturi, R.V. (1974).** Distribution of Residual Autocorrelations in Multiple Autoregressive Schemes, *Journal of the American Statistical Association*, 69, 928-934.
- Hosking, J.R.M. (1980).** The Multivariate Portmanteau Statistic, *Journal of the American Statistical Association*, 75, 602-608.
- Ioannidis, E.E. (2007).** Spectra of Bivariate VAR(p) Models, *Journal of Statistical Planning and Inference*, 137, 554-566.
- Li, W.K. and McLeod, A.I. (1981).** Distribution of the Residual Autocorrelations in Multivariate ARMA Time Series Models, *Journal of the Royal Statistical Society*, B43, No 2, 231-239.
- Lütkepohl, Helmut (1993).** Introduction to Multiple Time Series Analysis, 2nd edition, Springer – Verlag.
- Pope, A.L. (1990).** Biases of Estimators in Multivariate Non-Gaussian Autoregressions, *Journal of Time Series Analysis*, 11, 249-258.



Pukkila, T.M. and Krishnaiah, P.R. (1988). On the Use of Autoregressive Order Determination Criteria in Multivariate White Noise Tests, *IEEE, Transactions on Acoustics, Speech and Signal Processing*, ASSD-36, 1396-1403.

Quinn, B.G. (1980). Order Determination for a Multivariate Autoregression, *Journal of the Royal Statistical Society*, B42, 182-185.

Reinsel, Gregory C. (1997). Elements of Multivariate Time Series Analysis, 2nd edition, Springer.

Shaman, Paul and Stine, Robert A. (1989). A Fixed Point Characterisation for Bias of Autoregressive Estimators, *The Annals of Statistics*, 17, No 3, 1275-1284.



