شبكة المعلومات الجامعية

بسم اللّه الرحمن الرحيم

SALWA AKL

شبكة المعلومات الجامعية

شبكة المعلومات الجامعية

التوثيق الالكتروني والميكروفيلم

SALWA AKL

شبكة المعلومات الجامعية

جامعة عين شمس

التوثيق الإلكتروني والميكروفيلم

قسم

نقسم بالله العظيم أن المادة التي تم توثيقها وتسجيلها

علي هذه الأقراص المدمجة قد أعدت دون أية تغيرات

يجب أن

تحفظ هذه الأقراص المدمجة بعيدا عن الغبار

SALWA AKL

بعض الوثائق

الأصلية تالفة

SALWA AKL

بالرسالة صفحات

لم ترد بالأصل

SALWA AKL

Cairo University
Faculty of Economics and Political Science
Department of Statistics

BAYESIAN CLASSIFICATION WITH SEASONAL AUTOREGRESSIVE SOURCES.

By

Mohamed Youssef Sabry

Supervised by

Dr. Ahmed L. Daif
Associate Professor of Statistics
Department of Statistics

Dr. Azza A. El-Naggar
Assistant Professor of Statistics
Department of Statistics

Faculty of Economics and Political Science
Cairo University

A Thesis Submitted to The Faculty of Economics and
Political Science (Department of Statistics)
in Partial Fulfillment of The Requirement
for The M.Sc. Degree in Statistics
2002

Name: Mohamed Youssef Ezz Al Dien Zaky Sabry **Nationality:** Egyptian

Date and place of birth: 18/7/1970- Giza.

Specialization: Statistics.

Supervisors: Prof. Dr Ahmed Loutfy Daif Associate Professor

Dr. Azza Abdal Halim El-Naggar Assistant Professor

Department of Statistics
Faculty of Economics and Political Science
Cairo University

BAYESIAN CLASSIFICATION WITH SEASONAL AUTOREGRESSIVE SOURCES

Summary: In this research, four Bayesian Classification techniques are developed, using normal gamma and Jeffreys' rule, to classify a univariate time series into one of several seasonal autoregressive sources. The four techniques deal with different four cases of the sources' orders. The proposed techniques assume that the coefficients and precision are unknown. The idea of the four techniques is to classify a time series realization to a source r, when the marginal posterior mass function of a classification vector λ has its largest value at the r-th mass point. Numerical studies are carried out to check the performance of the proposed technique with respect to time series length, closeness between parameters, the choice effect of the false source orders and the choice effect of precision. Also, the proposed technique is illustrated with a case study on tourism.

Supervisors

Prof. Dr Ahmed Loutfy Daif

Dr Azza El-Naggar

عزة النجار

BAYESIAN CLASSIFICATION WITH SEASONAL AUTOREGRESSIVE SOURCES

ABSTRACT

In this research, four Bayesian Classification techniques are developed, using normal gamma and Jeffreys' rule, to classify a univariate time series into one of several seasonal autoregressive sources. The four techniques deal with different four cases of the sources' orders. The proposed techniques assume that the coefficients and precision are unknown. The idea of the four techniques is to classify a time series realization to a source r, when the marginal posterior mass function of a classification vector λ has its largest value at the r-th mass point. Numerical studies are carried out to check the performance of the proposed technique with respect to time series length, closeness between parameters, the choice effect of the false source orders and the choice effect of precision. Also, the proposed technique is illustrated with a case study on tourism.

Key Words

- Bayesian Classification
- Conditional Likelihood Function
- Marginal Posterior Function
- Jeffreys' Rule
- Unit roots
- Seasonal Autoregressive Sources
- Classification Vector
- Prior Distributions
- Univariate Time Series
- ARMA Models

Acknowledgment

During preparing this study, I have received valuable advice and helpful suggestions from several persons.

In particular, I wish to express my depth gratitude to Associate Professor Dr. Ahmed Loatfy who supervised this thesis and gave me encouragement, support and valuable comments.

It is a pleasure to acknowledge the effort of my supervisor Assistant Professor Dr. Azza El-Naggar for her effort, guidance and sincere advice.

I would like to take the opportunity to thank Professor Dr. Tarek Amira the head of statistics department and all staff of statistics department for their useful discussions and comments.

Also, depth gratitude is extended to my parents and my brother for their prayers and encouragement.

Abstract

Classification analysis is concerned with the problem of classifying objects or observations into one of several populations, which they must closely resemble on the basis of a number of measurements made on these objects or observations. In the classification problem, the measurements are presented via variables describing the objects. If the observed variables present time series, we will have a time series classification problem. Furthermore, if we have only one variable, it will be a univariate time series classification problem, otherwise we will call it a multiple time series classification problem. In the thesis, we propose a technique to solve one of univariate time series classification problem.

In this thesis, four Bayesian classification techniques are developed to classify a univariate time series into one of several seasonal autoregressive sources. The main difference between the four techniques is the orders of the sources. The *first* technique deals with sources that share a common known order (p,P).The *second* technique assumes that sources might have a different known order (p_r,P_r). The sources in the *third* technique are assumed to share a common unknown order (p,P). The *fourth* technique assumes that sources might have a different unknown order (p_r,P_r). In these four techniques, it is assumed that the coefficients and precision are not known. The concept of four techniques is based on classifying a time series realization to a source r, when the marginal posterior mass function of a classification vector λ has its largest value at the r-th mass point. A Bayes rule is used to derive the marginal posterior function via combining an appropriate conditional likelihood with a normal gamma prior distribution.

Numerical studies are carried out to test the adequacy of each proposed technique. Numerical studies demonstrate that the proposed techniques give good results in the majority of cases. Therefore, they are adequate for classification with seasonal autoregressive sources. The proposed techniques are illustrated by a case study on tourism using a real life data. Finally, some ideas are presented for the future work.

Table of Contents

Chapter I

Introduction

- **Literature Review**
- **Thesis Contributions**
- **Thesis Contents**

Chapter 1

Introduction

Observations on a variable that are made over time are known as time series data. These observations are dependent on each other and this dependence is important in itself in identifying time series models, Gaynor and Kirkapatric, (1994). The body of techniques available for the analysis of such series is referred to time series analysis. Time series data are commonplace in today's society. There are many examples for time series in life such as published data of gross net product, Dow Jones industrial average and consumer price index.

There is a common problem in real life applications that is how to assign, using an efficient rule, an object whose origin is unknown to one of several populations. This is called the discrimination and classification problem. Discrimination and classification are multivariate techniques concerned with separating distinct sets of objects (or observations) and with allocating new objects (observations) to previously defined groups. This problem is studied through discrimination and classification techniques. Describing, either graphically or algebraically, the differential features of objects and finding discriminants whose numerical values are such that the collections are separated as much as possible is known as discrimination, Johnson and Wichern, (1998). The classification is concerned with applying the allocation rule derived in discrimination to a new object of unknown population to assign it into one of available several sources.

1.1 Literature Review

Box and Jenkins have important contributions in time series analysis. In 1970, they introduced in their famous book the methodology and the theories that enabled statisticians later, to study time series in a highly systematic way. There are a lot of studies in non-Bayesian work, see Box and Jenkins (1970), Priestley (1981), and Tong (1990). Bayesian contributions started in 1970. Box and Jenkins (1970) introduced a Bayesian analysis of autoregressive process. Zellner (1971) began the systematic Bayesian analysis by deriving the posterior and predictive distributions of the first and the second order autoregressive process. In 1973, Newbold made an important step in Bayesian analysis by his study to ARMA-type transfer function models. Litterman (1980), introduced a Bayesian forecasting technique based on using a prior distribution. Many studies have been on Bayesian analysis in the eighties, for more details the reader can refer to Monahan (1983), Broemeling and Shaarawy (1986), and Cook (1985).

With regard to classification problem. Mahalanobis (1936) suggested a measure of "distance" between the populations, Morrison, (1976). Fisher (1936) introduced the linear discriminant function, Hand, (1981). Welch (1939) introduced the likelihood ratio technique, Morrison, (1976). Walker (1967) began the work of non-Bayesian classification of time series. He used the concept of Cox "tests of separate families" to establish an approximate discrimination method between ARMA sources. There are some important contributions, in applying non-Bayesian techniques on time series classification problem introduced by Gersch and Brotherton (1979), Gersch (1981) and Kedem and Slud (1982). Broemeling and Son (1987) initiated the work of Bayesian classification of time series. They introduced a Bayesian technique to assign a time series realization into one of several autoregressive sources with common known order. Shaarawy and Haroun (1992) developed a Bayesian technique to classify a time series realization into moving average sources with common known order. Daif and Zahran (1996) developed a Bayesian classification of ARMA sources with different orders. Abd-Elshafy (2000) introduced a Bayesian classification technique of multivariate ARMA models. All previous contributions have dealt with nonseasonal time series models, while contributions to classification of seasonal time series models are not considered.

1.2 Thesis Contributions

The Thesis has four contributions in Bayesian classification of seasonal univariate time series. It develops four Bayesian classification techniques based on using a classification vector to assign a realization into one of available several seasonal autoregressive sources, with unknown coefficients and precision. The *first* technique assumes that sources have common known orders. The *second* technique assumes that sources have different known orders. The *third* technique assumes that, sources have common unknown orders. The *fourth* technique assumes that sources have different unknown orders. The four techniques are developed using an appropriate conditional likelihood and an appropriate prior distribution. Each technique is derived using both of normal gamma and Jeffreys' prior.