ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

ANALYZING THE DEPENDENCY OF THE SPECTRAL MATRIX OF A VAR ON THE ROOTS OF ITS CHARACTERISTIC POLYNOMIAL

By

George A. Xronis

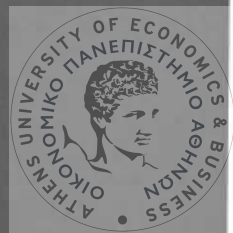
A THESIS

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
2002

Copyright © Athens, Greece, 2002 by Statistical Institute of Documentation,
Research and Analysis.
Department of Statistics, Athens University of Economics and Business

ISBN: 960-8287-43-X



ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ
ΒΙΒΛΙΟΘΗΚΗ
εισ. 72011
Αρ. 519.54
ταξ. ΧΡΟ

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

POSTGRADUATE PROGRAM

ANALYZING THE DEPENDENCY OF THE SPECTRAL MATRIX OF A VAR ON THE ROOTS OF ITS CHARACTERISTIC POLYNOMIAL

By

George A. Chronis

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ΒΙΒΛΙΟΘΗΚΗ

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
January 2003

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ
ΒΙΒΛΙΟΘΗΚΗ
εισ. 72011
Αρ. 579.54
ταξ. ΧΡΟ

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΤΜΗΜΑ ΣΤΑΤΙΣΤΙΚΗΣ

ΑΝΑΛΥΟΝΤΑΣ ΤΗΝ ΕΞΑΡΤΗΣΗ ΤΟΥ ΦΑΣΜΑΤΙΚΟΥ ΠΙΝΑΚΑ ΕΝΟΣ VAR ΜΟΝΤΕΛΟΥ ΜΕ ΤΙΣ ΡΙΖΕΣ ΤΟΥ ΧΑΡΑΚΤΗΡΙΣΤΙΚΟΥ ΤΟΥ ΠΟΛΥΩΝΥΜΟΥ

Γεώργιος Α. Χρόνης

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ΒΙΒΛΙΟΘΗΚΗ

ΔΙΑΤΡΙΒΗ

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

Αθήνα
Ιανουάριος 2003

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΟΙΚΟΝΟΜΙΚΟ
ΠΑΝΕΠΙΣΤΗΜΙΟ
ΑΘΗΝΩΝ
ΒΙΒΛΙΟΘΗΚΗ
εισ. 72011
Αρ.
ταξ.

ATHENS UNIVERSITY OF ECONOMICS AND BUSINESS

DEPARTMENT OF STATISTICS

A Thesis submitted in partial fulfilment of

the requirements for the degree of

Master of Science

ANALYZING THE DEPENDENCY OF THE SPECTRAL MATRIX OF A VAR ON THE ROOTS OF ITS CHARACTERISTIC POLYNOMIAL

George A. Xronis

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ ΒΙΒΛΙΟΘΗΚΗ

Approved by the Graduate Committee

E. Ioannides	N. Frangos	H. Pavlopoulos
Lecturer	Professor	Assistant Professor
Thesis Supervisor	Member of the Committe	

Athens, January 2003

Michael Zazanis, Associate Professor
Director of the Graduate Program

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ACKNOWLEDGEMENTS

I would like to thank my supervisor Lecturer E. Ioannidis for his support, patience and understanding, during the development of this thesis.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΒΙΒΛΙΟΘΗΚΗ

VITA

I was born in Athens in 1973. In 1992 and after my graduation from the Lyceum, I started my studies at the Department of Mathematics in the Aristotle University of Thessaloniki. Four years later I graduated and after the completion of my military obligations, I was accepted as a postgraduate student at the Department of Statistics in the Athens University of Economics and Business.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ
ΒΙΒΛΙΟΘΗΚΗ

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ABSTRACT

George Chronis

"ANALYZING THE DEPENDENCY OF THE SPECTRAL MATRIX OF A VAR ON THE ROOTS OF ITS CHARACTERISTIC POLYNOMIAL."

September 2002

When it comes to defining an AR model for a univariate time series with a spectrum that exhibits strong peaks at specific frequencies, one can easily do so by setting roots of its characteristic polynomial, at that frequency, with modulus close to one. Specifying the polynomial via its roots, in the univariate case, is a simple straightforward process.

In this thesis we try to extend the above idea to the multivariate context by studying the spectral behavior of a two dimensional VAR model.

The behavior of the VAR is driven by a 2×2 polynomial matrix, consisting of 4 polynomials. Our objective is to study that behavior and particularly the presence of 'extreme' spectra, in terms of the roots of the polynomial matrix and its determinant.

It turns out that the behavior of the spectrum of such a model and particularly the presence of strong peaks at certain frequencies, depend crucially on the roots of the determinant (roots that have modulus close to one). But not all of them will create a peak in the VAR's spectra. This also depends on whether they are also roots of the polynomials of which the polynomial matrix of the VAR consists.

We study exhaustively the possible outcomes for simple roots of the determinant and discuss certain aspects regarding the relation between spectral peaks and the coherence values at specific frequencies. As a result we obtain guidelines for the definition of a VAR with specific 'extreme' spectral features.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Furthermore an algorithm is presented, with the help of which one may define a VAR model by specifying the roots of the polynomials involved in the polynomial matrix and its determinant. The algorithm accepts the roots for these five polynomials as input, makes the appropriate checks, introduces some adjustments when needed, distributes the roots according to our instructions and finally finds the polynomials matrix's coefficients, thus defining the VAR model.

Some characteristic cases are presented as examples at the end of the thesis. For further experimentation, the code (created in s-plus) together with a short description, is presented at the appendix.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ΠΕΡΙΛΗΨΗ

Γεώργιος Χρόνης

"ΑΝΑΛΥΟΝΤΑΣ ΤΗΝ ΕΞΑΡΤΗΣΗ ΤΟΥ ΦΑΣΜΑΤΙΚΟΥ ΠΙΝΑΚΑ ΕΝΟΣ VAR ΜΟΝΤΕΛΟΥ ΜΕ ΤΙΣ ΡΙΖΕΣ ΤΟΥ ΧΑΡΑΚΤΗΡΙΣΤΙΚΟΥ ΤΟΥ ΠΟΛΥΩΝΥΜΟΥ."

Σεπτέμβριος 2002

Όταν ενδιαφερόμαστε να ορίζουμε ένα AR μοντέλο για μια μονομεταβλητή χρονοσειρά, με φάσμα που παρουσιάζει ισχυρές κορυφές σε συγκεκριμένες συχνότητες, μπορούμε απλά να ορίσουμε τις ρίζες του χαρακτηριστικού του πολυωνύμου, στην συγκεκριμένη συχνότητα, έτσι ώστε το μέτρο τους να είναι κοντά στην μονάδα. Το να προσδιοριστεί το πολυώνυμο μέσω των ριζών *του*, *στην μονομετα*βλητή περίπτωση, είναι μια απλή διαδικασία.

Σε αυτή την διατριβή προσπαθούμε να επεκτείνουμε την παραπάνω ιδέα σε πολλές διαστάσεις, μελετώντας την φασματική συμπεριφορά δυσδιάστατων VAR μοντέλων.

Η συμπεριφορά ενός VAR μοντέλου καθοδηγείται από έναν 2×2 πολυωνυμικό πίνακα που αποτελείται από 4 πολυώνυμα. Ο στόχος μας είναι να μελετήσουμε την συμπεριφορά και πιο συγκεκριμένα την παρουσία πολύ ισχυρών φασμάτων, σε σχέση με τις ρίζες του πολυωνυμικού πίνακα και της ορίζουσας του.

Αποδεικνύεται ότι η συμπεριφορά του φάσματος ενός τέτοιου μοντέλου και συγκεκριμένα η ύπαρξη πολύ μεγάλων τιμών σε αυτό, εξαρτάται από τις ρίζες αυτής της ορίζουσας (με μέτρο κοντά στην μονάδα). Παρόλα αυτά δεν προκαλούν όλες οι ρίζες της ορίζουσας το ίδιο αποτέλεσμα. Αποδεικνύεται ότι αυτό εξα*ρτάται από το αν* οι ρίζες αυτές είναι επίσης ρίζες των πολυωνύμων που απαρτίζουν τον πολυωνυμικό πίνακα του VAR μοντέλου.

Μελετούμε διεξοδικά τα πιθανά ενδεχόμενα για απλές ρίζες της ορίζουσας και συ*ζητούμε* μερικά ζητήματα σχετικά με την σχέση των υψηλών τιμών του φάσματος *και της* συνοχής, σε συγκεκριμένες συχνότητες. Ως αποτέλεσμα των παραπάνω,

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

λαμβάνουμε κατευθυντήριες γραμμές για τον ορισμό ενός VAR με συγκεκριμένα χαρακτηριστικά πολύ μεγάλων τιμών στο φάσμα του.

Επιπλέον παρουσιάζεται ένας αλγόριθμος, με την βοήθεια του οποίου μπορεί να οριστεί ένα VAR μοντέλο προσδιορίζοντας *τις ρίζες* των πολυωνύμων που περιέχονται στον πολυωνυμικό πίνακα και στην ορίζουσα. Ο αλγόριθμος δέχεται τις ρίζες των πέντε πολυωνύμων, κάνει τους αναγκαίους ελέγχους, εισάγει τις απαραίτητες διορθώσεις, κατανέμει τις ρίζες ανάλογα με τις οδηγίες μας και τελικά βρίσκει τους συντελεστές του πολυωνυμικού πίνακα, δηλαδή ορίζει πλήρως το VAR μοντέλο.

Κάποιες ιδιάζουσες περιπτώσεις παρουσιάζονται στο τέλος της διατριβής *σαν* παραδείγματα. Για περαιτέρω πειραματισμό, ο κώδικας (γραμμένος *σε* S-plus) μαζί με σύντομες επεξηγήσεις, παρουσιάζονται στο τέλος της διατριβής.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

TABLE OF CONTENTS

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Page

Chapter 6: An algorithm for VAR model specification

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

LIST OF TABLES

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

LIST OF PLOTS

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Chapter 1

Introduction

Let us recall the univariate AR(p) processes and their basic concepts. If $\{x(t)\}$ is a univariate time series, we consider it an AR process of order p if:

$x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - .... - \phi_p x_{t-p} = z_t$, where $\{z(t)\} \sim WN(0, \sigma^2)$. (1.1.1)

Or equivalently:

$\phi(B)x_t = z_t$, $t = 0, \pm 1, \pm 2, ...$ where ϕ is a p^{th} degree polynomial: $\phi(z) = 1 - \phi_1 z - \phi_2 z^2 - ... - \phi_p z^p$ which has all its roots outside the unit circle (B is the backshift operator defined by: $B^j x_t = x_{t-j}$).

The reason why the polynomial should have its roots outside the unit circle is to ensure the AR(p)'s stationarity. So, it turns out that the polynomial ϕ is crucial for the behavior of the process.

The significance of the polynomial ϕ is not only due to its involvement in the stationarity issue. As the following will show, the roots of ϕ play a crucial role in the spectral approach of the AR(p). This makes the study of the relation between the spectral density and the polynomial ϕ, vital.

The spectral density of $\{x(t)\}$ is the following:

$$f_x(\lambda) = \frac{\sigma^2}{2\pi \left| \phi(e^{-i\lambda}) \right|^2} \text{ with } -\pi \le \lambda \le \pi . \quad (1.1.2)$$

It turns out that the AR(p) processes are an important subset of the stationary processes. While any stationary model can be approximated arbitrarily well by an AR(p) model, for sufficiently high p, the AR(p)'s are particularly good in dealing

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

with series with strong waves at certain frequency bands (i.e have strong peaks at their spectrum), as will become evident below.

We know that when a polynomial has its constant term equal to one (which ϕ certainly has), it can be written as:

$$\phi(z) = 1 - \phi_1 z - \phi_2 z - \ldots - \phi_p z^p = (1 - \frac{z}{z_1}) \cdot (1 - \frac{z}{z_2}) \cdot \cdots \cdot (1 - \frac{z}{z_p}) \qquad (1.1.3)$$

where z_i , $i = 1, 2, \ldots, p$ are the roots of the ϕ.

So relation (1.1.3) provides an alternative form of the polynomial ϕ, which lies at the denominator of the spectral density f_x (relation (1.1.2)). From the above it is easily understood, that if a root $z_i = \rho_i \cdot e^{i\lambda_i}$ has $\rho_i \approx 1$, then the denominator of the spectral density f_x will be close to zero, thus the f_x will exhibit a peak at the specific frequency λ_i.

This information proves to be very useful, since now in order to define a AR(p) model in which certain frequency components have an extremely strong presence, we can use the roots of its characteristic polynomial and not attempt to directly define its coefficients. So, one can do so by just giving to the AR's polynomial $\phi(z) = 1 - \phi_1 z - \phi_2 z - \ldots - \phi_p z^p$, (which lies in the denominator of the spectrum), a root $z_i = \rho_i \cdot e^{i\lambda_i}$ with $\rho_i \approx 1$, where λ_i will be the frequency in question. It is obvious that at λ_i, the spectrum will exhibit a strong peak.

Such AR processes with extreme peaks at their spectrum (nearly unit root processes), have recently drawn attention, since they have proved to be extremely troublesome regarding their parameter estimation, due to their roots located close to the unit circle.

For example, Dahlhaus (1988), proves that when one considers a sequence of models with an AR root approaching the unit circle at speed $1/T$ --where T is the number of available observations-- the Whittle estimate --which in the AR case is identical to the Yule Walker estimate-- as well as the conditional Maximum Likelihood estimates may be biased, while the exact maximum likelihood estimate is

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

still optimal. In order to deal with the failure of the Yule-Walker estimate, tapered data should be taken. The resulting Yule-Walker estimates are consistent for the above model and have similar optimality properties with the exact Maximum Likelihood estimators.

In order to understand these phenomena let us take a simple model where the data obey:

$$X_t = \rho X_{t-1} + \varepsilon_t, \qquad \varepsilon_t \text{ are } N(0,\sigma^2) \tag{1.1.4}$$

It is well known that if $|\rho| < 1$ the process is stationary and the LS estimator for ρ is asymptotically normal with rate $T^{-1/2}$. If ρ=1, on the contrary the process is a non-stationary Unit-root process and the LS estimator for ρ is asymptotically non-gaussian with rate T^{-1}. The issue is how the LS estimator will behave when ρ is smaller but close to one. This will yield stationary AR models with spectrum at 0 frequency tending to infinity, the behavior of which should approximate the one of a Unit root process.

Chan and Wei (1987), study the transition of the distribution of the LS estimate of ρ under a sequence of models approaching a unit root process, at rate T^{-1}, and come up with a theorem providing a smooth transition between the two extremes (a $T^{-1/2}$-rate gaussian limit and a T^{-1}-rate non-gaussian limit distribution.

Another statistical model describing near integration has been introduced by Phillips, Moon and Xiao (2001). The model provides a more complete interface between I(0) and I(1) models and between $O\left(\sqrt{T}\right)$ and $O(T)$ asymptotics. The rate of convergence to the autoregressive coefficient is $O\left(T^a\right)$ for $a \in \left[\frac{1}{2},1\right]$ and varies in a continuous way between that of a stationary and nonstationary asymptotics.

Testing for situations where a root lies on the unit circle (unit root processes) against stationary alternatives has drawn lots of attention during the last two decades. Most of such unit root tests were based on a "time-domain-approach" setup ((Dickey and Fuller (1979)) and (Phillips and Perron (1988))). But there have also been some attempts to use a "frequency-domain-approach" setup for this testing problem. These are of interest in this context, as in this thesis we regard nearly integrated processes as such with extreme spectral properties.

Hassler (1993) compared in a simulation study, periodogram-regression with standard Unit root tests. It seems that this kind of test can be applied with both

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

fractionally integrated as well as in ARMA processes and furthermore the advantage of no nuisance parameters is at hand. What is also encouraging is that seasonality does not affect the test, which also has an acceptable performance in small datasets.

Another paper that examines the periodogram regression as a unit root test, is one from Akdi and Dickey (1998), in which the test statistic involving the periodogram as well as its distribution, are presented. Some interesting simulations and comments on the power of the test are at hand, together with its utilization on real data.

Finally, Choi and Phillips (1993) proposed tests for a unit root (which also use frequency domain methods), and do not involve nuisance parameters in their limiting distributions. The simulation results in their study show that these tests have good size characteristics in finite samples, although size distortions under negative serially correlated errors, are observed. Their conclusion is that the frequency domain tests have many convenient and appealing properties compared to the standard tests used.

All of the above clearly indicate that in AR processes having extreme peaks in the spectral density (i.e. characteristic polynomial's roots close to unity circle) is something more that a simple annoyance, since it interferes with matters like parameter estimation. So examining these aspects of an autoregressive process in the frequency domain could lead to some interesting conclusions.

What is not self-evident and easy, is the multivariate extension of the presence of extreme peaks. What kind of "extreme peaks" can be present in a VAR? Will all of its components have them simultaneously, or may they be present at only some of them? How will they affect the spectral characteristics reflecting the dependency of these components? Finally, how can one specify VAR(p) models with such features? This is a much more complicated and computationally, different goal.

First of all we will no longer deal with only a single time series but with a collection of series (i.e: a Vector Autoregressive model). So we not only we have n spectra to consider, but beyond those, the aspect of the dependency between these series is at hand. That dependency can be described by the concept of the cross-spectra.

For any pair of components $\{x_i(t)\}$ and $\{x_j(t)\}$ $, i, j = 1, 2, ..., n$, $t = 0, \pm 1, \pm 2, ...$, the cross-covariance function $\gamma_{ij}(h)$ of lag h, is defined as:

$$\gamma_{ij}(h) = Cov[x_i(t+h), x_j(t)] = E[(x_i(t+h) - \mu_i)^*(x_j(t) - \mu_j)] \qquad (1.1.5)$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

The cross-spectra at frequency λ, just like the spectrum in the univariate case, are the Fourier transforms of these cross-covariance functions:

$$f_{ij}(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma_{ij}(h) \qquad (1.1.6)$$

One step further in examining the dependency between two of the n components $\{x_i(t)\}$ and $\{x_j(t)\}$, is the coherence. It strongly resembles the ordinary correlation coefficient and actually indicates the relation between two series of a VAR model at a specific frequency.

So the coherence at a frequency λ is defined by:

$$K_{ij}^2(\lambda) = \frac{\left|f_{ij}(\lambda)\right|^2}{[f_{ii}(\lambda) f_{jj}(\lambda)]} = K_{ji}^2(\lambda), \; i, j = 1, 2, ..., n \qquad (1.1.7)$$

We will utilize all of the above in order to study the more specific case of a two dimensional VAR(p) model, consisted of two components $\{x(t)\}$ and $\{y(t)\}$, $t = 0, \pm 1, \pm 2, ...$ This is the multivariate analogue of the AR(p), and it is defined as:

$$X(t) = \mu + A_1 \cdot X(t-1) + A_2 \cdot X(t-2) + ... + A_p X(t-p) + \varepsilon(t) \text{ where, } (1.1.8)$$

$$X(t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix}, \; Cov\{\varepsilon(t)\} = \Sigma, \; t = 0, \pm 1, \pm 2, \pm 3, ... \text{ and,}$$

$$A_i = \begin{bmatrix} a_{xx}^i & a_{xy}^i \\ a_{yx}^i & a_{yy}^i \end{bmatrix}, \text{ with } i = 1, 2, ..., p$$

In this case, the spectral matrix of the process turns out to be equal to:

$$F(\lambda) = \begin{bmatrix} f_{xx}(\lambda) & f_{xy}(\lambda) \\ f_{yx}(\lambda) & f_{yy}(\lambda) \end{bmatrix} = \frac{1}{2\pi} \Phi^{-1}(e^{-i\lambda}) \cdot \Sigma \cdot \bar{\Phi}^{-1T}(e^{-i\lambda}) \qquad (1.1.9)$$

where Φ is a 2×2 polynomial matrix, called the characteristic polynomial matrix, defined below:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

$$\Phi(z) = I_2 - A_1 \cdot z - A_2 z^2 - ..... - A_p z^p = \begin{pmatrix} 1 - a_{11}^1 z - .... - a_{11}^p z^p & a_{12}^1 z - .... - a_{12}^p z^p \\ a_{21}^1 z - .... - a_{21}^p z^p & 1 - a_{22}^1 z - .... - a_{22}^p z^p \end{pmatrix} =$$

$$= \begin{pmatrix} \Phi_{xx}(z) & \Phi_{xy}(z) \\ \Phi_{yx}(z) & \Phi_{yy}(z) \end{pmatrix} \qquad (1.1.10)$$

It is obvious that this is much more complicated compared to the univariate case, since now we have in our hands four polynomials, plus the determinant of that matrix $F(\lambda)$ which is present at the denominator. In order to simplify the $F(\lambda)$ we will assume, in this thesis, that the covariance matrix $\Sigma = I$.

That way the $F(\lambda)$ and the coherence are equal to:

$$F(\lambda) = \frac{1}{2\pi} \cdot \Phi^{-1}(e^{-i\lambda}) \cdot I \cdot \bar{\Phi}^{-1^T}(e^{-i\lambda}) = \frac{1}{2\pi} \cdot (\Phi^T \cdot \Phi)^{-1} \qquad (1.1.11)$$

Similarly to the univariate case the condition $\det \Phi(z) \neq 0$, for $|z| \leq 1$ implies stationarity for the VAR. So, all of the roots should lie outside the unit circle. The question is what happens if we let these roots approach the boarder of non-stationarity.

Like in the univariate case, it turns out that depending on whether these roots are also roots of $\Phi_{xx}, \Phi_{xy}, \Phi_{yx}, \Phi_{yy}$ some (or all) of the following may occur: $f_{xx}(\lambda_0) \to \infty$, $f_{yy}(\lambda_0) \to \infty$ or $K_{xy}^2(\lambda_0) \to 1$. These are the 'extreme spectra' situations we will study.

In the univariate case --which has been extensively used in this context to motivate our study of extreme spectra of a VAR--, processes with a root close to the Unit Circle were termed "nearly integrated processes" and we saw that they have been used in the literature as local alternatives to Unit root --or integrated processes. A natural question to pose is: is there any analogue to this in the multivariate case? Are VAR(p) models with extreme spectrum (for which a root of det(Φ) is close to the unit circle) approaching some type of multivariate analog of integrated processes?

Our conjecture is that this in fact the case: if the two marginal spectra are approaching infinity at zero frequency, while the coherency tends to Unity, the VAR seems to be approaching a model of two co-integrated I(1) series. If, on the contrary

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

the coherency at zero does not tend to unity, then the VAR seems to be approaching a model of two non-co-integrated I(1) series. There are some indications in this direction. As we know the existence of cointegration relationship between two variables indicates that the series move together in the long run (for details see Engle and Granger (1987)). This, in the frequency domain, refers to a zero frequency relation of the series. It has been proved (Levy (2002)) that if two difference stationary series $x(t)$ and $y(t)$ are cointegrated with cointegrating vector [1, b], then the zero frequency coherence, phase and gain of their first differences will be equal to one, zero and $|b|$ respectively. Furthermore in that aspect (Granger and Weiss (1983)), the coherence at zero frequency tends to one if and only if the two series are cointegrated.

The above allow eventually a re-interpretation of the theorems that will be presented later on in the thesis, in terms of as to whether the VAR model considered approaches a model of two co-integrated series or not.

An algorithm that would take the above under consideration and enable us to create a VAR model with a pre-specified spectra and coherence could be of great importance since, once that is achieved, we will be in the position to construct any kind of a two-dimensional VAR model with the structure we want, something that it is not possible by simply defining its coefficients. The perspectives of the frequency domain analysis combined with the work presented in this thesis, give us an in-depth understanding of a VAR's structure.

In this thesis, starting at chapter two, we give a short outline of the univariate stationary processes and their spectral density issues.

Later on, in chapter three, we introduce the bivariate stationary processes. These are the simplest multivariate realizations and upon them the concepts of cross-spectra and correlation (i.e. coherence) between the time series, are analyzed.

In chapter four we focus on the family of Vector Autoregressive Processes (VAR) and some simple issues like stability, stationarity and estimation, are explored.

Our main results whatsoever are presented in chapter five, where we investigate the theoretical behaviour of the spectrum of a VAR process. The analysis of the characteristic polynomial of a VAR, is of most importance and reveals the complexity of the investigation attempted here. Besides to the theorems, that the analysis of the 'extreme spectra' situations leads to, a categorization of cases where the ranks of the

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

roots are small, is also achieved. Many interesting results occur, for example: if the rank of $z_i = \rho_i \cdot e^{i\lambda_i}$ as a root of the determinant $\det(\Phi)$ is one, it turns out that it is not possible for two series to have strong peaks at their spectra (in a given frequency) without being highly correlated in that frequency band and vice versa. This is not valid when the order of z_i is greater than one, where it is possible to obtain the two series which are highly correlated at a frequency band, without their spectra exhibiting strong power over that frequency.

Finally, in chapter six, we develop an algorithm in order to create VAR models simply by giving the roots for each of the five polynomials. Some examples from its utilization are presented and later on the algorithm itself is analyzed at the appendix, together with some comments on it structure.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Chapter 2

Univariate Stationary Processes

2.1 Stationary processes

A set of random variables {x(t)}, $t \in Z$ and its associated probability distributions is called a stochastic process. The 'family' of random processes includes a vast range of processes, rendering a unified approach in their study ineffective. Consequently more specialized subjects have been defined such as Markov processes, birth and death processes, diffusion and dilution processes and so on.

The classical theory of spectral analysis considers only a fraction of the family of random processes called 'stationary' processes. The main feature of these processes is that their statistical properties are shift invariant.

This means that if x(t) is such a process, then: …x(1), x(2), x(3), ..x(t),….. must each have the same probability density function. But that's not all that the above definition implies. It implies moreover that {x(1),x(4)} , {x(2),x(5)} , (x(3),x(6)} must have the same bivariate probability density and further more that {x(1),x(3),x(6)} , {x(2),x(4),x(7)} must have the same trivariate probability density function and so on. The formal definition is the following (Priestley, (1981)):

Definition 2.1.1: The process {x(t)} is said to be completely stationary if, for any $t_1, t_2, t_3, ..., t_n$ and any k, the joint probability distribution of $\{x(t_1), x(t_2), x(t_3), ...x(t_n)\}$ is identical with the joint probability of $\{x(t_1+k), x(t_2+k), ..., x(t_n+k)\}$.

That means that:

$$F_{x(t_1),....x(t_n)}(x_1, x_2, ..., x_n) \equiv F_{x(t_1+k),....x(t_n+k)}(x_1, x_2, ..., x_n) \qquad (2.1.1)$$

where $F(.)$ denotes the distribution function of the set of random variables which appear as suffixes.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

If we relax the above definition we have the weak stationarity:

Definition 2.1.2: The process {x(t)} is said to be stationary up to order m if, for any $t_1, t_2, t_3, ..., t_n$ and any k, all the joint moments up to order m of $\{x(t_1), x(t_2), x(t_3), ...x(t_n)\}$ exist and equal to the corresponding joint moments up to order m of $\{x(t_1+k), x(t_2+k), ..., x(t_n+k)\}$.

This means that

$$E[\{x(t_1)\}^{m_1}\{x(t_2)\}^{m_2}.....\{x(t_n)\}^{m_n}] = E[\{x(t_1+k)\}^{m_1}\{x(t_2+k)\}^{m_2}.....\{x(t_n+k)\}^{m_n}] \quad (2.1.2)$$

For any k, and all non-negative integer $m_1, m_2, ..., m_n$ satisfying $m_1 + m_2 + m_3 + ... + m_n \le m$.

2.2 The autocovariance function

If our observations originate from a time series they may not assumed to be independent. So the neighboring values in a time series are most likely correlated. Hence besides defining:

the mean value μ as: $\mu = E[x] = \int_{-\infty}^{+\infty} x f_x(x)dx$ (2.2.1)

and the variance σ^2 as: $\sigma^2 = E[(x-\mu)^2] = \int_{-\infty}^{+\infty} (x-\mu)^2 f_x(x)dx$ (2.2.2)

it is also necessary (for stationary series) to specify the autocovariance function (ACVF) as follows, (Jenkings and Watts, (1968)):

Definition 2.2.1: Let $\{x(t)\}$ be a stationary time series. The autocovariance function (ACVF) of $\{x(t)\}$ is $\gamma_x(h) = Cov[x(t+h), x(t)] = E[(x(t+h)-\mu)(x(t)-\mu)]$

From the above definition it is obvious that:

$$\gamma_x(0) = Var[x(t)] = \sigma_x^2 \text{ and } \gamma_x(h) = \gamma_x(-h). \quad (2.2.3)$$

Hence the autocovariance function is an even function.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

If a normalization of the autocovariance function is needed, possibly for comparing time series with different scales, the autocorrelation function (ACF) is introduced: $\rho_x(h) = \frac{\gamma_x(h)}{\gamma_x(0)} = Cor[x(t+h), x(t)]$.

Since $\rho_x(h)$ is a correlation coefficient it should be bounded: $|\rho_x(h)| \le 1$. And furthermore it should be symmetric attaining its maximum value at $h = 0$.

The autocovariance and autocorrelation functions are estimated by the sample autocovariance $\hat{\gamma}_x(h)$ and sample autocorrelation function $\hat{\rho}_x(h)$ respectively, where:

$$\hat{\gamma}(h) = n^{-1} \sum_{t=1}^{n-|h|} (x(t+|h|) - \bar{x})(x(t) - \bar{x}), \quad -\eta < h < \eta \text{ where } \bar{x} = \frac{1}{n} \sum_{t=1}^{n} x_t \qquad (2.2.4)$$

and

$$\hat{\rho}_x(h) = \frac{\hat{\gamma}_x(h)}{\hat{\gamma}_x(0)}. \qquad (2.2.5)$$

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΒΙΒΛΙΟΘΗΚΗ ΑΘΗΝΩΝ

2.3 Spectral density of univariate series

Let $\{x(t)\}$ be a zero mean stationary process (without loss of generality) and $\gamma(t)$ be the autocovariance function of that series, where $\sum_{h=-\infty}^{\infty} |\gamma(h)| < \infty$. Then, (Brockwell and Davis, (1996)), the spectral density of $\{x(t)\}$ is defined by:

$$f(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma(h), \quad -\infty < \lambda < \infty \qquad (2.3.1)$$

where $e^{i\lambda} = \cos(\lambda) + i\sin(\lambda)$ and $i = \sqrt{-1}$.

Definition 2.3.1: A function f is the spectral density of a stationary series $\{x(t)\}$ with autocovariance function γ if:

1. $f(\lambda) \ge 0$ for every $\lambda \in (0, \pi]$ (2.3.2)

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

2. $\gamma(h) = \int_{-\pi}^{\pi} e^{-ih\lambda} f(\lambda) d\lambda$ for all integers h. (2.3.3)

Properties of spectral density:

1 f is an even function, that is $f(-\lambda) = f(\lambda)$ (2.3.4)

2 $f(\lambda) \geq 0$ for every $\lambda \in (-\pi, \pi]$ (2.3.5)

3 $\gamma(k) = \int_{-\pi}^{\pi} e^{-ik\lambda} f(\lambda) d\lambda = \int_{-\pi}^{\pi} \cos(k\lambda) f(\lambda) d\lambda$ (2.3.6)

The above property actually expresses the autocovariance function of a stationary process, as a function of the spectral density. So even if $\sum_{h=-\infty}^{\infty} |\gamma(h)| = \infty$, a spectral density may exist.

In order to see the interpretation of the spectrum, a theorem concerning the spectral representation of a stationary process, should be presented. (Priestley page 247)

Theorem 2.3.1 (Priestley, (1981), page 246) Let $\{x(t)\}, t = 0, \pm 1, \pm 2, \ldots$ be a zero mean stationary process. There exists an orthogonal process $Z(\lambda)$ on the interval $(-\pi, \pi)$ such that for all integral t, $\{x(t)\}$ can be written in the form:

$$x(t) = \int_{\pi}^{\pi} e^{it\lambda} dZ(\lambda) \quad (2.3.7)$$

The process $Z(\lambda)$ has the following properties:

a) $E[dZ(\lambda)] = 0$ for $-\pi \leq \lambda \leq \pi$

b) $E[|dZ(\lambda)|^2] = f(\lambda) d\lambda$ for $-\pi \leq \lambda \leq \pi$ and,

c) for any two distinct frequencies λ, λ' the $Cov[dZ(\lambda), dZ(\lambda')] = E[dZ^*(\lambda) dZ(\lambda')] = 0$.

From the above theorem we can write the process $\{x(t)\}$ as:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

$x(t) = \int_{-\pi}^{\pi} e^{it\lambda} dZ(\lambda)$ and, (2.3.8)

where $\{dZ(\lambda)\}$ is the orthogonal processes having the properties described above.

The interpretation of the spectrum is straightforward. It is obvious that it describes in the frequency domain, exactly what the autocovariance function describes in the time domain. That is because the spectrum and the autocovariance function are Fourier transform pairs and one may just as well consider one aw the other. It can be considered as a decomposition of the time series into a set of frequency bands, which measure the relative importance of each of the bands in terms of the contribution of this band to the total variance of the series. That way the spectrum provides information on the stochastic structure of the process, which is extremely useful in out attempts in fitting a model. Furthermore it can be estimated by fairly simple numerical techniques, which do not require any specific assumptions on the structure of the process.

2.4 Estimation of spectral density

If γ is the autocovariance function of a stationary time series $\{x(t)\}$ that has spectral density f. The basis for the estimation of that density is given by the periodogram.

The definition of the Periodogram is the following:

Definition 2.4.1: The Periodogram of a process $\{x(t)\}$ is the following function:

$$I_n(\lambda) = \frac{1}{n}\left|\sum_{t=1}^{n} x_t e^{-it\lambda}\right|^2 = \frac{1}{n}\{[\sum_{t=1}^{n} x(t)\sin(\lambda t)]^2 + [\sum_{t=1}^{n} x(t)\cos(\lambda t)]^2\}. \quad (2.4.1)$$

In order to understand how the periodogram acts as an estimator of the spectral density, the Fourier frequencies must be used.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Let F_n be the set that consists of the frequencies $\omega_k = \frac{2\pi k}{n}$ with $k = -\left[\frac{n-1}{2}\right],\ldots\ldots,\left[\frac{n}{2}\right]$ and $[.]$ the integer part. Then F_n is a subset of $(-\pi,\pi]$ and consists of values that we call Fourier frequencies associated with sample size n.

The following theorem (Brockwell and Davies, (1996), page 122) shows that $I_n(\lambda)$ is the sample analogue of $2\pi f(\lambda)$. We must have in mind that $\sum_{h=-\infty}^{\infty}|\gamma(h)| < \infty$

thus $2\pi f(\lambda) = \sum_{h=-\infty}^{\infty} e^{-ih\lambda}\gamma(h),\ \lambda \in (-\pi,\pi]$ (2.4.2)

Theorem 2.4.1: If $\{x(t)\}$ is a real process and λ_k is any of the nonzero Fourier frequencies $\lambda_k = \frac{2\pi k}{n}$ in $(-\pi,\pi]$, then:

$$I_n(\lambda_k) = \sum_{|h|\le n} e^{-ih\lambda_k}\hat{\gamma}(h), \qquad (2.4.3)$$

where $\hat{\gamma}(h)$ is the sample autocovariance function of $\{x(t)\}$.

Theorem 2.4.2 (Brockwell and Davis, (1987), page 337) **:** It can be shown that:

If $\{x(t)\}$ is a linear process:

$$x(t) = \sum_{j=-\infty}^{\infty}\psi_j Z_{t-j} \text{ with } \{Z(t) \sim IID(0,\sigma^2) \qquad (2.4.4)$$

where $\sum_{j=-\infty}^{\infty}|\psi_j| < \infty$

and $\{x(t)\}$ has a strictly positive spectral density, then for any fixed frequencies $\lambda_1,\lambda_2,\lambda_3,\ldots,\lambda_m$ such that $0 < \lambda_1 < \lambda_2 < \lambda_3 \ldots < \lambda_m < \pi$ the periodogram ordinates $\{\frac{I_n(\lambda_1)}{2\pi f(\lambda_1)},\frac{I_n(\lambda_2)}{2\pi f(\lambda_2)},\ldots\ldots,\frac{I_n(\lambda_m)}{2\pi f(\lambda_m)}\}$ are approximately distributed as independent exponential random variables.

Thus the probability of an estimation error cannot be made arbitrary small by choosing a sufficiently large sample. This means that the periodogram is not a

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

consistent estimator. So it is common practice to smooth (or else window) the periodogram.

Towards that end we consider the class of estimators having the form:

$$\hat{f}(\lambda_i) = \frac{1}{2\pi} \sum_{|\kappa| \le m_n} I_n(\lambda_{i+\kappa}) \cdot W_{m_n}(\kappa) \qquad (2.4.5)$$

where $I_n(\lambda_i)$ is the periodogram,

W_{m_n} is a sequence of weight functions and,

In other words $\hat{f}(\lambda_i)$ corresponds to a "locally" weighted average of periodogram ordinates in the neighbourhood of the frequency λ_i. That smoothing is achieved with the function $W()$ which is called the "spectral window". The critical characteristic of a window is its width. It can be shown that in order to obtain a good estimate of a peak in the spectrum, the width of a spectral window must be of the same order of the width of the peak.

In order for this estimate of the spectral density to be consistent, (Brockwell and Davies, (1987)) we impose the following conditions:

a) $m_n \to \infty$ and $m/n \to 0$ as $n \to \infty$ (2.4.6)

b) $W(t) \ge 0$ for all t. (2.4.7)

c), $W(t) = W(-t)$ (2.4.8)

d) $\sum_{|t| \le m_n} W(t) = 1$ and, (2.4.9)

e) $\sum_{|t| \le m_n} W^2(t) \to 0$ as $n \to \infty$. (2.4.10)

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

If the above conditions hold, then:

***Theorem 2.4.3*:**

If $\{x(t)\}$ is a linear process:

$$x(t) = \sum_{j=-\infty}^{\infty} \psi_j Z_{t-j} \qquad \text{with} \qquad \{Z(t)\} \sim IID(0,\sigma^2) \tag{2.4.11}$$

$$\text{where} \sum_{j=-\infty}^{\infty} |\psi_j| < \infty \qquad \text{and} \qquad E[Z^4(t)] < \infty \tag{2.4.12}$$

Then:

$$\text{a) } \lim E\hat{f}(\lambda) = f(\lambda) \text{ when } n \to \infty \tag{2.4.13}$$

$$\text{b) } \lim \frac{Cov[\hat{f}(\lambda), \hat{f}(\omega)]}{\sum_{|i| \le m_n} W^2(i)} = \begin{cases} 2\hat{f}^2(\lambda) & \omega = \lambda = 0, or\ \pi \\ \hat{f}^2(\lambda) & \text{when } 0 < \omega = \lambda < \pi \\ 0 & \omega \neq \lambda \end{cases} \tag{2.4.14}$$

Furthermore it can be shown that the variance of the spectral estimator, is inversely proportional to the bandwidth of the spectral window. So small variance is associated with large bandwidths and large variance is associated with small bandwidths.

In practice the weight functions should be carefully chosen in order to sustain a satisfactory balance between the bias and the variance.

2.5 The autoregressive processes and their spectrum

Definition 2.5.1: The process $\{x(t)\}$, $t = 0,\pm1,\pm2,....$ is said to be a autoregressive process of order p {AR(p)} if $\{x(t)\}$ is stationary and if for every t:

$$x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} - .... - \phi_p x_{t-p} = z_t, \text{ where } \{z(t)\} \sim WN(0,\sigma^2) \tag{2.5.1}$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

In a different notation the AR(p) process may be written as: $\phi(B)x_t = z_t$ $t = 0,\pm1,\pm2,....$ where ϕ is a p^{th} degree polynomial: $\phi(z) = 1 - \phi_1 z - \phi_2 z - ... - \phi_p z^p$ and B is the backshift operator defined by: $B^j x_t = x_{t-j}$.

Regarding the spectral density of such a process the following theorem applies (Brockwell and Davis, (1987), page 121):

Theorem 2.5.1: Let $\{x(t)\}$ be an AR(p) process satisfying $\phi(B)x_t = z_t$ with $\{z(t)\} \sim WN(0,\sigma^2)$ where $\phi(z) = 1 - \phi_1 z - \phi_2 z - ... - \phi_p z^p$ has no zeroes on the unit circle. Then $\{x(t)\}$ has spectral density:

$$f_x(\lambda) = \frac{\sigma^2}{2\pi\left|\phi(e^{-i\lambda})\right|^2} \text{ with } -\pi \le \lambda \le \pi. \qquad (2.5.2)$$

It is usually assumed that the roots of the $\phi(z)$ lie outside of the unit circle, as in this case there is a causal solution to $\phi(B)x_t = z_t$. If the polynomial $\phi(z) = 1 - \phi_1 z - \phi_2 z - ... - \phi_p z^p$ has a root $z_0 = \rho_o e^{i\lambda_o}$, then the spectral density, will have a peak at λ_o, whose size will depend on the modulus of the root. The closer the modulus is to one, the stronger the peak and that is when particularly interesting cases are obtained.

Let as consider a AR(1) and a AR(2) processes

$x_t - \phi\, x_{t-1} = z_t$ and $x_t - \phi_1' x_{t-1} - \phi_2' x_{t-2} = z_t$

where $\phi = 0.8$ and $\phi_1' = 0.9$ and $\phi_2' = -0.81$ with $\{z(t)\} \sim WN(0,\sigma^2)$

The root of the first polynomial is 1.25 while the second one has a complex one (and its conjugate) $0.5555556 \pm 0.9622504i$.

The spectral density of the AR(1) and the AR(2) processes are presented below:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

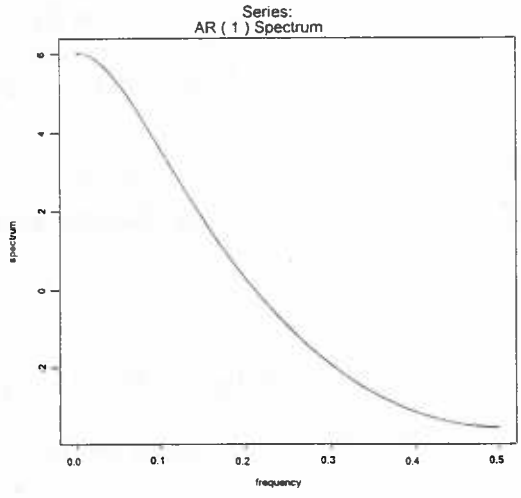


Figure 1.1

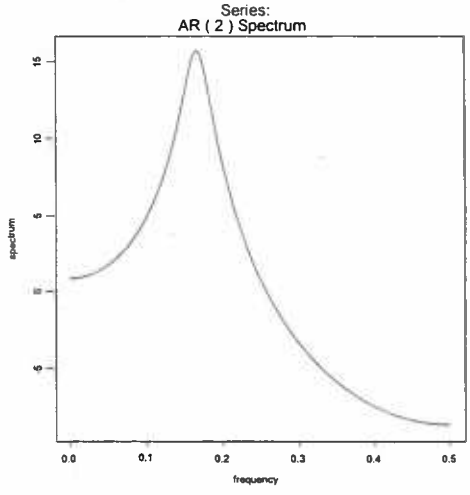


Figure 1.2

As expected the AR(2) process has a larger peak since the root of its polynomial is closer to unity. Furthermore the peak is shifted towards the right by $\pi/2$, due to the angle of the root.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Chapter 3

Bivariate Stationary Processes

3.1 Cross-covariance and cross-correlation functions

Let us consider two stochastic processes $\{x(t)\}$ and $\{y(t)\}$ $t = 0,\pm1,\pm2,....$ We say that $x(t), y(t)$ is a stationary bivariate process if:

I) $\{x(t)\}$ and $\{y(t)\}$ are univariate stationary processes and,

II) The $Cov[x(t), y(s)]$ is a function of $(t-s)$ only.

Provided the above conditions hold, we may define the covariance function for each of the two processes, (as mentioned in section 2.2). So:

$$\gamma_{xx}(h) = Cov[x(t+h), x(t)] = E[(x(t+h)-\mu_x)^*(x(t)-\mu_x)] \text{ and} \qquad (3.1.1)$$

$$\gamma_{yy}(h) = Cov[y(t+h), y(t)] = E[(y(t+h)-\mu_y)^*(y(t)-\mu_y)]. \qquad (3.1.2)$$

(The symbol * represents the complex conjugate of the factor.)

The corresponding autocorrelation functions are:

$$\rho_{xx}(h) = \frac{\gamma_{xx}(h)}{\gamma_{xx}(0)} = Cor[x(t+h), x(t)] \text{ and} \qquad (3.1.3)$$

$$\rho_{yy}(h) = \frac{\gamma_{yy}(h)}{\gamma_{yy}(0)} = Cor[y(t+h), y(t)] \qquad (3.1.4)$$

The above functions are used to describe the correlation structure within each process. In the case of a bivariate, or a multivariate, process the correlation structure between the processes has to be derived. In order to do so the cross-variance and the cross-correlation function are defined (Priestley, (1981)).

The cross-covariance functions $\gamma_{ij}(h)$, $i = \{x, y\}$ of lag h, are defined as:

$$\gamma_{xy}(h) = Cov[x(t+h), y(t)] = E[(x(t+h)-\mu_x)^*(y(t)-\mu_y)] \text{ or} \qquad (3.1.5)$$

$$\gamma_{yx}(h) = Cov[y(t+h), x(t)] = E[(y(t+h)-\mu_y)^*(x(t)-\mu_x)]. \qquad (3.1.6)$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

The basic property of the autocovariance function is that:

$$\gamma_{xy}(h) = Cov[x(t+h), y(t)] = E[(x(t+h)-\mu_x)^*(y(t)-\mu_y)] = E[(x(t)-\mu_x)^*(y(t-h)-\mu_y)] =$$

$$= E[(y(t-h)-\mu_y)(x(t)-\mu_x)^*] = Cov[y(t-h), x(t)] = \gamma_{yx}(-h) \qquad (3.1.7)$$

Thus the covariance structure of the two stochastic processes can be described by means of a single cross-covariance function.

Generally it may be needed to study the covariance structure of two processes with different scales and variances. For that purpose a normalized version of the cross-covariance function, the cross-correlation function $\rho_{ij}(h)$ $i = \{x, y\}$, is defined:

$$\rho_{xy}(h) = \frac{\gamma_{xy}(h)}{\sqrt{\gamma_{xx}(0)\gamma_{yy}(0)}} = \frac{\gamma_{xy}(h)}{\sigma_x \sigma_y} \qquad (3.1.8)$$

Like the correlation function mentioned in section 2.2, the cross-correlation function is a correlation coefficient as well, thus $|\rho_{xy}(h)| \leq 1$. But it is not symmetric and can attain its maximum value anywhere.

3.2 Spectra and cross-spectra

Once again, let us consider two stochastic processes $\{x(t)\}$ and $\{y(t)\}$ $t = 0, \pm 1, \pm 2, \ldots$ that have purely continuous spectral density functions:

$$f_{xx}(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma_{xx}(h) \text{ and} \qquad (3.2.1)$$

$$f_{yy}(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma_{yy}(h) \text{ respectively,} \qquad (3.2.2)$$

where γ_{xx} and γ_{yy} are the autocovariance functions for each process.

The above sums are actually the Fourier transforms of the respective autocovariance functions. It is obvious that the same transformation may be applied to the cross covariance function γ_{xy}, as defined in section 3.1. So we have (Priestley, (1981)):

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

***Definition 3.2.1*:** If $\{x(t)\}$ and $\{y(t)\}$ are two zero mean stationary processes, then the function:

$$f_{xy}(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} e^{-ih\lambda} \gamma_{xy}(h) \tag{3.2.3}$$

is called the cross-spectral density function of $\{x(t)\}$ and $\{y(t)\}$ and exists for every λ provided that $\sum_{h=-\infty}^{\infty} |\gamma_{xy}(h)| < \infty$.

From the above definition the spectral representation of the $\gamma_{ij}(h)$ are the following:

$$\gamma_{ij}(h) = \int_{-\pi}^{\pi} e^{ih\lambda} f_{ij}(\lambda) d\lambda \text{, with } i, j = \{x, y\} \tag{3.2.4}$$

Unlike the spectral densities $f_{xx}(\lambda)$ and $f_{yy}(\lambda)$, the cross spectrum $f_{xy}(\lambda)$ is typically a complex-valued function, since the cross-covariance function $\gamma_{xy}(h)$ is generally not symmetric about zero.

As we recall from theorem 2.3.1, $\{x(t)\}$ and $\{y(t)\}$ can be written as:

$$x(t) = \int_{-\pi}^{\pi} e^{it\lambda} dZ_x(\lambda) \text{ and,} \tag{3.2.5}$$

$$y(t) = \int_{-\pi}^{\pi} e^{it\lambda} dZ_y(\lambda), \tag{3.2.6}$$

where $\{Z_i(\lambda), -\pi \le \lambda \le \pi\}$, $i = 1,2$ are orthogonal increment processes.

Furthermore, from the same theorem:

$$f_i(\lambda) d\lambda = E[|dZ_i(\lambda)|^2], \ (i = \{x, y\}) \tag{3.2.7}$$

It can also be proved that the processes $\{Z_i(\lambda)\}$, $(i = \{x, y\})$ have an additional property:

$$E[dZ_i(\lambda), dZ_j^*(\mu)] = 0, \text{ for } \lambda \neq \mu \text{ and } (i = \{x, y\}) \tag{3.2.8}$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Substituting relations (3.2.5) and (3.2.6) in $\gamma_{xy}(h) = Cov[x(t+h), y(t)]$ we have that

$$\gamma_{xy}(h) = \int_{-\pi}^{\pi}\int_{-\pi}^{\pi} e^{-it\lambda}e^{-i(h+t)\mu}\mathrm{E}[dZ(\lambda)dZ^{*}(\mu)] \tag{3.2.9}$$

But based on the property of the orthogonal processes, we have that:

$$\gamma_{xy}(h) = \int_{-\pi}^{\pi} e^{ih\lambda}E[dZ_x(\lambda)dZ_y^{*}(\lambda)] \tag{3.2.10}$$

and of course similarly for $\gamma_{ii}(\lambda)$ $(i = \{x, y\})$:

$$\gamma_{ii}(\lambda) = \int_{-\pi}^{\pi} e^{ih\lambda}E[\left|dZ_i(\lambda)\right|^2] \tag{3.2.11}$$

Like in the univariate case where $\gamma_{ii}(h) = \int_{-\pi}^{\pi} e^{-ih\lambda}f_{ii}(\lambda)d\lambda$ (see definition 2.3.1), in the bivariate case wan obtain a similar relationship by inverting the definition relationship (3.2.3) which gives us:

$$\gamma_{xy}(h) = \int_{-\pi}^{\pi} e^{-ih\lambda}f_{xy}(\lambda)d\lambda \tag{3.2.12}$$

It is obvious that combining the relations (3.2.10) with (3.2.12) we end up with:

$$f_{xy}(\lambda)d\lambda = \mathrm{E}[dZ_x(\lambda)dZ_y^{*}(\lambda)] \tag{3.2.13}$$

which means that:

$$f_{xy}(\lambda) = f_{yx}^{*}(\lambda)\,. \tag{3.2.14}$$

Similarly we obtain:

$$f_{ii}(\lambda)d\lambda = \mathrm{E}[\left|dZ_i(\lambda)\right|^2] \tag{3.2.15}$$

These help us in the interpretation of the cross–spectral density, since $f_{xy}(\lambda)d\lambda$ represents the average value of the product of the coefficients of $e^{ih\lambda}$, in the spectral

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

representation of $\{x(t)\}$ and $\{y(t)\}$. This means that $f_{xy}(\lambda)d\lambda$ represents the $E[dZ_x(\lambda)dZ_y(\lambda)]$ which is actually the covariance between $dZ_x(\lambda)$ and $dZ_y(\lambda)$.

As in the bivariate case, the $f_{ii}(\lambda)d\lambda$, represent the average value of the square of the coefficient of $e^{ih\lambda}$ in the spectral representation of $\{x_i(t)\}$ i.e. $E[dZ_i(\lambda)]$ which is the variance of the $dZ_i(\lambda)$.

For bivariate series the spectral matrix can be defined as:

$$F(\lambda)=\begin{bmatrix} f_{xx}(\lambda) & f_{xy}(\lambda) \\ f_{yx}(\lambda) & f_{yy}(\lambda) \end{bmatrix} \tag{3.2.16}$$

3.3 Coherence and phase

As shown above, the $\{Z_i(\lambda), -\pi \le \lambda \le \pi\}, i=\{x,y\}$ are the orthogonal increment process in the spectral representation of the series $\{x(t)\}$ and $\{y(t)\}$. The squared absolute value of the correlation between $dZ_x(\lambda)$ and $dZ_y(\lambda)$ is called the squared coherence function $K_{xy}^2(\lambda)$.

$$\text{So } K_{xy}^2(\lambda)=\frac{\left|f_{xy}(\lambda)\right|^2}{[f_{xx}(\lambda)f_{yy}(\lambda)]}=K_{yx}^2(\lambda) \tag{3.3.1}$$

The resemblance with the ordinary correlation coefficient is obvious. Furthermore the coherence plays the role of a correlation coefficient at each frequency λ. As a correlation coefficient it is expected that $K_{xy}^2(\lambda)$ would lie between zero and one. Indeed the Cauchy-Schwarz inequality ensures that:

$$\left|f_{xy}(\lambda)\right|^2 \le f_{xx}(\lambda)f_{yy}(\lambda) \tag{3.3.2}$$

thus making $0 \le K_{xy}^2(\lambda) \le 1$ for $-\pi \le \lambda \le \pi$.

A high value of the coherence at a frequency λ, implies that the two processes are highly correlated at the specific frequency. In order to check at which frequency

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

bands this occurs, the coherence $K_{xy}^2(\lambda)$ is plotted against the frequency $-\pi \le \lambda \le \pi$, producing the so called coherence diagram.

Generally it can be proved (Brockwell and Davis, (1987)), that whenever $(x(t)\}$ and $(y(t)\}$ are related by a time invariant filter, that is:

$$y(t) = \sum_{j=-\infty}^{\infty} \psi_j \cdot x(t-j), \quad \text{where } \sum_j |\psi_j| < \infty \tag{3.3.3}$$

the coherence $K_{xy}^2(\lambda)$ equals to one, for every $-\pi \le \lambda \le \pi$.

Since the $f_{xy}(\lambda)$ is a complex valued function, it can be written as:

$$f_{xy}(\lambda) = C_{xy}(\lambda) - iQ_{xy}(\lambda) \tag{3.3.4}$$

where

$C_{xy}(\lambda) = \mathrm{Re}[f_{xy}(\lambda)]$

and

$Q_{xy}(\lambda) = -\mathrm{Im}[f_{xy}(\lambda)]$.

The above relations are actually the definitions of the cospectrum and the quadrature spectrum, so:

Definitions 3.3.1:

The cospectrum of the processes $\{x(t)\}$ and $\{y(t)\}$ is the function:

$C_{xy}(\lambda) = \mathrm{Re}[f_{xy}(\lambda)]$ and (3.3.5)

The quadrature spectrum of the processes $\{x(t)\}$ and $\{y(t)\}$ is the function:

$Q_{xy}(\lambda) = -\mathrm{Im}[f_{xy}(\lambda)]$. (3.3.6)

Alternatively, $f_{xy}(\lambda)$ can be expressed in a polar form, utilizing the cospectrum and the quadrature spectrum.:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

$$f_{xy}(\lambda) = A_{xy}(\lambda) \cdot e^{i\Phi_{xy}(\lambda)} \tag{3.3.7}$$

where:

$$A_{xy}(\lambda) = \left|f_{xy}(\lambda)\right| = \{C_{xy}^2(\lambda) + Q_{xy}^2(\lambda)\}^{1/2} \tag{3.3.8}$$

and

$$\Phi_{xy}(\lambda) = \tan^{-1}\{\frac{-Q_{xy}(\lambda)}{C_{xy}(\lambda)}\} \tag{3.3.9}$$

Definitions 3.3.2:

The function $A_{xy}(\lambda)$ is called the cross-amplitude spectrum and the function $\Phi_{xy}(\lambda)$ the phase spectrum.

By combining (3.3.1) with the definition relation (3.3.7) we can see that the squared coherence function $K_{xy}^2(\lambda)$ depends on the cross-amplitude spectrum, since:

$$K_{xy}^2(\lambda) = \frac{A_{xy}^2(\lambda)}{[f_{xx}(\lambda) f_{yy}(\lambda)]} \tag{3.3.10}$$

In order to interpret the $A_{xy}(\lambda)$ and $\Phi_{xy}(\lambda)$ we use the polar form of the $dZ_i(\lambda)$:

$$dZ_i(\lambda) = \left|dZ_i(\lambda)\right| \cdot e^{i\Phi_i(\lambda)}, \text{ with } i = \{x, y\}. \tag{3.3.11}$$

In this way the $\left|dZ_i(\lambda)\right|, i = \{x, y\}$, is the amplitude and $\Phi_i(\lambda)$ is the phase of the component with frequency λ, in the spectral representation of $\{x(t)\}$ or $\{y(t)\}$.

If we assume, for simplicity, that $\left|dZ_i(\lambda)\right|$ and $\Phi_i(\lambda)$ are independent random variables, ($i = \{x, y\}$) it is then apparent that:

$$E[dZ_x(\lambda) \cdot dZ_y^*(\lambda)] = E[\left|dZ_x(\lambda)\right| \cdot \left|dZ_y(\lambda)\right|] \cdot E[e^{i[\Phi_x(\lambda) - \Phi_y(\lambda)]}] \tag{3.3.12}$$

But from the definition of the $A_{xy}(\lambda)$ and $\Phi_{xy}(\lambda)$, it is clear that:

$$E[dZ_x(\lambda) \cdot dZ_y^*(\lambda)] = f_{xy}(\lambda) = A_{xy}(\lambda) \cdot e^{i\Phi_{xy}(\lambda)} \tag{3.3.13}$$

which makes:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

$$A_{xy}(\lambda) = E[|dZ_x(\lambda)| \cdot |dZ_y(\lambda)|] \text{ and} \tag{3.3.14}$$

$$\Phi_{xy}(\lambda) = E[e^{i[\Phi_x(\lambda)-\Phi_y(\lambda)]}] \tag{3.3.15}$$

This relation actually reveals that $A_{xy}(\lambda)$ is the average value of the product of the amplitudes of the components with frequency λ in the spectral representation of the processes $\{x(t)\}$ and $\{y(t)\}$, i.e. $E[|dZ_x(\lambda)| \cdot |dZ_y(\lambda)|]$.

Similarly, $\Phi_{xy}(\lambda)$ is the average value of the phase shift: $\{\Phi_x(\lambda) - \Phi_y(\lambda)\}$ between the components of frequency λ in the spectral representation of the processes $\{x(t)\}$ and $\{y(t)\}$, i.e. $E[e^{i[\Phi_x(\lambda)-\Phi_y(\lambda)]}]$.

3.4 Estimation of spectra and cross-spectra

In the estimation of the cross-spectral density function we can use the notions mentioned in the univariate case. A natural generalization will enable us to deal with the bivariate case. Let us examine this approach (Priestley, (1981)):

Let us suppose that N observations of a real valued stationary bivariate time series consisting of datasets: $X(t) = \{x(t), y(t)\}$ are available.

Let $\Gamma(h)$ be the covariance matrix:

$$\Gamma(h) = \begin{bmatrix} \gamma_{xx}(h) & \gamma_{xy}(h) \\ \gamma_{yx}(h) & \gamma_{yy}(h) \end{bmatrix} \tag{3.4.1}$$

with absolutely summable components and $F(h)$ the spectral matrix mentioned in section 3.2:

$$F(\lambda) = \begin{bmatrix} f_{xx}(\lambda) & f_{xy}(\lambda) \\ f_{yx}(\lambda) & f_{yy}(\lambda) \end{bmatrix} \tag{3.4.2}$$

The $F(\lambda)$ is defined as:

$$F(\lambda) = \frac{1}{2\pi} \sum_{h=-\infty}^{\infty} \Gamma(h) e^{-ih\lambda}, \quad -\pi \le \lambda \le \pi \tag{3.4.3}$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

As mentioned in the univariate case (see section 2.4), in order to define the multivariate periodogram we must utilize the Fourier frequencies.

So let $\omega_k = \frac{2\pi k}{n}$, with $-[(n-1)/2] \le k \le [n/2]$ be the Fourier frequencies. The discrete Fourier transform of the $X(t)$ is defined by:

$$J(\omega_k) = \frac{1}{\sqrt{n}} \sum_{t=1}^{N} X(t) \cdot e^{-it\omega_k} \tag{3.4.4}$$

The periodogram of the $X(t)$ is defined at each of the Fourier frequencies ω_k as a $2X2$ matrix:

$$I_n(\omega_k) = J(\omega_k) \cdot J^*(\omega_k) \tag{3.4.5}$$

Expanding the univariate case, the function $I_{xy}(\omega)$ is called the cross-periodogram and at the Fourier frequencies ω_k is has the value:

$$I_{xy}(\omega_k) = \frac{1}{\sqrt{n}} \left(\sum_{t=1}^{n} x(t) \cdot e^{-it\omega_k}\right) \cdot \frac{1}{\sqrt{n}} \left(\sum_{t=1}^{n} y(t) \cdot e^{-it\omega_k}\right)^* \tag{3.4.6}$$

The periodogram of a bivariate series in matrix form is the following:

$$I_n(\lambda) = \begin{bmatrix} I_n^{xx}(\lambda) & I_n^{xy}(\lambda) \\ I_n^{yx}(\lambda) & I_n^{yy}(\lambda) \end{bmatrix} \tag{3.4.7}$$

where for $i = j$, I_n^{ii} is the periodogram of $\{x(t)\}$ or $\{y(t)\}$, while for $i \neq j$, (thus $i = x, j = y$ or $i = y, j = x$) I_n^{ij} is the cross-periodogram between $\{x(t)\}$ and $\{y(t)\}$.

Asymptotic Properties of the periodogram

Theorem 3.4.1: If ω_j is any non-zero Fourier frequency and $\overline{X}_i = \frac{1}{n} \sum_{t=1}^{n} X_i(t)$, then:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

$$I_N(\omega_j) = \sum_{|k|<n} \Gamma(k) \cdot e^{-ik\omega_j} \tag{3.4.8}$$

where:

- if $k \geq 0$: $\hat{\Gamma}(k) = \frac{1}{n} \sum_{t=1}^{(n-\kappa)} [X(t+k) - \overline{X}_n] \cdot [X(t) - \overline{X}_n]'$, and (3.4.9)
- if $k < 0$: $\hat{\Gamma}(k) = \hat{\Gamma}(-k)$ (3.4.10)

The periodogram at frequency zero is:

$$I_n(0) = N \cdot \overline{X}_n \cdot \overline{X}_n' \tag{3.4.11}$$

***Theorem 3.4.2*:** If $\{X(t)\}$ is a stationary bivariate time series with mean μ and covariance matrices $\Gamma(h)$ having absolutely summable components then:

$$\text{i) } E[I_n(0)] \longrightarrow n \cdot \mu \cdot \mu' \longrightarrow 2\pi \cdot f(0) \tag{3.4.12}$$

$$\text{ii) } E[I_n(\omega)] \longrightarrow 2\pi \cdot f(\omega), \text{ if } \omega \neq 0 \tag{3.4.13}$$

where $f(\cdot)$ is the spectral matrix function of $\{X(t)\}$.

Regarding the asymptotic distribution and asymptotic covariances of the periodogram values, the following theorems are presented.

***Theorem 3.4.3*:** (Brockwell and Davis, (1987), page 431). Let us assume that $\{x(t)\}$ is a linear process:

$$x(t) = \sum_{k=-\infty}^{\infty} C_k \cdot Z(t-k) \qquad \text{with } \{Z(t)\} \sim IID(0, \Sigma) \text{ and } E[Z_i(t)] < \infty \tag{3.4.14}$$

where Σ is non-singular and the components of the matrices C_k satisfy:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

$$\sum_{k=-\infty}^{\infty} |C_k(i.j)| \cdot |k|^{1/2} < \infty \ , \quad \text{where} \quad i,j = \{x,y\} \text{ and} \tag{3.4.15}$$

Let $I_n(\lambda) = [I_n^{ij}(\lambda)]$, $i,j = \{x,y\}$, $-\pi \le \lambda \le \pi$ denote the periodogram of $\{Z(t)\}$. Then:

I) If $0 < \lambda_1 < \lambda_2 < \lambda_3 ... < \lambda_m < \pi$ then the matrices $\{I_n(\lambda_1), I_n(\lambda_2), ..., I_n(\lambda_m)\}$ converge jointly in distribution as $n -> \infty$ to independent random matrices, the k^{th} of which is distributed as $W_k \cdot W_k^*$ where $W_k \sim N(0, 2\pi \cdot f(\lambda_k))$ and f is the spectral matrix of $\{x(t)\}$.

II) If $\omega_j = 2\pi j/n \in [0,\pi]$, and $i = 1,2$ and $\omega_k = 2\pi k/n \in [0,\pi]$ then:

$$Cov[I^{pq}(\omega_j), I^{rs}(\omega_k)] = \tag{3.4.16}$$

$$= \begin{cases} (2\pi)^2 \{f_{pr}(\omega_j) \cdot f_{sq}(\omega_j) + f_{ps}(\omega_j) \cdot f_{qr}(\omega_j)\} + O(n^{-1/2}) & \omega_j = \omega_k = 0, or \pi \\ (2\pi)^2 f_{pr}(\omega_j) \cdot f_{sq}(\omega_j) + O(n^{-1/2}) & \text{if } 0 < \omega_j = \omega_k < \pi \\ O(n^{-1}) & \omega_j \neq \omega_k \end{cases}$$

where the terms $O(n^{-1/2})$ and $O(n^{-1})$ can be bounded uniformly in j and k by $c_1 n^{-1/2}$ and $c_1 n^{-1}$ respectively for some positive constants c_1 and c_2.

It would be useful to examine the above relations. The covariance between each estimator goes to zero, when dealing with different frequencies ω_k, ω_j, which implies that the spectrum estimators are independent in this case.

When dealing with the same frequencies we have::

$$Cov[I^{xx}(\omega), I^{yy}(\omega)] = (2\pi)^2 f_{xy}(\omega) \cdot f_{yx}(\omega) = (2\pi)^2 |f_{xy}(\omega)|^2$$

$$\text{if } 0 < \omega_j = \omega_k < \pi \tag{3.4.17}$$

while

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

$$Cov[I^{xx}(\omega), I^{yy}(\omega)] = (2\pi)^2 \{f_{xy}(\omega) \cdot f_{yx}(\omega) + f_{xy}(\omega) \cdot f_{xy}(\omega)\} = (2\pi)^2 \{|f_{xy}(\omega)|^2 + |f_{xy}(\omega)|^2\}$$

if $\omega_j = \omega_k = 0$ or π. (3.4.18)

Similarly we have that:

$$Cov[I^{xy}(\omega), I^{xy}(\omega)] = (2\pi)^2 f_{xx}(\omega) \cdot f_{yy}(\omega)$$

if $0 < \omega_j = \omega_k < \pi$ (3.4.19)

and

$$Cov[I^{xy}(\omega), I^{xy}(\omega)] = (2\pi)^2 \{f_{xx}(\omega) \cdot f_{yy}(\omega) + f_{xy}(\omega) \cdot f_{yx}(\omega)\}$$

if $\omega_j = \omega_k = 0$ or π (3.4.20)

and the covariance between a periodogram and a cross-periodogram is:

$$Cov[I^{xy}(\omega), I^{xy}(\omega)] = (2\pi)^2 f_{xx}(\omega) \cdot f_{yx}(\omega)$$

if $0 < \omega_j = \omega_k < \pi$ (3.4.21)

$$Cov[I^{xy}(\omega), I^{xy}(\omega)] = (2\pi)^2 \{f_{xx}(\omega) \cdot f_{yx}(\omega) + f_{xy}(\omega) \cdot f_{xx}(\omega)\} = 2 \cdot \{(2\pi)^2 \cdot f_{xx}(\omega) \cdot f_{yx}(\omega)\}$$

if $\omega_j = \omega_k = 0$ or π (3.4.22)

Smoothing the periodogram

In section 2.4 we introduced the smoothing of the periodogram for a univariate series. Similarly, in the bivariate case we will estimate the smoothed spectral density estimator $\hat{F}(\lambda)$. So, we consider the class of estimators having the form:

$$\hat{F}(\lambda_i) = \frac{1}{2\pi} \sum_{|\kappa| \le m_n} I_n(\lambda_{i+\kappa}) \cdot W_{m_n}(\kappa) \quad (3.4.23)$$

where $I_n(\lambda_i)$ is the periodogram,

W_{m_n} is a sequence of weight functions and,

Everything mentioned in section 2.4 about the sequence of weight functions W_b and the bandwidth b are also valid in the bivariate case. So we still consider the

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

estimator $\hat{F}(\lambda_i)$ as a weight average of the periodogram and responsible for the weighting is the spectral window $W()$.

It should be noted that generally we deal with processes that have unknown mean μ. The periodogram for those series, regardless of their dimension, is estimated for the mean corrected series $X(t)-\overline{X}(t)$. This causes no implications since the periodograms of the original and the zero-mean process, are identical in the Fourier frequencies with the exception of the value zero.

So in order to estimate the periodogram for $\lambda = 0$ we use the above formula:

$$\hat{F}(0) = \frac{1}{2\pi} \cdot \mathrm{Re}[W_b(0) I_n(\lambda_1) + 2\sum_{k=1}^{b} W_b(k) I_n(\lambda_{k+1})] \qquad (3.4.24)$$

From the above it is obvious that we have applied the same weight function for all of the four components of $I_n(\lambda)$, but is not necessary. It is possible to use different functions when the components $I_{ij}(\lambda)$ have different characteristics.

Like in section 2.4, the following theorem dictates the consistency of the estimator $\hat{F}(\lambda)$:

Theorem 3.4.4: For the estimate of the spectral density of a bivariate time series to be consistent, (Brockwell and Davis, (1987)), we impose the following conditions to the spectral window W():

a) $m_n \to \infty$ and $m/n \to 0$ as $n \to \infty$ (3.4.25)

b) $W(t) \geq 0$ for all t. (3.4.26)

c), $W(t) = W(-t)$ (3.4.27)

d) $\sum_{|t| \leq m_n} W(t) = 1$ and, (3.4.28)

e) $\sum_{|t| \leq m_n} W^2(t) \to 0$ as $n \to \infty$. (3.4.29)

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

If the above conditions hold, then:

$$\text{a) } \lim E\hat{f}(\lambda) = f(\lambda) \text{ when } n \to \infty \tag{3.4.30}$$

$$\text{b) } \lim \frac{Cov[\hat{f}_{ij}(\lambda), \hat{f}_{kl}(\omega)]}{\sum_{|i|\le m_n} W^2(i)} = \begin{cases} \hat{f}_{ik}(\lambda)\hat{f}_{kj}(\lambda) + \hat{f}_{il}(\lambda)\hat{f}_{jk}(\lambda) & \omega = \lambda = 0, or\pi \\ \hat{f}_{ik}(\lambda)\hat{f}_{lj}(\lambda) & \text{when } 0 < \omega = \lambda < \pi \\ 0 & \omega \neq \lambda \end{cases} \tag{3.4.31}$$

Finally, like in the univariate processes, small variance is associated with large bandwidths and large variance is associated with small bandwidths.

3.5 Estimation of phase and coherence

Recalling the definition of the phase spectra, relation (3.3.9), it is obvious that it can be estimated by (Brockwell and Davis, (1987)):

$$\hat{\Phi}_{ij}(\lambda) = \tan^{-1}\{\frac{-\hat{Q}_{ij}(\lambda)}{\hat{C}_{ij}(\lambda)}\} \tag{3.5.1}$$

where

$$\hat{C}_{ij}(\lambda) = \mathrm{Re}[\hat{f}_{ij}(\lambda)] \tag{3.5.2}$$

and

$$\hat{Q}_{ij}(\lambda) = -\mathrm{Im}[\hat{f}_{ij}(\lambda)]. \tag{3.5.3}$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

It is obvious that the phase spectra is a non-linear function of the real and imaginary part of the cross-spectra. If the coherence $K_{ij}^2(\lambda) > 0$ then it can be proved that:

$$\hat{\Phi}_{ij}(\lambda) \text{ is } AN(\Phi_{ij}(\lambda), A_n^2 A_{xy}^2(\lambda)[K_{xy}^{-2}(\lambda) - 1]/2) \tag{3.5.4}$$

$$\text{where } A_n^2(\lambda) = \sum_{|t| \le m_n} W^2(t) \tag{3.5.5}$$

So it is obvious that the asymptotic variance of $\hat{\Phi}_{ij}(\lambda)$ is large when the coherence $K_{xy}^2(\lambda)$ is small and becomes zero when the coherence reaches one.

Thus, at frequencies where the coherence is low, the estimates of the phase spectrum may have extremely large variance. This can be explained, since when the coherence is small, we are actually trying to estimate the average difference between the phase of two (effectively) independent complex valued random variables.

Furthermore, if n is large and $K_{xy}^2(\lambda) = 0$ the phase spectrum $\hat{\Phi}_{ij}(\lambda)$ is approximately uniformly distributed on $(-\pi, \pi)$. When $K_{xy}^2(\lambda) > 0$ the distribution of the $\hat{\Phi}_{ij}(\lambda)$ is no longer approximately uniform but it is the ratio of two asymptotically normal variables ($\hat{Q}_{ij}(\lambda)$ and $\hat{C}_{ij}(\lambda)$) with non-zero means. That is the reason why the distribution of $\hat{\Phi}_{ij}(\lambda)$ has a peak in $(-\pi, \pi)$.

Regarding the estimation of the coherence $K_{ij}(\lambda)$, we can use the estimators of the real and imaginary parts of the cross-spectra as well as the estimators of the cross-spectra itself. So:

$$\hat{K}_{ij}^2(\lambda) = \left| \frac{\hat{C}_{ij}^2(\lambda) + \hat{Q}_{ij}^2(\lambda)}{[\hat{f}_{ij}(\lambda)\hat{f}_{ij}(\lambda)]} \right|^{1/2} \tag{3.5.6}$$

It can be shown that if $K_{xy}^2(\lambda) > 0$ then:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

$$\left|\hat{K}_{xy}(\lambda)\right| \text{ is } AN(\left|K_{xy}(\lambda)\right|, A_n^2[1-K_{xy}^2(\lambda)]^2/2) \qquad (3.5.7)$$

Since $K_{xy}^2(\lambda)$ is actually a correlation coefficient it should be expected to have a distribution closely related to that of an ordinary correlation coefficient.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Chapter 4

Vector Autoregressive Processes

4.1 Basic definitions and assumptions

Generally the Vector Autoregressive Processes of order p, are models with the form (Lutkepohl, (1993)):

$$x(t) = \mu + A_1 \cdot x(t-1) + A_2 \cdot x(t-2) + ... + A_p x(t-p) + \varepsilon(t), \qquad (4.1.1)$$

with $t = 0, \pm 1, \pm 2, \pm 3, ...$

Where:

$x(t) = (x_1(t), x_2(t), ..., x_k(t))'$ is a $(k \times 1)$ random vector,

the A_i are fixed $(k \times k)$ coefficient matrices,

$\mu = (\mu_1, \mu_2, ..., \mu_k)'$ is a fixed $(k \times 1)$ vector of intercept terms and finally,

$\varepsilon(t) = (\varepsilon_1(t), \varepsilon_2(t), ...\varepsilon_k(t))'$ is a k-dimensional white noise process.

The later means that $E[\varepsilon(t)] = 0, E[\varepsilon(t) \cdot \varepsilon'(t)] = \Sigma_\varepsilon$ and finally that $E[\varepsilon(t) \cdot \varepsilon'(s)] = 0$ for $t \neq s$. The covariance matrix Σ_ε is assumed to be non-singular.

For simplicity reasons we will deal with the bivariate case of these multivariate processes, thus the case where k=2. Every conclusion regarding the bivariate processes is valid for any multivariate process as well.

Let us examine the relation (4.1.1) closely. Our interest lies in the investigation of the assumptions that must be made, in order for that model to be well-defined. When the order of the autoregressive process is one we have the following VAR(1) model:

$$x(t) = \mu + A_1 \cdot x(t-1) + \varepsilon(t) \qquad (4.1.2)$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

with $t = 0,\pm1,\pm2,\pm3,\ldots.$

If the generating process starts at the time point $t = 1$ then we have:

$$x(1) = \mu + A_1 \cdot x(0) + \varepsilon(1)$$

$$x(2) = \mu + A_1 \cdot x(1) + \varepsilon(2) = \mu + A_1 \cdot [\mu + A_1 \cdot x(0) + \varepsilon(1)] + \varepsilon(2) =$$

$$= \mu \cdot (I_2 + A_1) + A_1^2 \cdot x(0) + A_1 \cdot \varepsilon(1) + \varepsilon(2)$$

.....

.....

$$x(t) = \mu + A_1 \cdot x(t-1) + \varepsilon(t) = \mu \cdot (\mathrm{I}_2 + \mathrm{A}_1 + \ldots + \mathrm{A}_1^{t-1}) + A_1^t \cdot x(0) + \sum_{i=0}^{t-1} A_1^i \cdot \varepsilon(t-i) \qquad (4.1.3)$$

In the above relation it is clear that the 2x1 vector $x(t) = (x_1(t), x_2(t))'$, is uniquely determined by $x(0) = (x_1(0), x_2(0))'$, $\varepsilon(1) = (\varepsilon_1(1), \varepsilon_2(1))'$, $\varepsilon(2) = (\varepsilon_1(2), \varepsilon_2(2))', \ldots\ldots\ldots\ldots$, and $\varepsilon(t) = (\varepsilon_1(t), \varepsilon_2(t))'$.

Considering that the starting point of the process usually is the infinite past, the question is what kind of process would be consistent with the mechanism presented in the relation:

$$x(t) = \mu + A_1 \cdot x(t-1) + \varepsilon(t) = \mu \cdot (\mathrm{I}_2 + \mathrm{A}_1 + \ldots + \mathrm{A}_1^j) + A_1^{j+1} \cdot x(t-j-1) + \sum_{i=0}^{j} A_1^i \cdot \varepsilon(t-i) \qquad (4.1.4)$$

It can be proved (Lutkepohl, (1993)), that if all of the eigenvalues of the matrix A_1 have moduli less than one, the sequence A_1^i with $i = 0,1,2,\ldots.$ is absolutely summable, which makes the sum :

$\sum_{i=0}^{j} A_1^i \cdot \varepsilon(t-i)$ exists in mean square when $j \to \infty$.

Furthermore:

$$\mu \cdot (\mathrm{I}_2 + \mathrm{A}_1 + \ldots + \mathrm{A}_1^j) \to (I_2 - A_1)^{-1} \mu, \text{ when } j \to \infty. \qquad (4.1.5)$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

The only term of the relation (4.1.4) left unexamined is the $A_1^{j+1} \cdot x(t-j-1)$, but it is obvious that the quantity A_1^{j+1} converges to zero quite fast as $j \to \infty$, thus the above term can be omitted.

So summarizing, we can state that if all eigenvalues of the coefficient matrix A_1 have moduli less than one we can say that $x(t)$ is a well defined VAR(1) process:

$$x(t) = \mu' + \sum_{i=0}^{\infty} A_1^i \varepsilon(t-i), \; t = 0, \pm 1, \pm 2, \ldots \qquad (4.1.6)$$

where $\mu' = (\mathrm{I}_2 - \mathrm{A}_1)^{-1} \mu$.

4.2 Stable and stationary Vector Autoregressive Processes

From the previous section it is obvious that the condition for the eigenvalues of the VAR(1) process is crucial. That is why the existence of that condition characterizes the process as a stable one. So a VAR(1) process is called a stable one when all of the eigenvalues of the coefficient matrix A_1 have moduli less than one, or equivalently when:

$$\det(I_2 - A_1 z) \neq 0 \text{ for } |z| \leq 1. \qquad (4.2.1)$$

It should be noted that a process may be well defined even if the stability condition is not met, but we would not be able to consider the starting point of the process to be $-\infty$. So only for a stable process this is possible and that is the reason that these processes are preferred.

In section 4.1 a VAR(1) process was studied but it can be shown that these results can easily be extended is an order p model. This is due to the fact that any VAR(p) model can be written in an VAR(1) form (Lutkepohl, (1993)). So the stability condition (4.2.1) can be extended to:

$$\det(I_2 - A_1 \cdot z - A_2 z^2 - \ldots - A_p z^p) \neq 0 \quad \text{for } |z| \leq 1 \qquad (4.2.2)$$

This polynomial is called the *reverse characteristic polynomial* of the VAR(p) process. So in other words a VAR(p) process is stable if the reverse characteristic polynomial has no roots in and on the complex unit circle.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

The implications of the stability condition to the behavior of a VAR process resemble the ones for a stationary process that is, steady fluctuation around a mean value and time independent variance. This is due to the fact that:

Stationarity condition

A stable Vector Autoregressive Process $x(t)$, $t = 0,\pm1,\pm2,....$ is stationary

(The converse is not true, since there are stationary VAR processes that are not stable.)

4.3 Estimation of Vector Autoregressive Processes

In this section the estimation of a VAR(p) process will be discussed. There are two ways to achieve that, the classical Least Squares Estimation and the Maximum Likelihood Estimation. We begin with the:

Multivariate Least Squares Estimation (Lutkepohl, (1993), page 62):

Lets assume we have a sample of size N of a two dimensional time series: $x(1), x(2),....., x(N)$. We define:

$$X := (x(1), x(2),....., x(N)) \text{ which is a } 2\times N \text{ matrix,} \quad (4.3.1)$$

$$B := (\mu, A_1, A_2,.....A_p)' \text{ which is a } 2\times(2\cdot p+1) \text{ matrix,} \quad (4.3.2)$$

$$Z_t := (1, x(t),....., x(t-p+1))' \text{ which is a } 2(p+1)\times 1 \text{ matrix,} \quad (4.3.3)$$

$$Z := (Z_0, Z_0,....Z_{N-1}) \text{ which is a } 2(p+1)\times N \text{ matrix,} \quad (4.3.4)$$

$$\varepsilon^* := (\varepsilon_1, \varepsilon_2,....., \varepsilon_N) \text{ which is a } 2\times N \text{ matrix.} \quad (4.3.5)$$

Using the above notation the standard VAR(p) model:

$$x(t) = \mu + A_1 \cdot x(t-1) + A_2 \cdot x(t-2) + ... + A_p x(t-p) + \varepsilon(t), \quad (4.3.6)$$

with $t = 0,\pm1,\pm2,\pm3,....$.

can be written as:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

$$X = BZ + \varepsilon^* \tag{4.3.7}$$

It can be found that the multivariate estimator of B, the $\hat{B}$ is:

$$\hat{B} = XZ'(ZZ')^{-1} = (BZ + \varepsilon^*)Z'(ZZ')^{-1} = B + \varepsilon^* Z'(ZZ')^{-1} \tag{4.3.8}$$

This can be achieved by multiplying with $Z'(t-1)$ the following relation:

$x(t) = B \cdot Z(t) + \varepsilon(t)$.

Taking expectations we have:

$$E[x(t) \cdot Z'(t-1)] = B \cdot E(Z(t-1)Z'(t-1)]. \tag{4.3.9}$$

Since:

$$E[x(t) \cdot Z'(t-1)] = \frac{1}{N}\sum_{t=1}^{N} x(t) \cdot Z'(t-1) = \frac{1}{N} XZ' \text{ and,} \tag{4.3.10}$$

$$E(Z(t-1)Z'(t-1)] = \frac{1}{N}\sum_{t=1}^{N} Z(t-1) \cdot Z'(t-1) = \frac{1}{N} ZZ' \tag{4.3.11}$$

the relation (4.3.9) becomes:

$$\frac{1}{N} XZ' = \hat{B}\frac{1}{N} ZZ' \Rightarrow \hat{B} = XZ'(ZZ')^{-1} \tag{4.3.12}$$

Asymptotic properties

The consistency and the asymptotic normality of the Least Square Estimator $\hat{B}$ can be easily proved if the following hold:

i) The $\Gamma := \lim ZZ'/N$ exists and is non-singular and,

ii) $\frac{1}{\sqrt{N}}(Z \otimes I_2) \cdot \tilde{\varepsilon} \xrightarrow{N \to \infty} \mathrm{N}(0, \Gamma \otimes \Sigma_\varepsilon)$

where $\tilde{\varepsilon}$ is a column vector $2N \times 1$ with all the $\varepsilon(1), \varepsilon(2), ..., \varepsilon(N)$.

If $\varepsilon(t)$ is a normally distributed white noise, it can be proved that the consistency and the asymptotic normality of the Least Square Estimator for a stable normally distributed VAR(p) process, are ensured.

In most cases the mean value of the process is not known therefore is should be estimated by the vector of the sample means:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

$$\bar{x} = \frac{1}{N}\sum_{t=1}^{N} x(t) \qquad (4.3.13)$$

It can be proved that:

Theorem 4.3.1

If a VAR(p) process $x(t)$ is stable and $\varepsilon(t)$ is a standard white noise, then

$\sqrt{N}(\bar{x} - \mu) \xrightarrow{d} N(0, \Sigma_{\bar{x}})$, where

$\Sigma_{\bar{x}} = (I_2 - A_1 - .... - A_p)^{-1} \cdot \Sigma \cdot (I_2 - A_1 - ..... - A_p)'^{-1}$

In particular $\lim \bar{x} = \mu$.

The Yule-Walker Estimator

The Least Squares Estimators derives from the Yule-Walker equations, which imply that:

$\Gamma_x(0) = [A_1, ..., A_p] \cdot [\Gamma_x(-1), ..., \Gamma_x(-p)]' + \Sigma_\varepsilon$, for $h = 0$ and, (4.3.14)

$\Gamma_x(h) = [A_1, ..., A_p] \cdot [\Gamma_x(h-1), ..., \Gamma_x(h-p)]'$, for $h > 0$ (4.3.15)

Therefore A can be estimated by:

$A = [\hat{\Gamma}_x(1), ..., \hat{\Gamma}_x(p)] \cdot \hat{\Gamma}_x(0)^{-1}$ (4.3.16)

where the $\hat{\Gamma}_x(\cdot)$'s can be estimated using all of the available data (for details see Lutkepohl, (1993), page 62).

The above estimator has the same asymptotic properties with the Least Square one but has less attractive small sample properties and this is the reason why the first estimator is preferred.

Maximum Likelihood Estimation (Brockwell and Davis, (1987), page 417)

There is an alternative procedure besides the classical Least Square Estimation and that is the maximum likelihood one. It can be implemented when the distribution of the process is known.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Under the assumption that the VAR(p) process in question is a Gaussian one, with mean value zero and covariance matrices:

$$K(i,j) = E[x(i) \cdot x(j)] \tag{4.3.17}$$

the exact likelihood of the process $\{x(1), x(2), ..., x(N)\}$ can be estimated. Let us consider the $2N \times 1$ column observation vector $X := [x'(1), x'(2), ,.....x'(N)]$ and let $\hat{X} := [\hat{x}'(1), \hat{x}'(2), ,.....\hat{x}'(N)]$ be the one step predictors (for details see (Brockwell, (1987), page 412) . Then it can be proved that the likelihood of $\{x(1), x(2), ..., x(N)\}$ is:

$$L = (2\pi)^{-N} (\prod_{j=1}^{N} \det[V(j-1)])^{-1/2} \exp\{-\frac{1}{2} \sum_{j=1}^{N} [x(j) - \hat{x}(j)]' \cdot V^{-1}(j-1) \cdot [x(j) - \hat{x}(j)]\} \tag{4.3.18}$$

where $V(j)$ are the covariance matrices. The first partial derivatives of the likelihood can be calculated in order to obtain the estimators for the mean value, the coefficient matrices A_i and the covariance matrix Σ_ε. Nevertheless the estimation of the above is a difficult task due to the potentially large number of parameters involved in relation (4.3.18). Non linear optimization algorithms are implemented for that cause.

Asymptotic properties

When $x(t)$ is a stable Gaussian VAR(p) process, the Maximum Likelihood estimators of the mean value, the coefficient and covariance matrices, are consistent ones.

Furthermore the estimator of the mean value is asymptotically independent of the estimators of the coefficient and covariance matrix.

Similarly the estimator of the coefficient matrix is asymptotically independent of the other two estimators respectively.

ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΒΙΒΛΙΟΘΗΚΗ

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Chapter 5
Extreme VAR spectra

5.1 Spectrum of Vector Autoregressive models

Let us consider the VAR(p) model presented in section 4.1:

$$x(t) = \mu + A_1 \cdot x(t-1) + A_2 \cdot x(t-2) + \ldots\ldots\ldots + A_p x(t-p) + \varepsilon(t), \qquad (5.1.1)$$

with $t = 0, \pm 1, \pm 2, \pm 3, \ldots\ldots$.

Let us write the above model in a different way, assuming (without loss of generality) that the process has zero mean. Our model can be written as:

$$\Phi(B) \cdot x(t) = \varepsilon(t) \qquad (5.1.2)$$

where B denotes the backshift operator and:

$$\Phi(z) := I_2 - A_1 z - \ldots\ldots - A_p z^p \qquad (5.1.3)$$

$$x(t) := \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}, \qquad (5.1.4)$$

$$\varepsilon(t) := \begin{bmatrix} \varepsilon_1(t) \\ \varepsilon_2(t) \end{bmatrix}. \qquad (5.1.5)$$

The spectral density matrix of the above process (Brockwell and Davis, (1987), page 418) is:

$$F(\lambda) = \frac{1}{2\pi} \Phi^{-1}(e^{-i\lambda}) \cdot \Sigma \cdot \overline{\Phi}^{-1^T}(e^{-i\lambda}) \qquad (5.1.6)$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

5.2 Analyzing the characteristic polynomial

What we are interested in, is to analyze the spectral matrix in terms of the roots of the polynomials which are elements of $\Phi(z)$. In particular we are interested in the conditions under which we obtain a peak in the coherence and the spectrum.

As we recall from section 2.5, in the univariate case, if the root of the polynomial was close to the unit circle, a peak was produced whose size depend on the moduli of the root. So specifying the model by roots rather than coefficients enables us to give to it interesting features. Similarly we can produce interesting features in a multivariate model by determining $\Phi(z)$ in terms of its roots rather by its coefficients.

The first step is to investigate the relations and possible interactions between the spectral densities and the coherence. Conditions under which a peak in each density, will have an impact in the between them coherence, should be discovered. That way the expected behaviour of the coherence will be revealed.

In chapter 4 the VAR(p) model and it's characteristic polynomial were presented. Let us examine a VAR(p) model where:

$$x(t) = \mu + A_1 \cdot x(t-1) + A_2 \cdot x(t-2) + ... + A_p \cdot x(t-p) + \varepsilon(t), \qquad (5.2.1)$$

where $Cov\{\varepsilon(t)\} = \Sigma$

and it's characteristic polynomial:

$$\Phi(z) = I_2 - A_1 \cdot z - A_2 z^2 - ..... - A_p z^p = \begin{pmatrix} 1 - a_{xx}^1 z - .... - a_{xx}^p z^p & a_{xy}^1 z - .... - a_{xy}^p z^p \\ a_{yx}^1 z - .... - a_{yx}^p z^p & 1 - a_{yy}^1 z - .... - a_{yy}^p z^p \end{pmatrix} =$$

$$= \begin{pmatrix} \Phi_{xx}(z) & \Phi_{xy}(z) \\ \Phi_{yx}(z) & \Phi_{yy}(z) \end{pmatrix} \qquad (5.2.2)$$

where:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

$$A_i = \begin{pmatrix} a^i_{xx} & a^i_{xy} \\ a^i_{yx} & a^i_{yy} \end{pmatrix}, \; i = 1,2 \tag{5.2.3}$$

Since the spectral density matrix of the VAR(p) process is the following:

$$F(\lambda) = \begin{bmatrix} f_{xx}(\lambda) & f_{xy}(\lambda) \\ f_{yx}(\lambda) & f_{yy}(\lambda) \end{bmatrix} = \frac{1}{2\pi} \cdot \Phi^{-1}(e^{-i\lambda}) \cdot \Sigma \cdot \bar{\Phi}^{-1T}(e^{-i\lambda}) \tag{5.2.4}$$

We will assume in what follows, that the covariance matrix equals:

$$\Sigma = I \tag{5.2.5}$$

This restriction simplifies the following calculation without undermining its applicability. Nevertheless it should always kept in mind, that any deviation from this restriction will render all the conclusions deriving from now on, misleading.

$$F(\lambda) = \begin{bmatrix} f_{xx}(\lambda) & f_{xy}(\lambda) \\ f_{yx}(\lambda) & f_{yy}(\lambda) \end{bmatrix} = \frac{1}{2\pi} \cdot \begin{pmatrix} \Phi_{xx}(e^{-i\lambda}) & \Phi_{xy}(e^{-i\lambda}) \\ \Phi_{yx}(e^{-i\lambda}) & \Phi_{yy}(e^{-i\lambda}) \end{pmatrix}^{-1} \cdot I \cdot \begin{pmatrix} \bar{\Phi}_{xx}(e^{-i\lambda}) & \bar{\Phi}_{xy}(e^{-i\lambda}) \\ \bar{\Phi}_{yx}(e^{-i\lambda}) & \bar{\Phi}_{yy}(e^{-i\lambda}) \end{pmatrix}^{-1T} \tag{5.2.6}$$

For simplicity reasons we will just use the notation:

$$\Phi = \begin{pmatrix} \Phi_{xx} & \Phi_{xy} \\ \Phi_{yx} & \Phi_{yy} \end{pmatrix} \tag{5.2.7}$$

instead of the more complex one :

$$\Phi(e^{-i\lambda}) = \begin{pmatrix} \Phi_{xx}(e^{-i\lambda}) & \Phi_{xy}(e^{-i\lambda}) \\ \Phi_{yx}(e^{-i\lambda}) & \Phi_{yy}(e^{-i\lambda}) \end{pmatrix}. \tag{5.2.8}$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Our objective is to reach to a more detailed form of the $\Phi^{-1}(e^{-i\lambda})\cdot\Sigma\cdot\overline{\Phi}^{-1^{\mathrm{T}}}(e^{-i\lambda})$ product, in order to investigate the relationship and the interactivity between the Φ_{ij} $i,j=\{x,y\}$ terms. So:

$$\Phi^{-1}(e^{-i\lambda})\cdot\Sigma\cdot\overline{\Phi}^{-1^{\mathrm{T}}}(e^{-i\lambda})=\Phi^{-1}\cdot I\cdot\overline{\Phi}^{-1^{T}}=\left(\overline{\Phi}^{T}\cdot\Phi\right)^{-1}=$$

$$=\begin{pmatrix}\left|\Phi_{xx}\right|^2+\left|\Phi_{yx}\right|^2 & \overline{\Phi}_{xx}\Phi_{xy}+\Phi_{yy}\overline{\Phi}_{yx}\\ \Phi_{xx}\overline{\Phi}_{xy}+\overline{\Phi}_{yy}\Phi_{yx} & \left|\Phi_{yy}\right|^2+\left|\Phi_{xy}\right|^2\end{pmatrix}^{-1} \qquad (5.2.9)$$

and since: $\det(\overline{\Phi}^{T}\cdot\Phi)=\left|\Phi_{xx}\Phi_{yy}-\Phi_{xy}\Phi_{yx}\right|^2$ we have that:

$$2\cdot\pi\cdot F(\lambda)=\left(\overline{\Phi}^{T}\cdot\Phi\right)^{-1}=\begin{pmatrix}\dfrac{\left|\Phi_{yy}\right|^2+\left|\Phi_{xy}\right|^2}{\left|\Phi_{xx}\Phi_{yy}-\Phi_{xy}\Phi_{yx}\right|^2} & -\dfrac{\overline{\Phi}_{xx}\Phi_{xy}+\Phi_{yy}\overline{\Phi}_{yx}}{\left|\Phi_{xx}\Phi_{yy}-\Phi_{xy}\Phi_{yx}\right|^2}\\ -\dfrac{\Phi_{xx}\overline{\Phi}_{xy}+\overline{\Phi}_{yy}\Phi_{yx}}{\left|\Phi_{xx}\Phi_{yy}-\Phi_{xy}\Phi_{yx}\right|^2} & \dfrac{\left|\Phi_{xx}\right|^2+\left|\Phi_{yx}\right|^2}{\left|\Phi_{xx}\Phi_{yy}-\Phi_{xy}\Phi_{yx}\right|^2}\end{pmatrix}$$

(5.2.10)

The above matrix includes the spectra of the two series. The (1,1) element is actually the spectrum of the first series while the (2,2) is the spectrum of the second one.

So:

$$f_{xx}=\frac{\left|\Phi_{yy}\right|^2+\left|\Phi_{xy}\right|^2}{\left|\Phi_{xx}\Phi_{yy}-\Phi_{xy}\Phi_{yx}\right|^2} \qquad \text{and} \qquad f_{yy}=\frac{\left|\Phi_{xx}\right|^2+\left|\Phi_{yx}\right|^2}{\left|\Phi_{xx}\Phi_{yy}-\Phi_{xy}\Phi_{yx}\right|^2}$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

This means that f_{xx} has peaks at those roots of the determinant of $(\overline{\Phi}^T \cdot \Phi)$ which have modulus close to one, $(\det(\overline{\Phi}^T \cdot \Phi) = |\Phi_{xx}\Phi_{yy} - \Phi_{xy}\Phi_{yx}|^2)$ if these are not simultaneously also zeros of Φ_{yy} and Φ_{xy}.

Similarly f_{yy} has peaks at the zeros of the determinant of $(\overline{\Phi}^T \cdot \Phi)$, if these are not simultaneously the zeros of Φ_{xx} and Φ_{yx}.

5.3 Analyzing the coherence

We are now in the position to investigate the Coherence:

$$K_{xy}^2(\lambda) = K_{yx}^2(\lambda) = \frac{|f_{xy}(\lambda)|^2}{[f_{xx}(\lambda) f_{yy}(\lambda)]} \tag{5.3.1}$$

Once again we will omit the frequency λ and the above relation will become:

$$K_{xy}^2 = K_{yx}^2 = \frac{|f_{xy}|^2}{[f_{xx} f_{yy}]} = \frac{|\overline{\Phi}_{xx}\Phi_{xy} + \Phi_{yy}\overline{\Phi}_{yx}|^2}{(|\Phi_{yy}|^2 + |\Phi_{xy}|^2) \cdot (|\Phi_{xx}|^2 + |\Phi_{yx}|^2)} \tag{5.3.2}$$

But:

$$\begin{aligned}
&(|\Phi_{yy}|^2 + |\Phi_{xy}|^2) \cdot (|\Phi_{xx}|^2 + |\Phi_{yx}|^2) = \\
&= |\Phi_{yy}|^2 \cdot |\Phi_{xx}|^2 + |\Phi_{yy}|^2 \cdot |\Phi_{yx}|^2 + |\Phi_{xx}|^2 \cdot |\Phi_{xy}|^2 + |\Phi_{xy}|^2 \cdot |\Phi_{yx}|^2 = \\
&= (|\Phi_{yy}|^2 \cdot |\Phi_{yx}|^2 + |\Phi_{xx}|^2 \cdot |\Phi_{xy}|^2) + 2\,\mathrm{Re}(\Phi_{xx}\overline{\Phi}_{xy} \cdot \Phi_{yy}\overline{\Phi}_{yx}) + \\
&+ |\Phi_{yy}|^2 \cdot |\Phi_{xx}|^2 + |\Phi_{xy}|^2 \cdot |\Phi_{yx}|^2 - 2\,\mathrm{Re}(\Phi_{xx}\overline{\Phi}_{xy} \cdot \Phi_{yy}\overline{\Phi}_{yx}) = \\
&= |\overline{\Phi}_{xx}\Phi_{xy} + \Phi_{yy}\overline{\Phi}_{yx}|^2 + |\Phi_{yy}|^2 \cdot |\Phi_{xx}|^2 + |\Phi_{xy}|^2 \cdot |\Phi_{yx}|^2 - 2\,\mathrm{Re}(\Phi_{xx}\overline{\Phi}_{xy} \cdot \Phi_{yy}\overline{\Phi}_{yx}) = \\
&= |\overline{\Phi}_{xx}\Phi_{xy} + \Phi_{yy}\overline{\Phi}_{yx}|^2 + |\Phi_{yy}\Phi_{xx} - \Phi_{xy}\Phi_{yx}|^2
\end{aligned} \tag{5.3.3}$$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

So we have that the coherence:

$$K_{xy}^2 = K_{yx}^2 = \frac{|f_{xy}|^2}{[f_{xx}f_{yy}]} = \frac{\left|\bar{\Phi}_{xx}\Phi_{xy} + \Phi_{yy}\bar{\Phi}_{yx}\right|^2}{(|\Phi_{yy}|^2 + |\Phi_{xy}|^2)\cdot(|\Phi_{xx}|^2 + |\Phi_{yx}|^2)} =$$

$$= \frac{1}{1 + \left|\frac{\Phi_{yy}\Phi_{xx} - \Phi_{xy}\Phi_{yx}}{\bar{\Phi}_{xx}\Phi_{xy} + \Phi_{yy}\bar{\Phi}_{yx}}\right|^2} = \frac{1}{1 + \frac{\det(\bar{\Phi}^T\cdot\Phi)}{\left|\bar{\Phi}_{xx}\Phi_{xy} + \Phi_{yy}\bar{\Phi}_{yx}\right|^2}} \qquad (5.3.4)$$

The last line holding true for those z for which $\bar{\Phi}_{xx}\Phi_{xy} + \Phi_{yy}\bar{\Phi}_{yx} \neq 0$.

It is obvious that if $z_0 = \rho_0 \cdot e^{i\lambda_0}$ is a root of $\det(\bar{\Phi}^T \cdot \Phi)$, and the co-spectrum is not zero in a neighbourhood of $e^{i\lambda_0}$ (i.e. $\bar{\Phi}_{xx}\Phi_{xy} + \Phi_{yy}\bar{\Phi}_{yx} \neq 0$), the coherence reaches one as ρ_0 approaches one.

5.4 Extreme spectra and the roots of the polynomial of Φ

In this chapter our interest focuses on the conditions under which "extreme" spectra may be generated in the multivariate case. "Extreme" in this context means that as the modulus of z_0 approaches unity, $f_{xx}(\lambda_0)$ (and/or) $f_{yy}(\lambda_0)$ will tend to infinity, while in certain situations the coherence will tend to unity.

Whether such a peak will be present depends on whether this specific root $z_0 = \rho_0 \cdot e^{i\lambda_0}$ of $\det(\Phi^T \cdot \Phi)$ is also a root of $|\Phi_{yy}|^2$ and $|\Phi_{xy}|^2$, because it is the relation between them that decides whether a peak will be present in f_{xx} or not. That is why we will study cases concerning whether z_0 is simultaneously a root of one (or some) of the other polynomial, under which several types of "extreme" spectra will be obtained.

It is on such extreme spectra that we will be focusing in this section. In order to investigate this relation, the formula (5.2.10) should be at hand:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

$$2\cdot\pi\cdot F(\lambda)=\left(\overline{\Phi}^{T}\cdot\Phi\right)^{-1}=\begin{pmatrix}\dfrac{\left|\Phi_{yy}\right|^{2}+\left|\Phi_{xy}\right|^{2}}{\left|\Phi_{xx}\Phi_{yy}-\Phi_{xy}\Phi_{yx}\right|^{2}} & -\dfrac{\overline{\Phi}_{xx}\Phi_{xy}+\Phi_{yy}\overline{\Phi}_{yx}}{\left|\Phi_{xx}\Phi_{yy}-\Phi_{xy}\Phi_{yx}\right|^{2}}\\ -\dfrac{\Phi_{xx}\overline{\Phi}_{xy}+\overline{\Phi}_{yy}\Phi_{yx}}{\left|\Phi_{xx}\Phi_{yy}-\Phi_{xy}\Phi_{yx}\right|^{2}} & \dfrac{\left|\Phi_{xx}\right|^{2}+\left|\Phi_{yx}\right|^{2}}{\left|\Phi_{xx}\Phi_{yy}-\Phi_{xy}\Phi_{yx}\right|^{2}}\end{pmatrix} \tag{5.4.1}$$

So we will be looking at a root of the determinant at $z_0=\rho_0\cdot e^{i\lambda_0}$,(since the roots of $\det(\Phi^{T}\cdot\Phi)$ are decisive for our study) and will be assuming that $\rho_0\approx 1$. Then we will examine the impact of z_0 being a root of the Φ_{ij}'s $i,j\in\{x,y\}$, on the spectra of the two series and the between them coherence.

First, the aspect of the order of z_0 should be clarified. It is obvious that if: $r_d, r_{xx}, r_{xy}, r_{yx}, r_{yy}$ are the orders of z_0 as a root for the $\det(\Phi), \Phi_{xx}, \Phi_{xy}, \Phi_{yx}, \Phi_{yy}$ respectively, taking under consideration the relations (5.3.4) and (5.4.1) we have that:

- as $|z_0|\to 1$ then $f_{xx}(\lambda_0)\to\infty$ if and only if $r_d>\min(r_{yy},r_{xy})$
- as $|z_0|\to 1$ then $f_{yy}(\lambda_0)\to\infty$ if and only if $r_d>\min(r_{xx},r_{yx})$
- as $|z_0|\to 1$ the $K_{xy}^2\to 1$ if and only if $r_d>order_{(z_0)}(\Phi_{xx}\cdot\overline{\Phi}_{xy}+\overline{\Phi}_{yy}\cdot\Phi_{yx})$.

In the following we will think of a fixed root z_0 with $|z_0|\approx 1$ and will refer to the above situations as "peak in f_{xx}", "peak in f_{yy}" and "coherence reaching one".

From the above it is obvious that if the order of $z_0=\rho_0\cdot e^{i\lambda_0}$ as root of the determinant is relatively large compared to the order of z_0 at the Φ_{ij}'s $i,j=\{x,y\}$, a peak will be present in both f_{xx}, f_{yy} and the coherence will approach one. So the cases of most interest are the ones where the order of the determinant does not meet with all of the above conditions.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

We will analyze exhaustively the cases where $r_d = 1$ and $r_{xx}, r_{xy}, r_{yx}, r_{yy} \leq 1$. It turns out that under these assumptions the values of $r_d, r_{xx}, r_{xy}, r_{yx}, r_{yy}$ are sufficient in order to describe the behaviour of the spectra at λ_0. Therefore no assumptions on the $order_{(z_0)}(\Phi_{xx} \cdot \bar{\Phi}_{xy} + \bar{\Phi}_{yy} \cdot \Phi_{yx})$ needs to be made.

Later on, the cases where the order of z_0 as root of the determinant is allowed to take higher values, will also be mentioned. There, the behaviour of the spectra will in general depend on the $order_{(z_0)}(\Phi_{xx} \cdot \bar{\Phi}_{xy} + \bar{\Phi}_{yy} \cdot \Phi_{yx})$. Nevertheless, for certain situations a classification was made which will be presented below, giving many interesting results about the relation of the high values of a VAR's spectrum with its coherence.

The conclusions will be presented after the classification, which was made depending on whether z_0 is the root of the Φ_{ij}'s or not. Starting with the Φ_{xx} and Φ_{yy}, we can create four subcategories and within each one , the order of z_0 for the Φ_{ij}'s, $i, j = \{x,y\}$ can be either one or zero. So we have:

Investigating the cases where $r_d = 1$ *and* $r_{xx} = r_{yy} = r_{xy} = r_{yx} = 0$ or 1

First category: $\Phi_{xx}(z_0) = \Phi_{yy}(z_0) = 0$

If $\Phi_{xy}(z_0) = 0$, $\Phi_{yx}(z_0) \neq 0$ and $r_d = r_{xx} = r_{yy} = r_{xy} = 1$, $r_{yx} = 0$, then a peak is observed in f_{yy} only, while the coherence does not reach one. In a similar way when $\Phi_{xy}(z_0) \neq 0$, $\Phi_{yx}(z_0) = 0$ and $r_d = r_{xx} = r_{yy} = r_{yx} = 1$, $r_{xy} = 0$ only f_{xx} has a peak and this time also the coherence does not reach its maximum value. The case where $\Phi_{xy}(z_0) = 0$, $\Phi_{yx}(z_0) = 0$ and $r_d = r_{yx} = r_{xx} = r_{yy} = r_{xy} = 1$, can not be considered, since that could only be possible if the order of z_0 as a root of the determinant was two.

Finally the case where $\Phi_{xy}(z_0) \neq 0$, $\Phi_{yx}(z_0) \neq 0$ and $r_d = r_{xx} = r_{yy} = 1$, $r_{xy} = r_{yx} = 0$ is not a valid one since that way z_0 could not be the root of the determinant.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Second category: $\Phi_{xx}(z_0)=0$, $\Phi_{yy}(z_0)\neq 0$

If $\Phi_{xy}(z_0)=0$, $\Phi_{yx}(z_0)\neq 0$ and $r_d=r_{xx}=r_{xy}=1$, $r_{yy}=r_{yx}=0$, then both f_{xx} and f_{yy} have a peak and coherence reaches its maximum value. This behaviour is not observed when $\Phi_{xy}(z_0)\neq 0$, $\Phi_{yx}(z_0)=0$ and $r_d=r_{xx}=r_{yx}=1$, $r_{yy}=r_{xy}=0$ where only a peak at f_{xx} is present and the coherence does not reach one. . Exactly the same behaviour, (a peak only in f_{xx}) is present when $\Phi_{xy}(z_0)=0$, $\Phi_{yx}(z_0)=0$ and $r_d=r_{xx}=r_{xy}=r_{yx}=1$, $r_{yy}=0$.

Once again the above three are the only valid cases, since the case where $\Phi_{xy}(z_0)\neq 0$, $\Phi_{yx}(z_0)\neq 0$ and $r_d=r_{xx}=r_{yy}=1$, $r_{xy}=r_{yx}=0$ is not a valid one since that way z_0 could not be the root of the determinant.

Third category: $\Phi_{xx}(z_0)\neq 0$, $\Phi_{yy}(z_0)=0$

This category is symmetrical to the second one, so in general we expect that both f_{xx} and f_{yy} should have a peak and that the coherence should reach one when $\Phi_{xy}(z_0)\neq 0$, $\Phi_{yx}(z_0)=0$ and $r_d=r_{yy}=r_{yx}=1$, $r_{xx}=r_{xy}=0$. Similarly, in a perfect symmetry with the second case, when $\Phi_{xy}(z_0)=0$, $\Phi_{yx}(z_0)\neq 0$ and $r_d=r_{yy}=r_{xy}=1$, $r_{xx}=r_{yx}=0$ a peak should be observed in the f_{yy} while once again the coherence should not reach one.

It is obvious that the case where $\Phi_{xy}(z_0)=0$, $\Phi_{yx}(z_0)=0$ and $r_d=r_{yx}=r_{yy}=r_{xy}=1$, $r_{xx}=0$ is again similar to the previous one, thus a peak is observed only in f_{yy} while the coherence does not reach unity.

Finally, the case where $\Phi_{xy}(z_0)\neq 0$, $\Phi_{yx}(z_0)\neq 0$ and $r_d=r_{xx}=r_{yy}=1$, $r_{xy}=r_{yx}=0$ is not a valid one for the same reasons mentioned in the previous section.

Fourth category: $\Phi_{xx}(z_0)\neq 0$, $\Phi_{yy}(z_0)\neq 0$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

This category if the simplest of all, since when: ($\Phi_{xy}(z_0)=0$ and $\Phi_{yx}(z_0)=0$) or ($\Phi_{xy}(z_0)\neq 0$ and $\Phi_{yx}(z_0)=0$) or ($\Phi_{xy}(z_0)=0$ and $\Phi_{yx}(z_0)\neq 0$) it is obvious that z_0 cannot be the root of the determinant. Thus the only valid case is the one in which $\Phi_{xy}(z_0)\neq 0$ and $\Phi_{yx}(z_0)\neq 0$, where z_0 can be the root of the determinant only if $\Phi_{xx}(z_0)\Phi_{yy}(z_0)=\Phi_{xy}(z_0)\Phi_{yx}(z_0)$. In this case a peak is observed in the spectra of both series and their coherence reaches one.

The above holds true because our assumptions imply:

$$\overline{\Phi}_{xx}(z_0)\Phi_{xy}(z_0)+\Phi_{yy}(z_0)\overline{\Phi}_{yx}(z_0)\neq 0 \qquad (5.4.2)$$

To prove this , we rewrite our restrictions:

$$\Phi_{xx}(z_0)\neq 0\ ,\ \Phi_{yy}(z_0)\neq 0\ ,\ \Phi_{xy}(z_0)\neq 0\ ,\ \Phi_{yx}(z_0)\neq 0 \quad \text{and} \qquad (5.4.3)$$

$$\Phi_{xx}(z_0)\Phi_{yy}(z_0)=\Phi_{xy}(z_0)\Phi_{yx}(z_0) \qquad (5.4.4)$$

So we are examining whether it is possible to have:

$$\overline{\Phi}_{xx}(z_0)\Phi_{xy}(z_0)+\Phi_{yy}(z_0)\overline{\Phi}_{yx}(z_0)=0 \qquad (5.4.5)$$

from (5.4.4) we have that:

$$\Phi_{xx}(z_0)=\frac{\Phi_{xy}(z_0)\Phi_{yx}(z_0)}{\Phi_{yy}(z_0)} \qquad (5.4.6)$$

so if indeed relation (5.4.4) is valid, replacing (5.4.6) in (5.4.5) leads to:

$$\frac{\Phi_{xy}(z_0)\Phi_{yx}(z_0)}{\Phi_{yy}(z_0)}\cdot\Phi_{xy}(z_0)+\Phi_{yy}(z_0)\overline{\Phi}_{yx}(z_0)=0$$

$$\left|\Phi_{xy}(z_0)\right|^2\overline{\Phi}_{yx}(z_0)+\left|\Phi_{yy}(z_0)\right|^2\overline{\Phi}_{yx}(z_0)=0$$

$$\left(\left|\Phi_{xy}(z_0)\right|^2+\left|\Phi_{yy}(z_0)\right|^2\right)\cdot\overline{\Phi}_{yx}(z_0)=0$$

which is not possible, since for the case considered here, all $\Phi_{ij}(z_0)\neq 0$.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

These were all the possible combination given the specified restriction, regarding the orders r_d , r_{xx} , r_{yy} , r_{xy} , r_{yx}. The results from this classification are summarized in the following theorem.

Theorem 5.4.1: If $z_0 = \rho_0 \cdot e^{i\lambda_0}$ has $r_d(z_0) \leq 1$, $r_{xx} = r_{yy} = r_{xy} = r_{yx} = 0$ or 1 , then as $\rho_0 \to 1$ the following cases may arise:

Case A: $f_{xx}(\lambda_0) \to \infty$, $f_{yy}(\lambda_0) \to \infty$ and $K_{xy}(\lambda_0) \to 1$ simultaneously. This may arise if:

- z_0 is not a root of any of the other polynomials: Φ_{xx} , Φ_{yy} , Φ_{xy} , Φ_{yx} or,
- z_0 is a root of the polynomials driving exactly one of the series x or y (ie: (Φ_{xx} and Φ_{xy}) or (Φ_{yy} and Φ_{yx})) , which consist of exactly one line of the Φ matrix.

Case B: Only one of the $f_{xx}(\lambda_0)$, $f_{yy}(\lambda_0)$ will tend to ∞, while the coherence will not tend to unity. This behaviour is expected in any other case.

□

Note: Under the conditions of the above theorem, the coherence reaches one at z_0 if and only if both spectra f_{xx} and f_{yy} have a peak at z_0. This means that a high correlation between two series in a specific frequency occurs if and only if, this frequency is strongly represented in both series. We will see in the next section that this need not be the case when the order of the z_0 as root of the $\det(\Phi)$ is equal or greater than two.

Let us now proceed with the examination of a more relaxed condition.

The case $r_d \geq 2$

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Let us recall relation (5.3.4) which gives us the coherence K_{12}^2 in relation with the determinant $\det(\bar{\Phi}^T \cdot \Phi)$ and the polynomial $\bar{\Phi}_{xx}\Phi_{xy} + \Phi_{yy}\bar{\Phi}_{yx}$ and relation (5.4.1) which presents the relation of the f_{ij} 's $i,j=\{x,y\}$ with the determinant.

So, relation (5.3.4) proofs that the behaviour of the coherence depends on the order of z_0 as root of the polynomial $\Phi_{xx} \cdot \bar{\Phi}_{xy} + \bar{\Phi}_{yy} \cdot \Phi_{yx}$. As mentioned before, when $r_d > order_{(z_0)}(\Phi_{xx} \cdot \bar{\Phi}_{xy} + \bar{\Phi}_{yy} \cdot \Phi_{yx})$ the coherence will approach unity at λ_0.

The theorem below, concerns a useful particular case. It states that (if $r_d \geq 2$) it is possible to have peaks at both individual spectra, while the coherence does not approach one, something which was not possible for $r_d = 1$.

Theorem 5.4.2: If $\min(r_{xx} + r_{xy}\ ,\ r_{yy} + r_{yx}) \geq r_d$ then the coherence does not reach one at λ_0. While if $r_d > \min(r_{yy}, r_{xy})$ and $r_d > \min(r_{xx}, r_{yx})$ then the individual spectra f_{xx} , f_{yy} will tend to infinity at λ_0.

□

The above can be easily satisfied if, for example, we set: $\Phi_{xx}(z_0) = \Phi_{yy}(z_0) = \Phi_{xy}(z_0) = \Phi_{yx}(z_0) = 0$. Then $order_{(z_0)}(\Phi_{xx} \cdot \bar{\Phi}_{xy} + \bar{\Phi}_{yy} \cdot \Phi_{yx}) \geq 2$ while $r_d = 2$, thus we will obtain peaks at f_{xx} and f_{yy} and the coherence will not approach unity.

The converse, as the following theorem states, is not possible. This is very interesting, because it is possible to have a two dimensional VAR model with a strong presence of a specific frequency λ_0 in both its series, without the two series to be highly 'correlated' at λ_0.

Theorem 5.4.3: When $|z_0| \to 1$, it is impossible for the coherence to reach unity if even one of the spectra f_{xx} or f_{yy} does not have a peak at λ_0.

This is easily understood since when, for example:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

$f_{xx}(\lambda_0) \nrightarrow \infty$ and $f_{yy}(\lambda_0) \rightarrow \infty$ then, (5.4.7)

$r_d \leq \min(r_{yy}, r_{xy})$, (5.4.8)

which means that we also have:

$r_d \leq \min(r_{yy} + r_{yx}, r_{xy} + r_{xx})$ thus, (5.4.9)

$r_d \leq \min(r_{yy} + r_{yx}, r_{xy} + r_{xx}) \leq order(\Phi_{xx} \cdot \bar{\Phi}_{xy} + \bar{\Phi}_{yy} \cdot \Phi_{yx})$. (5.4.10)

From the above relation it is obvious that the coherence can not reach unity when (5.4.7) stands.

□

What theorem 5.4.3 actually states, is that (when the restriction $r_d \geq 2$ is met) one could not encounter a VAR model in which the coherence reaches one for a specific frequency λ_0, while even one of the spectra f_{ii}'s $i = \{x, y\}$ does not tend to infinity (i.e has a strong peak) at the same frequency λ_0. So a coherence reaching its maximum value, indicates high powered spectrum for both series.

Another issue to be clarified is that the above theorems apply only to VAR models. This is obvious when examining the following remark (Brockwell and Davis 1987 page 422) which states that:

When two series $x(t)$ and $y(t)$ have squared coherence $K_{xy}^2(\lambda)$ and if linear filters are applied to each process giving:

$X(t) = \sum_{j=-\infty}^{\infty} a_j \cdot x(t-j)$ and, (5.4.11)

$Y(t) = \sum_{j=-\infty}^{\infty} b_j \cdot y(t-j)$ (5.4.12)

where $\sum_j |a_j| < \infty$ and $\sum_j |b_j| < \infty$ then,

the $X(t)$ and $Y(t)$ have the same squared coherence $K_{xy}^2(\lambda) = K_{XY}^2(\lambda)$.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

From the above remark it is obvious that if we consider a two dimensional (non VAR) time series with f_{xx} and f_{yy} having strong peaks at λ_0 and the coherence reaching one:

$$f_{xx}(\lambda_0) \to \infty \ , \ f_{yy}(\lambda_0) \to \infty \text{ and } K^2_{xy}(\lambda_0) \to 1 \qquad (5.4.13)$$

and apply an appropriately selected time invariant linear filter to one of the series, we can obtain $f_{xx}(\lambda_0) \not\to \infty$ without changing the between them coherence. This means that we could have:

$$f_{xx}(\lambda_0) \not\to \infty \ , \ f_{yy}(\lambda_0) \to \infty \text{ and } K^2_{xy}(\lambda_0) \to 1 \qquad (5.4.14)$$

which contradict the theorems stated before for VAR models.

Given the above conclusions, we are now in the position to proceed to the definition of a VAR model , based on its given roots.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Chapter 6

An algorithm for VAR model specification

6.1 Calculating the coefficients of the characteristic polynomial

The next step in our analysis is to create an algorithm that will accept as input the roots of the VAR's polynomial matrix and produce an output with the coefficients of the VAR's characteristic matrix. Such an algorithm will enable us to create a two dimensional VAR model with pre-specified extreme features of the desired type. With the proper manipulation of the roots of the VAR's polynomial, many interesting models may be constructed.

Let us analyze what that algorithm should be able to do. First of all let's consider a two dimensional VAR(p) model like the one presented in (5.2.1) with the unity matrix I_2 as a covariance matrix of the innovations and zero mean values. That is $\Sigma = I_2$ and $\mu=0$.

Objective:

One should always keep in mind that our goal is to define an algorithm which allows the definition of VARs with pre-specified structures, that is strong peaks for the one or the other spectrum and coherence close to one, at certain frequencies.

It is easy to define a VAR by specifying Φ_{xx} , Φ_{yy} , Φ_{xy} , Φ_{yx} via their roots, but in such approach the user would not control directly the roots of the determinant, which are of most importance when we wish to decide the behaviour of the spectra and the coherence.

So, we need an algorithm where the roots of the determinant are given as input, while, inevitably, the roots of the Φ_{xx} , Φ_{yy} , Φ_{xy} , Φ_{yx} will not all be free. The nature of the restrictions imposed to them will be analyzed in the following.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

The idea is to:

a) specify the roots of the determinant (in such a way that their modulus will be greater than one to ensure stationarity),

b) specify the roots of Φ_{xx} , Φ_{yy} except from one of the roots of the Φ_{yy}, which may be eventually added automatically by the algorithm (see step 2 below) and finally,

c) calculate the roots of $\Phi_{xy} \cdot \Phi_{yx}$ which they will be allocated accordingly, by the user, to Φ_{xy} and Φ_{yx} . That distribution is not a random one, since we intend to create the desired structure of the spectra, using the theorems mentioned in the previous section.

So it is obvious that pre-specified spectra and coherence structures should characterize our model, thus the appropriate coefficients $\{a_{i,j}^{k}\}$ with $i, j, k = 1,2$ mentioned in (5.2.3), should be the output of our algorithm.

That will be the result from the utilization of the following steps. In the following lets us consider:

Polynomial	*Number of roots*	Roots
$\det(\bar{\Phi}^{T} \cdot \Phi)$	$n_{\det}$	z_i
Φ_{xx}	$n_{\Phi_{xx}}$	w_i
Φ_{yy}	$n_{\Phi_{yy}}$	s_i
Φ_{xy}	$n_{\Phi_{xy}}$	p_i
Φ_{yy}	$n_{\Phi_{yx}}$	q_i

Table 6.1.1: The polynomial's roots and their representation

The following table shows the given and the retrieved information:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Input	z_i , $i=1,2,...n_{\det}$, w_i , $i=1,2,...n_{\Phi_{xx}}$, s_i , $i=1,2,...,(n_{\Phi_{yy}}-1)$
Output	$\{a_{i,j}^k\}$ with $i,j,=1,2$ and $k=1,2,...p$ from the A_k matrices

Table 6.1.2: The input and the output for the algorithm's utilization

Step one:

The roots z_i , $i=1,2,...n_{\det}$ of the determinant $\det(\Phi)$ will be given. From these roots the coefficients of that polynomial $\det(\Phi)=\Phi_{xx}\cdot\Phi_{yy}-\Phi_{xy}\cdot\Phi_{yx}$ will be calculated. It is obvious that for all $\det(\Phi)$, Φ_{xx} , Φ_{yy} the constant term should be one.

Step two:

The roots w_i , $i=1,2,...n_{\Phi_{xx}}$ and s_i , $i=1,2,...n_{\Phi_{yy}}$ of the Φ_{xx} and Φ_{yy} polynomials respectively, will be given. Similarly their coefficients will be calculated.

It is obvious from the definition of the determinant $\det(\Phi)=\Phi_{xx}\cdot\Phi_{yy}-\Phi_{xy}\cdot\Phi_{yx}$, that knowing the coefficients of the $\det(\Phi)$, Φ_{xx} and Φ_{yy} we actually know the coefficients of the product $\Phi_{xy}\cdot\Phi_{yx}$ since:

$$\Phi_{xy}\cdot\Phi_{yx}=\Phi_{xx}\cdot\Phi_{yy}-\det(\Phi) \qquad (6.1.1)$$

Since the polynomials Φ_{xy} and Φ_{yx} should not have constant terms it is obvious that a restriction about their product $\Phi_{xy}\cdot\Phi_{yx}$ is imperative. Since both of them have a constant term equal to zero, their product should not have either a constant or a linear term.

From relation (6.1.1) we realize that manipulating Φ_{xx} and Φ_{yy} allows us to alter the coefficients of the product $\Phi_{xy}\cdot\Phi_{yx}$. So we can define one of those polynomials (say Φ_{yy}) such that no constant and linear term would be present in the product in question.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

The problem actually exists only for the linear term, since all $\det(\Phi)$, Φ_{xx} and Φ_{yy} have constant terms equal to one, thus the term $\Phi_{xx}\cdot\Phi_{yy}-\det(\Phi)$ has constant term equal to zero.

Regarding the linear term, if:

$C_{\det(\Phi)}$ is the linear term of the $\det(\Phi)$,

$C_{\Phi_{xx}}$ is the linear term of the Φ_{xx} and,

$C_{\Phi_{yy}}$ is the linear term of the Φ_{yy},

we need to make sure that:

$$C_{\Phi_{yy}} = C_{\det(\Phi)} - C_{\Phi_{xx}} \tag{6.1.2}$$

Since Φ_{yy} has constant term equal to unity it can be written as:

$$\Phi_{yy} = (1-\frac{z}{s_1})\cdot(1-\frac{z}{s_2})\cdots(1-\frac{z}{s_{n_{\Phi_{yy}}}}) =$$

$$= 1-(\frac{1}{s_1}+\frac{1}{s_2}+....+\frac{1}{s_{n_{\Phi_{yy}}}})\cdot z+(..)\cdot z^2+...(...)\cdot z^{n_{\Phi_{yy}}} \tag{6.1.3}$$

Considering the definition of $C_{\Phi_{yy}}$ given above, it is obvious that:

$$C_{\Phi_{yy}} = -(\frac{1}{s_1}+\frac{1}{s_2}+....+\frac{1}{s_{n_{\Phi_{yy}}}}) \tag{6.1.4}$$

Replacing (6.1.4) in (6.1.2) we have:

$-(\frac{1}{s_1}+\frac{1}{s_2}+....+\frac{1}{s_{n_{\Phi_{yy}}}}) = C_{\det(\Phi)} - C_{\Phi_{xx}}$, thus

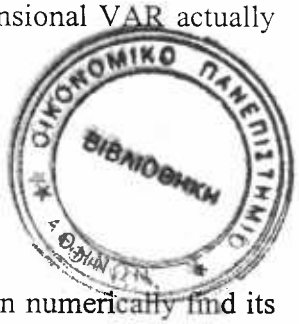
$$\frac{1}{s_{n_{\Phi_{yy}}}} = C_{\Phi_{xx}} - C_{\det(\Phi)} - \frac{1}{s_1} - \frac{1}{s_2} - .... - \frac{1}{s_{(n_{\Phi_{yy}}-1)}} \tag{6.1.5}$$

So, from the total of $n_{\Phi_{yy}}$ roots of Φ_{yy} we will initially give only the first $(n_{\Phi_{yy}}-1)$ (as stated in 6.1.2) and then calculate the last one with the help of (6.1.5).

Furthermore in the case where $\Phi_{xy}\cdot\Phi_{yx}\equiv 0$, the user should be able to set one of the polynomials (or both) to zero and input arbitrarily roots for the other one. A

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

typical case where that may be at hand, is when our two dimensional VAR actually consists of two uncorrelated series, meaning that $\Phi_{xy} = \Phi_{yx} = 0$.



We may now proceed to the next step where,

Step three:

Knowing the coefficients of the polynomial $\Phi_{xy} \cdot \Phi_{yx}$ we can numerically find its roots. These will be a unification of Φ_{xy}'s and Φ_{yx}'s roots. The user may then arbitrarily allocate the p_i , $i = 1,2,...,n_{\Phi_{xy}}$ roots to the Φ_{xy} polynomial and the rest q_i , $i = 1,2,...,n_{\Phi_{yx}}$ to the Φ_{yx} one. For each one of them its coefficients will be calculated.

By now the polynomial matrix (5.2.7) is well defined and its polynomials have the desired properties. This means that we have completed the creation of a two dimensional VAR(p) model based entirely on the roots of its polynomial matrix.

For example, let us create a VAR in which only the f_{yy} should have a peak at $\pi/2$. Examining section 5.4 we see that one possible way to do so is by giving the $z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{2}}$ as a single root to the $\det(\Phi)$, Φ_{xx} , Φ_{yy} and Φ_{xy}. So we:

- Give root z_0 to the $\det(\Phi)$, Φ_{xx} , Φ_{yy} and obtain from the algorithm the roots of the product $\Phi_{xy} \cdot \Phi_{yx}$. In these roots, z_0 will definitely be present as a single one. The next step is to:
- Assign z_0 to polynomial Φ_{xy} and the other roots, that are of no interest to us, to Φ_{yx}. Executing the algorithm for a second time will produce the characteristic polynomial matrix $\Phi(z)$, which means that we will actually know the coefficient matrices A_k , $k = 1,2,....., p$.

The above scheme was utilized four times, producing the examples that are presented in the next section.

6.2 *The results from the algorithm's utilization*

Let us consider the following examples representing some of the possible combinations that are scattered throughout the cases presented in section 5.4.

From the second category analyzed in 5.4, we consider the VAR(3) model:

$$x(t) = \mu + A_1 \cdot x(t-1) + A_2 \cdot x(t-2) + A_3 x(t-3) + \varepsilon(t), \qquad (6.2.1)$$

with $t = 0,\pm1,\pm2,\pm3,\ldots..$

$\det[(\bar{\Phi}^T \cdot \Phi)(z_0)] = 0$, $\Phi_{xx}(z_0) = 0$, $\Phi_{yy}(z_0) \neq 0$, $\Phi_{xy}(z_0) = 0$, $\Phi_{yx}(z_0) \neq 0$ and

$r_d = r_{xx} = r_{xy} = 1$, $r_{yy} = r_{yx} = 0$.

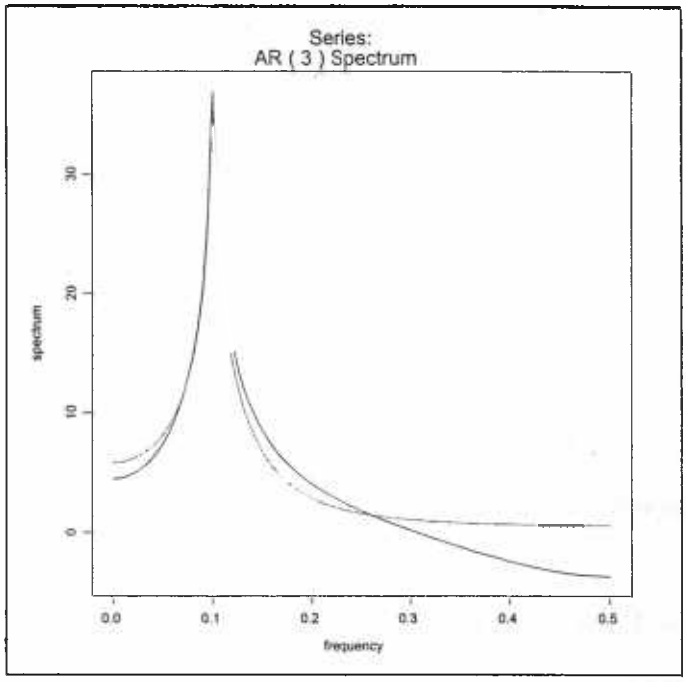
We can create a two dimensional VAR model with the above characteristics, utilizing the given algorithm. Let us produce two series such that the frequency component at $\pi/5$, will be responsible for a large percent of the models variance for both of them, (i.e. a peak will be present in these series at $\pi/5$). Giving the following roots:

Polynomial	***Root***
$\det(\bar{\Phi}^T \cdot \Phi)$	$z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}$
Φ_{xx}	$z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}$
Φ_{yy}	$z_0 = (0.5)^{-1} \cdot e^{i\frac{\pi}{10}}$
Φ_{xy}	$z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}$
Φ_{yx}	$z_0 = (0.363)^{-1} \cdot e^{i\cdot 0}$

Table 6.2.1: The polynomial's roots for the (6.2.1) VAR(3) model

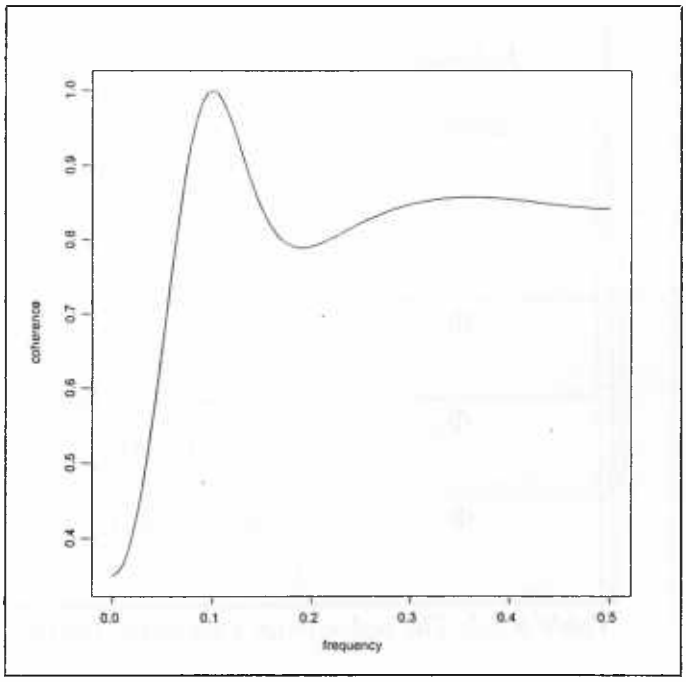
a very large peak will be present at the f_{11} and f_{22} (plot 6.2.1).

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ



Plot 6.2.1: The spectrum for the (6.2.1) VAR(3) model

The coherence for that particular model will also approach unity at $\pi/5$ as shown in the plot below:



Plot 6.2.2: The coherence for the (6.2.1) VAR(3) model

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

One aspect that should be investigated is how the size of the spectrum peaks, predetermine the value of the coherence. It should be clarified whether the coherence's approach of its maximum value, varies with the power of the spectrum at the specific frequency.

A simple way to investigate that, is by utilizing the last mentioned model (6.2.1) but with a small variation regarding the moduli of the root z_0. So let us replace $z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}$ with $z_0 = (0.85)^{-1} \cdot e^{i\frac{\pi}{5}}$. It is obvious that the size of the spectrum peaks at $\pi/5$ will be greatly reduced, but one should ask what will happen to the coherence. So once again we have the same model but with different root z_0:

$$x(t) = \mu + A_1 \cdot x(t-1) + A_2 \cdot x(t-2) + A_3 x(t-3) + \varepsilon(t), \qquad (6.2.2)$$

with $t = 0, \pm 1, \pm 2, \pm 3, \ldots$.

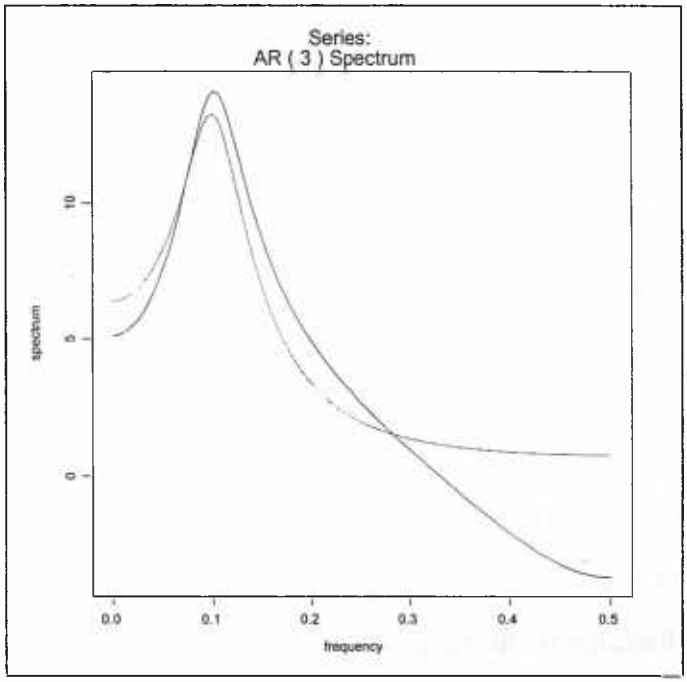
$\det[(\bar{\Phi}^T \cdot \Phi)(z_0)] = 0$, $\Phi_{xx}(z_0) = 0$, $\Phi_{yy}(z_0) \neq 0$, $\Phi_{xy}(z_0) = 0$, $\Phi_{yx}(z_0) \neq 0$ and $r_d = r_{xx} = r_{xy} = 1$, $r_{yy} = r_{yx} = 0$.

Polynomial	***Root***
$\det(\bar{\Phi}^T \cdot \Phi)$	$z_0 = (0.85)^{-1} \cdot e^{i\frac{\pi}{5}}$
Φ_{xx}	$z_0 = (0.85)^{-1} \cdot e^{i\frac{\pi}{5}}$
Φ_{yy}	$z_0 = (0.5)^{-1} \cdot e^{i\frac{\pi}{10}}$
Φ_{xy}	$z_0 = (0.85)^{-1} \cdot e^{i\frac{\pi}{5}}$
Φ_{yx}	$z_0 = (0.363)^{-1} \cdot e^{i\cdot 0}$

Table 6.2.2: The polynomial's roots for the (6.2.2) VAR(3) model

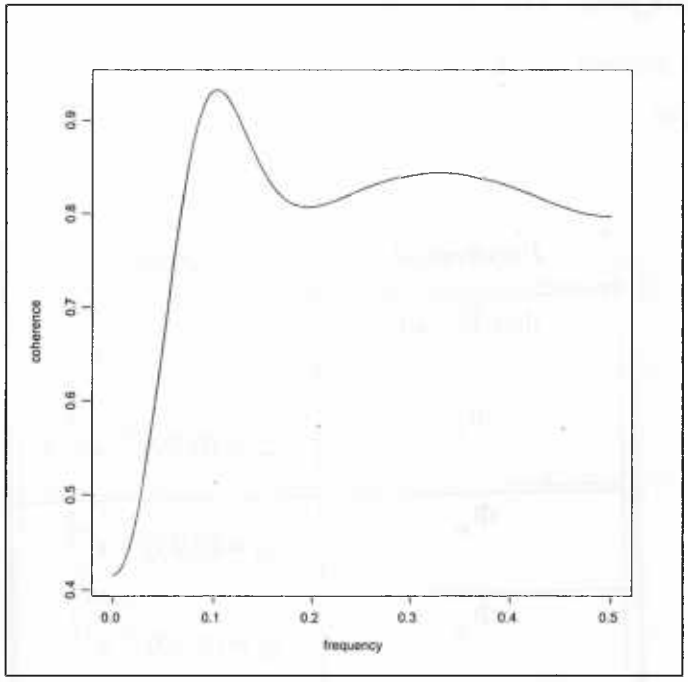
As it is show below the strength of the spectrum peaks at $\pi/5$ is greatly reduced:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ



Plot 6.2.3: The spectrum for the (6.2.2) VAR(3) model

Examining the coherence plot is obvious that its values are still very close to one with a small reduction (plot 6.2.4).



Plot 6.2.4: The coherence for the (6.2.2) VAR(3) model

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

This example forces us to conclude that the existence of even a small wave in a specific frequency in two series, will result in a high correlation between them, at that frequency. That high valued correlation (i.e. coherence) is not easily influenced by any change at the size of the spectrum peak.

Another characteristic case is where only one series has a peak at $\pi/5$ and the coherence does not reach one at $\pi/5$. For example let us consider the VAR(3) case where:

$$x(t) = \mu + A_1 \cdot x(t-1) + A_2 \cdot x(t-2) + A_3 x(t-3) + \varepsilon(t), \qquad (6.2.3)$$

with $t = 0, \pm 1, \pm 2, \pm 3, \ldots$.

$\det[(\bar{\Phi}^T \cdot \Phi)(z_0)] = 0$, $\Phi_{xx}(z_0) = 0$, $\Phi_{yy}(z_0) = 0$, $\Phi_{xy}(z_0) = 0$, $\Phi_{yx}(z_0) \neq 0$ and

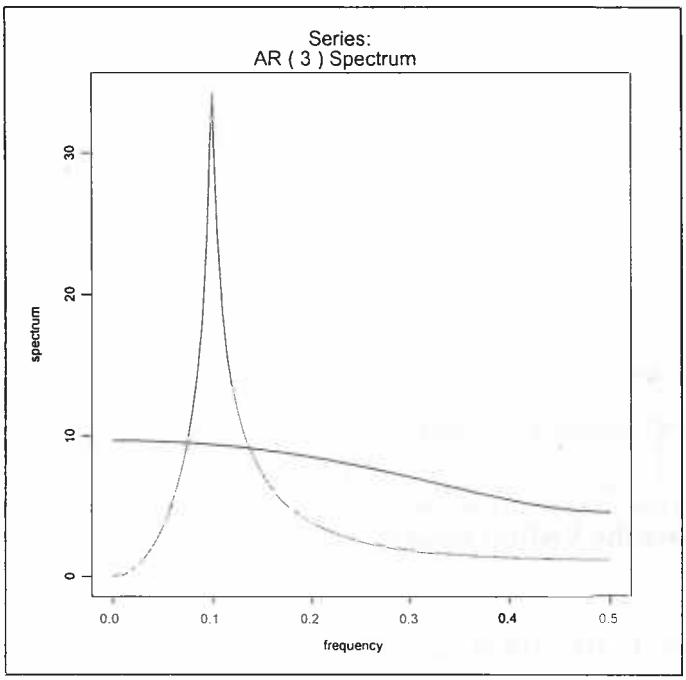
$r_d = r_{xx} = r_{yy} = r_{xy} = 1$, $r_{yx} = 0$.

We create a two dimensional VAR model with the above characteristics, that is the frequency component at $\pi/5$ will be responsible for a large percent of the model's variance for one series. Once again by giving the root $z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}$, a very large peak will be present at $\pi/5$ (plot 6.2.5). The roots as they were given to the polynomials are:

Polynomial	***Root***
$\det(\bar{\Phi}^T \cdot \Phi)$	$z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}$
Φ_{xx}	$z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}$
Φ_{yy}	$z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}$
Φ_{xy}	$z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}$
Φ_{yx}	$z_0 = (0.99)^{-1} \cdot e^{i\cdot 0}$

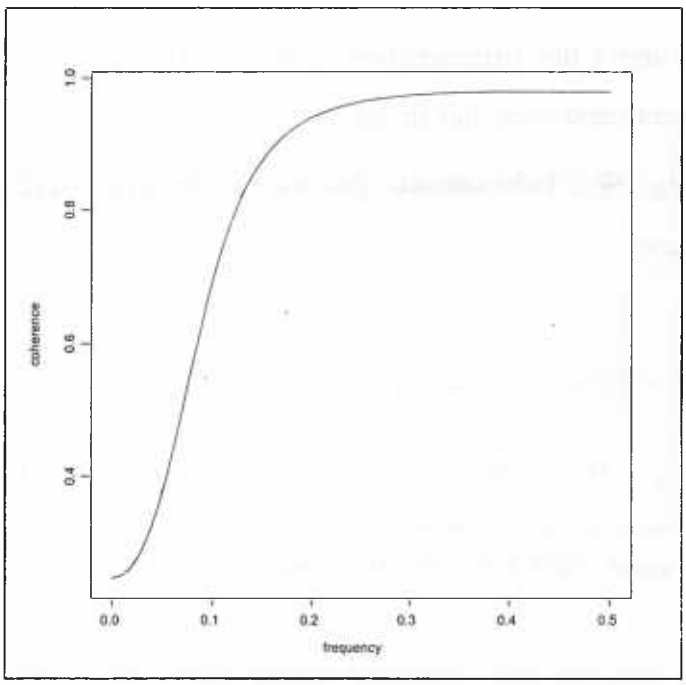
Table 6.2.3: The polynomial's roots for the (6.2.3) VAR(3) model

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ



Plot 6.2.5: The spectrum for the (6.2.3) VAR(3) model

The coherence for that model clearly does not reach its maximum value at $\pi/5$:



Plot 6.2.6: The coherence for the (6.2.3) VAR(3) model

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

The three situations presented above represent all the possible cases that are scattered throughout the exhaustive study where $r_d = 1$ and $r_{xx} = r_{yy} = r_{xy} = r_{yx} = 0$ or 1.

These cases verify the remark made in the same paragraph, stating that "a high correlation between two series in a specific frequency occurs if and only if, this frequency is strongly represented in these series".

But the question in hand is whether this is valid in general. So is this true for the cases where, for example, $r_d \geq 2$? As stated in 5.4 this is not the true and a case in which does not occur will be presented below. This example, exactly like the previously mentioned ones, simply verifies the theoretical conclusions drawn in 5.4.

So we have the VAR(4) model:

$$x(t) = \mu + A_1 \cdot x(t-1) + A_2 \cdot x(t-2) + A_3 x(t-3) + A_4 x(t-4) + \varepsilon(t), \qquad (6.2.4)$$

with $t = 0, \pm 1, \pm 2, \pm 3, \ldots$.

$\det[(\bar{\Phi}^T \cdot \Phi)(z_0)] = 0$, $\Phi_{xx}(z_0) = 0$, $\Phi_{yy}(z_0) = 0$, $\Phi_{xy}(z_0) = 0$, $\Phi_{yx}(z_0) = 0$ and $r_d = 2$, $r_{xx} = r_{yy} = r_{xy} = r_{yx} = 1$.

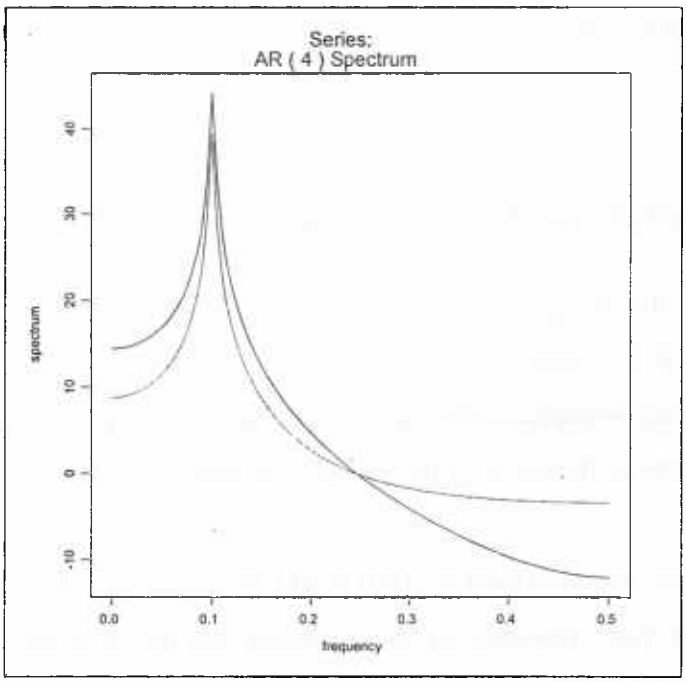
As noted above the differentiation of the (6.2.4) VAR(4) model in respect to the previously mentioned ones, lies in the order of z_0 as a root of the $\det(\bar{\Phi}^T \cdot \Phi)$ and the Φ_{xx}, Φ_{yy}, Φ_{xy}, Φ_{yx} polynomials. So the roots that were assigned at these polynomials are:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Polynomial	***Root***
$\det(\bar{\Phi}^T \cdot \Phi)$	$z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}, r_d = 2$
Φ_{xx}	$z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}$
Φ_{yy}	$z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}$
Φ_{xy}	$z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}$
Φ_{yx}	$z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}$

Table 6.2.4: The polynomial's roots for the ***(6.2.4) VAR(4) model***

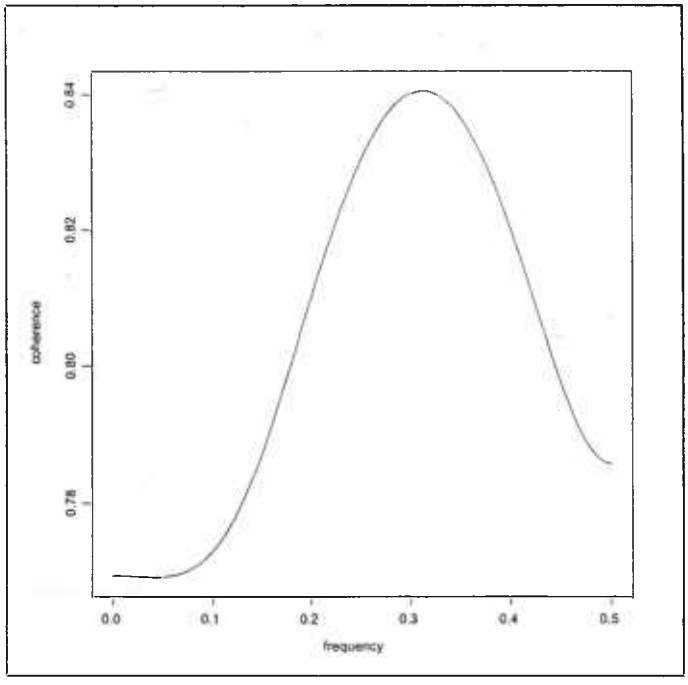
As expected the spectra for the model exhibit large peaks at $\pi/5$:



Plot 6.2.7: The spectrum for the (6.2.4) VAR(4) model

But the coherence's value at the specific frequency does not even come close to one.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ



Plot 6.2.8: The coherence for the (6.2.4) VAR(4) model

These four VAR model presented in this section are representatives of all the possible combinations that take place when we consider the cases where $r_d = 1$ and $r_{xx} = r_{yy} = r_{xy} = r_{yx} = 0$ or 1.

6.3 Topics for further investigation

All of the above, open up a wide range of issues that require further research. For example, the VAR model in question is a two dimensional one. The next step is to examine a more complex form of a VAR utilization and try to comprehend the interaction between the roots of the polynomial matrix (and its determinant), with the cross-spectra.

Furthermore, that interaction itself could be expanded. That is, to examine the impact of these roots not only on the coherence and the cross-spectra but also on the phase between the series.

One last issue, is the restriction imposed in (5.2.5) that the covariance matrix of the VAR(p) process is the unity matrix ($\Sigma = I$). As stated before, this makes our approach easier to comprehend without damaging the applicability of the results, but

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

one could focus on examining the impact of the existence of a different covariance matrix. Some special matrix structures, could create interesting variations on the conclusions presented in this thesis.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Appendix A

Presentation of the algorithm

The algorithm presented below was written in S-Plus. The first part is a function called "calculate.coefficients" that accepts as input the inverse roots of a polynomial and returns its coefficients.

```
#************************ SPECIFYING A TWO DIMENSIONAL VAR MODEL
#************************ VIA POLYNOMIAL ROOTS



#***************** roots conversion ******************

calculate.coefficients <- function(a.polynom)
{
a.polynom$order <- 0
a.polynom$coefs <- c(1, rep(0,2*a.polynom$inv.roots.number+1) )
k <- 0
while (k<a.polynom$inv.roots.number)
    {
    k <- k+1

    if ((a.polynom$inv.roots[k,2]==0) || (a.polynom$inv.roots[k,2]==pi))
        {
        a.polynom$order <- a.polynom$order + 1
        j <- a.polynom$order+1
        while (j>1)
        {
            a.polynom$coefs[j] <- a.polynom$coefs[j] -
                    a.polynom$coefs[j-1]*a.polynom$inv.roots[k,1]*
                        cos(a.polynom$inv.roots[k,2])
            j <- j-1
        }
        }
    else
        {
        a.polynom$order <- a.polynom$order + 2
        j <- a.polynom$order+1
        while (j>2)
            {
            print(k)
             print(j)
            a.polynom$coefs[j] <- a.polynom$coefs[j] -
                    2*a.polynom$coefs[j-1]*a.polynom$inv.roots[k,1]*
                                cos(a.polynom$inv.roots[k,2])+
                    a.polynom$coefs[j-2]*((a.polynom$inv.roots[k,1])^2)
            j <- j-1
            }
        a.polynom$coefs[2] <- a.polynom$coefs[2] -
                    2*a.polynom$inv.roots[k,1]*cos(a.polynom$inv.roots[k,2])
        }

    }
a.polynom
}
```


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

The above function accepts a list containing all the necessary information regarding the polynomial's roots. These are the number of (inverse) roots and the roots themselves, first their inverse moduli and secondly their argument. For example: $z_0 = (0.99)^{-1} \cdot e^{i\frac{\pi}{5}}$ is given below as `c(.99,pi/5)`. The returned coefficients are stored into a matrix called "CoefMatrix".

```
#************* INPUT ************************************************
N <-15                    #*********** MAXIMUM ORDER OF VAR***********
library(Matrix)
CoefMatrix <- Matrix(0,7,N)


Determinant.polynom<-list()
Determinant.polynom$inv.roots.number <- 2
Determinant.polynom$inv.roots <- matrix(c(.99,pi/5), ncol=2)
Determinant.polynom$inv.roots <-rbind(Determinant.polynom$inv.roots, c(0.99,pi/5))
Determinant.polynom


Fxx.polynom<-list()
Fxx.polynom$inv.roots.number <- 2
Fxx.polynom$inv.roots <- matrix(c(.99,pi/5), ncol=2)
Fxx.polynom$inv.roots<-rbind(Fxx.polynom$inv.roots, c(.5,pi/10))
Fxx.polynom

Fyy.polynom<-list()
Fyy.polynom$inv.roots.number <- 2
Fyy.polynom$inv.roots <- matrix(c(.99,pi/5), ncol=2)
Fyy.polynom$inv.roots<-rbind(Fyy.polynom$inv.roots, c(0.9,pi/5))
Fyy.polynom


Fxx.Fyy.polynom<-list()
Fxx.Fyy.polynom$inv.roots.number <-
    Fxx.polynom$inv.roots.number+Fyy.polynom$inv.roots.number
Fxx.Fyy.polynom$inv.roots<-rbind(Fxx.polynom$inv.roots,Fyy.polynom$inv.roots)
Fxx.Fyy.polynom


Fxy.polynom<-list()
Fxy.polynom$inv.roots.number <- 2
Fxy.polynom$inv.roots <- matrix(c(0.99,pi/5), ncol=2)
Fxy.polynom$inv.roots<-rbind(Fxy.polynom$inv.roots, c(0.3632713,0))
Fxy.polynom$inv.roots<-rbind(Fxy.polynom$inv.roots, c(0.9,pi/6))
Fxy.polynom


Fyx.polynom<-list()
Fyx.polynom$inv.roots.number <- 2
Fyx.polynom$inv.roots <- matrix(c(0.99,pi/5), ncol=2)
Fyx.polynom$inv.roots<-rbind(Fyx.polynom$inv.roots, c(0.9,pi/5))
Fyx.polynom
```

At first the roots of $\det(\bar{\Phi}^T \cdot \Phi)$, Φ_{xx} and the Φ_{yy} are given and the algorithm finds the roots of the product $\Phi_{xx} \cdot \Phi_{yy}$. Let us see how the function "calculate.coefficients" gives the coefficients which are stored into the "CoefMatrix".

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

But first a small checkpoint is introduced in order to ensure that the roots of the determinant are not within the unit circle.

```
#*************** CHECKPOINT No1. Checking the Determinant*******************

for (j in 1:(Determinant.polynom$inv.roots.number))
{
if  (Determinant.polynom$inv.roots[j,1]>1 )
{
    stop(message="The Determiant has roots in the unit circle!!")

}
}
#****************End of checkpoint No 1************************************


#************* Finding coefficients of Determinant ***********

Determinant.polynom<- calculate.coefficients(Determinant.polynom)
Determinant.polynom

CoefMatrix[1,]<-Determinant.polynom$coefs
for (j in (length(Determinant.polynom$coefs)+1):N)
{
    CoefMatrix[1,j]<-0
}


#************* Fxx
Fxx.polynom <- calculate.coefficients(Fxx.polynom)

CoefMatrix[3,]<-Fxx.polynom$coefs
for (j in (length(Fxx.polynom$coefs)+1):N)
{
    CoefMatrix[3,j]<-0
}


#************** Fyy


Fyy.polynom <- calculate.coefficients(Fyy.polynom)

CoefMatrix[4,]<-Fyy.polynom$coefs
for (j in (length(Fyy.polynom$coefs)+1):N)
{
    CoefMatrix[4,j]<-0
}


#************* Finding coefficients of Fxx*Fyy ***********
Fxx.Fyy.polynom <- calculate.coefficients(Fxx.Fyy.polynom)
CoefMatrix[2,]<-Fxx.Fyy.polynom$coefs
for (j in (length(Fxx.Fyy.polynom$coefs)+1):N)
{
    CoefMatrix[2,j]<-0
}


#************* Finding coefficients of  Fxy*Fyx ***********

CoefMatrix[5,]<-CoefMatrix[2,]-CoefMatrix[1,]
```


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

After the coefficients of the $\Phi_{xy} \cdot \Phi_{xy}$ are calculated, a second checkpoint is in order since it is imperative that Φ_{xy} and Φ_{yx} do not have a zero order term or equivalently their product should not have zero and first order terms. If this is the case the algorithm informs us for the situation and adds another root to the Φ_{yy} such that the problem disappears. Of course if this happens the coefficients of the $\Phi_{xx} \cdot \Phi_{yy}$ (and therefore of the $\Phi_{xy} \cdot \Phi_{yx}$ too) are re-calculated. (One should also note that in this stage the possibility $\Phi_{xy} \cdot \Phi_{yx} \equiv 0$ should be examined as mentioned in section 6.1. This checkpoint is omitted for simplicity reasons since it refers to a trivial case)

So:

```
#********************** Checkpoint No 2. Checking constant and linear term of Fxy.Fyx

if (CoefMatrix[5,1] !=0 && CoefMatrix[5,2]!=0) then
{

#stop(message="The roots for the Fyy are not acceptable!! The program will choose one
    for you!")


#********************** Defining Fyy ************************************

Subtracting.quantity<-0

for (j in (1:(Fyy.polynom$inv.roots.number)))
{
    if (Fyy.polynom$inv.roots[j,2]==0)
    {
        Subtracting.quantity<-Subtracting.quantity+Fyy.polynom$inv.roots[j,1]
    }

    if (Fyy.polynom$inv.roots[j,2] !=0)
    {

    Subtracting.quantity<-
    Subtracting.quantity+2*(cos(Fyy.polynom$inv.roots[j,2])*Fyy.polynom$inv.roots[j,1]
    )

    }
}


suggested.Fyy.inv.root<-CoefMatrix[3,2]-CoefMatrix[1,2]-Subtracting.quantity
Fyy.polynom$inv.roots <-rbind(Fyy.polynom$inv.roots, c(suggested.Fyy.inv.root,0))
Fyy.polynom$inv.roots.number <- (Fyy.polynom$inv.roots.number)+1
Fyy.polynom <- calculate.coefficients(Fyy.polynom)

CoefMatrix[4,]<-Fyy.polynom$coefs
for (j in (length(Fyy.polynom$coefs)+1):N)
{
    CoefMatrix[4,j]<-0

}
```


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

```
#********** SECOND definition of Fxx.Fyy **********

Fxx.Fyy.polynom<-list()
Fxx.Fyy.polynom$inv.roots.number <-
    Fxx.polynom$inv.roots.number+Fyy.polynom$inv.roots.number
Fxx.Fyy.polynom$inv.roots<-rbind(Fxx.polynom$inv.roots,Fyy.polynom$inv.roots)
Fxx.Fyy.polynom

#************* Finding coefficients of Fxx*Fyy for the SECOND time***********

Fxx.Fyy.polynom <- calculate.coefficients(Fxx.Fyy.polynom)
CoefMatrix[2,]<-Fxx.Fyy.polynom$coefs
for (j in (length(Fxx.Fyy.polynom$coefs)+1):N)
{   CoefMatrix[2,j]<-0
}



#************* Finding coefficients of  Fxy*Fyx for the SECOND time***********

CoefMatrix[5,]<-CoefMatrix[2,]-CoefMatrix[1,]



}

#************* END of Checkpoint No 2 **********************************
```

The next step is to find the roots of the $\Phi_{xy} \cdot \Phi_{yx}$:

```
#************* Finding roots of  Fxy*Fyx ***********
roots <-polyroot(CoefMatrix[5,])
NumberofRoots <-length(roots)


print('Your roots are.....')
for (j in 1:(NumberofRoots))

{
z<- roots[j]
#print[j]
print(1/Mod(z))
print(pi/Arg(z))
print('*****************')
}
```

That brings the first execution of the program to an end since the roots that were estimated above should be manually 'distributed' to Φ_{xy} and Φ_{yx} at the beginning of the program.

After the assignment of every root different that zero to each polynomial the algorithm is executed again where now it continues further on. The zero root is always present with order two, so the algorithm assigns one to each polynomial:

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

```
#************* Finding coefficients of Fxy ***********

Fxy.polynom <- calculate.coefficients(Fxy.polynom)

if (Fxy.polynom$inv.roots.number==0)
{
        Fxy.polynom$coefs[1]<-0
}

CoefMatrix[6,]<-Fxy.polynom$coefs
for (j in (length(Fxy.polynom$coefs)+1):N)
{
    CoefMatrix[6,j]<-0
}



#************* Finding coefficients of Fyx ***********

Fyx.polynom <- calculate.coefficients(Fyx.polynom)

if (Fyx.polynom$inv.roots.number==0)
{
        Fyx.polynom$coefs[1]<-0
}

CoefMatrix[7,]<-Fyx.polynom$coefs
for (j in (length(Fyx.polynom$coefs)+1):N)
{
    CoefMatrix[7,j]<-0
}


#**********************Adding the zero root to Fxy
for (i in 1:(N-1))
{
    CoefMatrix[6,N+1-i] <- CoefMatrix[6,N-i]
}
CoefMatrix[6,1] <-0
CoefMatrix[6,]<- CoefMatrix[6,]*CoefMatrix[5,3]


#**********************Adding the zero root to Fyx
for (i in 1:(N-1))
{

    CoefMatrix[7,N+1-i] <- CoefMatrix[7,N-i]

}
CoefMatrix[7,1] <-0
```

Our final checkpoint is introduced here. This is imperative since there is a chance that the roots of the $\det(\bar{\Phi}^T \cdot \Phi)$, Φ_{xx} and the Φ_{yy} are given in such way that the polynomial $\Phi_{xy} \cdot \Phi_{yx} = \Phi_{xx} \cdot \Phi_{yy} - \det(\Phi)$ is the zero one, while one of the Φ_{xy} or the Φ_{yx} is not identically zero. The checkpoint below makes that check and if indeed this is the case, it forces Φ_{yx}'s coefficients to be zero.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

```
#********************** Checkpoint No 3. Checking Fxy.Fyx *********************
counter.Fxy.Fyx<-0
counter.Fxy<-0
counter.Fyx<-0
for (j in 1:N)
{
counter.Fxy.Fyx<-counter.Fxy.Fyx+abs(CoefMatrix[5,j])
counter.Fxy<-counter.Fxy+abs(CoefMatrix[6,j])
counter.Fyx<-counter.Fyx+abs(CoefMatrix[7,j])
}


if (counter.Fxy.Fyx==0)
{
if (counter.Fxy!=0 & counter.Fyx!=0)
{
stop(message="There is a problem with the FxyFyx polynomial!")

#******* Making Fyx zero polynomial **************

for (j in 1:N)
{
    CoefMatrix[7,j]<-0
}


}
}
#************** END of Checkpoint No 3 ************************************
```

This is the end of the polynomial's manipulation. Now it is time to estimate the order p of the VAR(p) model.

```
#******************************** Estimating the ORDER ***************************


i<-N

while (CoefMatrix[3,i]==0 & CoefMatrix[4,i]==0 & CoefMatrix[6,i]==0 &
    CoefMatrix[7,i]==0 )

{
    order.of.Var<-(i-2)
    i<-(i-1)
}

order.of.Var
```

Having the order of the VAR(p) model, it is time to proceed to the estimation of the p coefficient matrices. These matrices together with the covariance matrix, are all that we need in order to estimate the theoretical spectrum of the model.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

```
#**************** Estimation of the coefficient matrices ****************

CovarianceMatrix <- matrix(c(1,0,0,1),k)

Var.Matrices <-array(0,dim=c(order.of.Var,2,2))

for (i in 1:order.of.Var)
{

j<-i+1
Var.Matrices[i,,] <- array(c(-CoefMatrix[3,j],-CoefMatrix[7,j],-CoefMatrix[6,j],-
    CoefMatrix[4,j]),dim=c(k,k))
print(Var.Matrices[i,,])
}
Var.Matrices[1,,]
```

It is now easy to estimate the true spectrum of the Var(p):

```
#****************** REAL VAR SPECTRUM ********************

timeseries.coeff <-list()
timeseries.coeff$ar <- Var.Matrices
timeseries.coeff$var.pred <-CovarianceMatrix
timeseries.coeff$order <- order.of.Var


real.var.spectrum<-spec.ar(timeseries.coeff,plot=T)
reducedphase1 <- real.var.spectrum$phase%%(2*pi)
plot(real.var.spectrum$freq,real.var.spectrum$coh,type="l",xlab="frequency",ylab="cohe
    rence")
plot(real.var.spectrum$freq,reducedphase1, type="l",xlab="frequency",ylab="phase")
```

With the algorithm presented above all of the examples presented in 6.2 where easily created. It is obvious that a two dimensioned Var(p) model with a specific structure can now be created at will.

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

References

Akdi, Y. and Dickey, D. A. (1998). Periodograms of unit root time series: Distributions and tests, *Communications in Statistics*, 27, 69-87

Brockwell, P. J. and Davis, R. A. (1987). Time series: Theory and Methods. *Springer-Verlang, New York*

Brockwell, P. J. and Davis, R. A. (1996). Introduction to time series and forecasting. *Springer-Verlang, New York*

Chan, N. H. and Wei, C. Z. (1987). Asymptotic inference for nearly nonstationary AR(1) processes, *Annals of Statistics*, 15, 1050-1063

Choi, I. and Phillips, P. C. B. (1993). Testing for a unit root by frequency domain regression, *Journal of Econometrics*, 59, 263-286

Dahlhaus, R. (1988). Small sample effects in time series analysis: A new asymptotic theory and a new estimate, *Annals of Statistics*, 16, 808-841

Dickey, D. A. and Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root, *Journal of the American Statistical Association*, 74, 427-431

Engle, R. F. and Granger, C. W. J. (1987). Co-integration end error correction: representation, estimation and testing, *Econometrica*, 55, 251-276

Granger, C. W. J. and Weiss, A. A. (1983). Time Series Analysis of Error-Correcting Models. Studies in Econometrics, Time Series and Multivariate Statistics. *New York: Academic Press*

Hassler, U. (1993). Unit roots tests: The autoregressive approach in comparison with the periodogram regression, *Statistical Papers*, 34, 67-82

Jenkins, G. and Watts, D. (1968). Spectral analysis and its applications. *Holden-Day, San Francisco*

Levy, D. (2002). Cointegration in frequency domain, *Journal of Time Series Analysis*, 23, 333-339

Lutkepohl, H. (1993). Introduction to multiple time series analysis. *Springer-Verlang, New York*

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Phillips, P. C. B. and Perron, P. (1988). Testing for a unit root in time series regression, *Biometrika,* 75, 335-346

Phillips, P. C. B. and Moon, H. R. and Xiao, Z. (2001). How to estimate autoregressive roots near unity, *Econometric Theory*, 17, 29-69

Priestley, M. (1981). Spectral analysis and time series. *Academic Press, London*

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ

Δωρεά


80025 75540


ATHENS UNIVERSITY OF ECONOMICS & BUSINESS
ΟΙΚΟΝΟΜΙΚΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΑΘΗΝΩΝ