TIME SERIES ANALYSIS WITH INTERVENTIONS:

A BAYESIAN APPROACH

By

JOAQUIN DIAZ-SAIZ

Bachelor of Science
Instituto Tecnologico y de Estudios
Superiores de Monterrey
Monterrey, N. L. Mexico
1966

Master of Science
Centro Interamericano de
Ensenanza de Estadistica
Santiago, Chile
1968

Submitted to the Faculty of the Graduate College
of the Oklahoma State University
in partial fulfillment of the requirements
for the Degree of
DOCTOR OF PHILOSOPHY
July, 1985

TIME SERIES ANALYSIS WITH INTERVENTIONS:

A BAYESIAN APPROACH

Thesis Approved:

Thesis Adviser

Dean of The Graduate College

ACKNOWLEDGEMENTS

I would like to express my sincere appreciation to Dr. Lyle D. Broemeling for serving as my major adviser and wish to thank him for his valuable guidance during my graduate studies.

I would also express my gratitude to the other committee members, Dr. J. Leroy Folks, Dr. Nitis Mukhopadhyay, and Dr. John D. Rea for their time and assistance, and to the other faculty members of the Oklahoma State University that contributed to make my stay a very valuable experience.

I would like to gratefully acknowledge the support provided by the Office of Naval Research through contract number N00014-82-K-0292, and the Universidad Nacional Autonoma de Mexico.

Last, but certainly not least, I would like to express my most sincere gratitude to my wife, Aida, and to my parents, Francisco Diaz Ponte and Alicia Saiz de Diaz for their constant encouragement and support.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Data arising from research in a wide variety of fields occur in the form of time series. Frequently, the occurrence of one or more events produces changes in the model that generated the series in various ways.

The analysis of changing time series has been the subject of many recent publications. In this paper, a Bayesian procedure will be proposed to analyze the case in which a change of the level of the series has been produced at a known point in time. This problem is known in the literature as Intervention Analysis.

Most of the work done in the area relies on the use of difference equation models and maximum likelihood procedures. Due to the nonlinearities involved, the implementation of those procedures requires the use of approximations which are based on the large sample properties of the estimators obtained, or on the behavior of the likelihood function when the number of observations is large.

The method proposed in this paper relies on the use of a large sample approximation of the likelihood function. This approximation eliminates some of the nonlinearities and produces posterior distributions which, at least conditionally correspond to multivariate t-distributions, if a suitable family of priors is used. These conditional posterior distributions will be used to model the effect of an intervention, following the stages used in the classical approach of time series analysis.

A review of the literature in the general topic of the analysis of changing time series, and in particular in the area of intervention analysis is presented in Chapter II.

The description of the problem and the models to be used, together with the derivation of the approximate posterior and predictive distributions is the subject of Chapter III, while Chapter IV contains the methodology proposed to model the intervention effect by means of conditional posterior distributions, assuming that the noise model has previously identified and its parameters estimated through the observations prior to the time in which the intervention takes place.

In Chapter V, the modeling procedures will be applied to two examples and the results obtained will be analyzed. A summary of the research done will be given in Chapter VI.

CHAPTER II

REVIEW OF THE LITERATURE

Extensive work has been done to analyze and model time series. Perhaps the most cited reference in the area is the book by Box and Jenkins (1970) in which a methodology was proposed to model the time series through the so called autoregressive-moving average (ARMA) models. This approach has been, so far, the most widely used by scientists in many different fields.

From the Bayesian point of view, an approach that constitutes a very natural one to use in this problem, we can mention the book by Zellner (1971) in which a chapter is devoted to present the analysis of the first and second order autoregressive processes. Further work was presented by Newbold (1973), Smith (1979), Broemeling and Land (1984), and Cook (1983). An extensive work to develop a complete Bayesian time series analysis procedure has been done by Broemeling and Shaarawy (1984a, 1984b), Shaarawy (1984), and Shaarawy and Broemeling (1984a, 1984b, 1985). Their work includes the analysis of multiple time series and is based on an approximation to the likelihood function which eliminates the nonlinearities through the maximum likelihood estimation of the residuals. This procedure has advantage of eliminating the need of the second order derivatives to approximate the likelihood function. Also, conjugate priors can easily be used to incorporate the prior information available and produce posterior distributions which correspond to multivariate t-distributions and are easy to work with.

A problem that has interested several people recently, is related to the analysis of the case in which the model that generated a time series could have changed during the recording of the data. In this regard, the papers by Salazar (1982) and Cook (1983) study the case in which the point in time in which the change occurred is unknown.

In the area of intervention analysis, the point in time at which the change could have taken place is known and, if in fact there was a change it affected the level of the series but not the noise structure. Probably the first paper related to this topic is the one by Box and Tiao (1965) in which they study the problem of analyzing the change in level of an integrated moving average process. A Bayesian analysis of the problem was included there.

An extensive coverage of the topic appears in the book by Glass, Willson and Gottman (1975) as part of the area of analysis of time series experiments. Box and Tiao (1975) proposed the use of difference equation models to represent the dynamic intervention effect and the noise. The proposed models are applied there to a series of monthly averages of the level of 0_3 in the atmosphere of Downtown Los Angeles to determine the effects of three different interventions, and also to a series of the monthly rate of change of the consumer price index where two interventions are considered. The analysis is done in this paper using maximum likelihood procedures and, although the authors mention that the Bayesian analysis can be done after approximating the posterior density, they do not report any detailed results. The publications mentioned before deal with univariate series.

In the multivariate case, Abraham (1980) obtained approximate maximum likelihood estimators for the parameters of the model. The approximation is needed due to the difficulties that arise from the nonlinearities.

The intervention analysis approach has been widely used in different fields, and as examples we can mention the papers by Stoline and Huitema (1978) in the social sciences, Shahabudin (1980) in the area of finance, and Lasarre and Tau (1982) in the area of public safety, among others.

In this report, the model suggested by Box and Tiao will be analyzed from the Bayesian point of view, using an approximation to the likelihood function which is essentially an extension to the one proposed first by Shaarawy (1984) for the case of the moving average models. A complete conditional modeling procedure for the intervention effect will be presented.

CHAPTER III

POSTERIOR AND PREDICTIVE ANALYSIS

The model proposed by Box and Tiao (1975) to analyze time series with interventions corresponds to

$$Y_t = \sum_{j=1}^{K} \{\omega_j(B)/\delta_j(B)\}\xi_{tj} + N_t \quad t = 1, 2, \ldots, N \tag{3.1}$$

where $\omega_j(B)$ and $\delta_j(B)$ are polynomials in the backshift operator B, of degrees r_j and s_j respectively, for $j = 1, 2, \ldots, K$. The roots of $\omega_j(B)$, $j = 1, 2, \ldots, K$, are normally required to be outside the unit circle and, the roots of $\delta_j(B)$, $j = 1, 2, \ldots, K$, are required to be, outside or on the unit circle.

The noise N_t, $t = 1, 2, \ldots, N$ is assumed to be generated by the autoregressive moving average (ARMA) process

$$N_t = \frac{\theta(B)}{\phi(B)} a_t \qquad t = 1, 2, \ldots, N \tag{3.2}$$

where $\theta(B)$ and $\phi(B)$ are polynomials in B of degrees q and p respectively, with roots outside the unit circle.

The posterior and predictive analysis of this model can be very complicated mainly because the form of the posterior distributions does not correspond to any standard one, and the number of parameters makes the numerical analysis very difficult to handle, unless it is restricted to some particular cases. This chapter presents the analysis of two models that are known to be useful in practice. The first corresponds to the case of a single intervention with no restrictions on the polynomials in the transfer function, and will be the topic of section 3.1. Section 3.2 will be devoted to the analysis of what we call the linear intervention model, which includes several interventions but the polynomials in the denominators of the corresponding transfer functions are all equal to one.

For the analysis, an approximation to the likelihood function will be used, where estimators of the residuals are calculated recursively using the maximum likelihood estimators of the parameters. This procedure was proposed by Shaarawy (1984) for the analysis of moving average models. Some comments on the validity of this approach will be given in the appendix of this paper. To model the prior information, normal-inverted gamma and "noninformative" priors will be used.

3.1 Single Intervention Model

The Approximate Likelihood Function

The model described by (3.1), for the case of a single intervention corresponds to

$$Y_t = (\omega(B)/\delta(B))\xi_t + N_t \quad (t = 1, 2, \ldots, N) \tag{3.3}$$

where $\omega(B)$ and $\delta(B)$ are polynomials in the backshift operator B of degrees r and s respectively, ξ_t is the intervention variable that corresponds to an indicator function of certain subset of the time domain where the intervention is taking place, and N_t is the random noise assumed to be generated by the ARMA(p,q) model (3.2).

If the polynomials in (3.2) and (3.3) are given by

$$\omega(B) = \omega_0 + \omega_1 B + \ldots + \omega_r B^r \tag{3.4}$$

$$\delta(B) = 1 + \delta_1 B + \ldots + \delta_s B^s \tag{3.5}$$

$$\phi(B) = 1 + \phi_1 B + \ldots + \phi_p B^p \tag{3.6}$$

$$\theta(B) = 1 + \theta_1 B + \ldots + \theta_q B^q \tag{3.7}$$

the residuals at time t can be given in terms of the observations $Y_1, Y_2, \ldots, Y_t$ and the residuals $a_1, a_2, \ldots, a_{t-1}$ by the equation

$$\begin{aligned} a_t &= \phi(B)Y_t - (\theta(B)-1)a_t + \delta_1(\phi(B)Y_{t-1} - \theta(B)a_{t-1}) \\ &+ \cdots + \delta_s(\phi(B)Y_{t-s} - \theta(B)a_{t-s}) - \omega_0\phi(B)\xi_t - \omega_1 \phi(B)\xi_{t-1} - \\ &\ldots - \omega_r \phi(B)\xi_{t-r} \end{aligned} \tag{3.8}$$

or

$$\begin{aligned} a_t &= (\delta(B)Y_t - \omega(B)\xi_t - (\delta(B)-1)a_t) - (\phi_1(\omega(B)\xi_{t-1}-\delta(B)Y_{t-1}) \\ &+ \cdots + \phi_p(\omega(B)\xi_{t-p}-\delta(B)Y_{t-p})) + \theta_1\delta(B)a_{t-1} + \cdots + \theta_q\delta(B)a_{t-q} \end{aligned} \tag{3.9}$$

Since the Jacobian of the transformation is equal to one, the likelihood function of the parameters given $Y_1, Y_2, \ldots, Y_{p+s}$ and assuming that $a_1, a_2, \ldots, a_{p+s}$ are all equal to zero, corresponds to

$$\ell(\underline{\omega},\underline{\delta},\underline{\phi},\underline{\theta},\sigma^2) \propto \frac{1}{(\sigma^2)^{\frac{N-(p+s)}{2}}} \exp\{-\frac{1}{2\sigma^2}(\sum_{t=p+s+1}^{N} a_t^2)\} \tag{3.10}$$

where $\underline{\omega}' = (\omega_o,\omega_1, \ldots, \omega_2)$, $\underline{\delta}' = (\delta_1, \ldots, \delta_s)$, $\underline{\phi}' = (\phi_1, \ldots, \phi_p)$, $\underline{\theta}' = (\theta_1, \ldots,\theta_q)$ and, a_t, t=p+s+1, ..., N are given by (3.8) or (3.9).

Although the exponent in (3.10) can be written in terms of the observations, the nonlinearities involved produce a likelihood function which is very difficult to work with. We propose here the approximation of the likelihood by using the observable variables $\hat{a}_t$, t=p+s, ..., N which correspond to the values of the variables a_t, t=p+s, ..., N obtained recursively from equations (3.8) or (3.9), using the maximum likelihood estimators of the parameters as the parameter values. The approximate likelihood corresponds to

$$\ell^*(\underline{\omega},\underline{\delta},\underline{\phi},\underline{\theta},\sigma^2|Y)$$

$$\alpha \frac{1}{(\sigma^2)^{\frac{N-(s+p)}{2}}} \quad \exp\{-\frac{1}{2\sigma^2}(\underline{z} - X\underline{\gamma})' \ (\underline{z} - X\underline{\gamma})\} \tag{3.11}$$

where the components of $\underline{z}$ are

$$z_t = \phi(B)Y_t - (\theta(B) -1)a_t \tag{3.12}$$

the elements of X are

$$X_{t,j} = \phi(B)\xi_{t-j+1} \quad j = 1, 2, \ldots, r + 1 \tag{3.13}$$

$$X_{t,j} = (\theta(B)\hat{a}_{t-j+r+1} - \phi(B)Y_{t-j+r+1}) \quad j = r+2, \ldots, r+s+1 \tag{3.14}$$

for t = p+s+1, ..., N and $\underline{\gamma}' = (\omega_o, \omega_1, \ldots, \omega_r, \delta_1, \delta_2, \ldots, \delta_s)$ and $\hat{a}_t = 0$, t = 1, 2, ..., p+s. Equation (3.11) can be obtained from (3.8) and (3.10) and will be used to obtain the conditional posterior distribution of $\underline{\gamma}$ given $\underline{\eta}' = (\underline{\phi}', \underline{\theta}')$. Notice that $\underline{z}$ and X are functions of $\underline{\eta}$.

Using (3.9) and (3.10), it can be shown that (3.11) is equivalent to

$$\ell^*(\underline{\omega}, \underline{\delta}, \underline{\phi}, \underline{\theta},\sigma^2|\underline{Y})$$

$$\alpha \frac{1}{(\sigma^2)^{\frac{N-(s+p)}{2}}} \exp \{- \frac{1}{2\sigma^2} (\underline{U}-Q\underline{\eta})'(\underline{U}-Q\underline{\eta})\} \tag{3.15}$$

where the components of $\underline{U}$ are given by

$$u_t = \delta(B)Y_t - \omega(B)\xi_t - (\delta(B) -1)\hat{a}_t \tag{3.16}$$

and, the elements of the matrix Q are

$$q_{t,j} = \omega(B)\xi_{t-j} - \delta(B)Y_{t-j} \quad j = 1, 2, \ldots, p$$

$$q_{t,j} = \delta(B)\hat{a}_{t-j+p} \quad j = p + 1, \ldots, p + q \tag{3.17}$$

for $t = p + s + 1, \ldots, N$. Equation (3.15) will be useful in deriving the conditional posterior distribution of $\underline{\eta}$ given $\underline{\gamma}$. Notice that, in this case, $\underline{U}$ and Q are functions of the parameter vector $\underline{\gamma}$.

From now on, by likelihood function we mean the approximate likelihood given by (3.11) or (3.15). It is important to mention here that, even though some of the nonlinearities have been removed through the approximation procedure, some of them remain because of the cross products between some of the intervention and noise parameters. This problem cannot be eliminated but, the conditional posterior distributions, which correspond to multivariate t-distributions, will enable us to develop some inferential procedures for estimation and testing.

The Prior Distribution

It will be assumed in this chapter that the prior distribution of the vector of parameters $\underline{\mu}' = (\underline{\gamma}', \underline{\eta}')$ given σ^2 is normal with mean $\underline{\mu}'_o = (\underline{\gamma}'_o, \underline{\eta}'_o)$ and variance-covariance matrix $\sigma^2\Sigma$, and that the prior distribution of σ^2 is

$$\Pi(\sigma^2) \propto \frac{1}{(\sigma^2)^{k+1}} \exp \{ \frac{-\alpha}{2\sigma^2} \} . \tag{3.18}$$

Under these assumptions, the joint prior distribution can be written as

$$\Pi(\underline{\gamma},\underline{\eta},\sigma^2) \; \alpha \; \frac{1}{(\sigma^2)^{\frac{p+q+r+s+2k+1}{2}+1}}$$

$$\exp\left\{-\frac{1}{2\sigma^2}\left[(\underline{\gamma}-\underline{\gamma}_o(\underline{\eta}))'\,\Sigma_{11.2}^{-1}\,(\underline{\gamma}-\underline{\gamma}_o(\underline{\eta})) + (\underline{\eta}-\underline{\eta}_o)'\,\Sigma_{22}^{-1}\,(\underline{\eta}-\underline{\eta}_o) + \alpha\right]\right\} \tag{3.19}$$

where $\underline{\gamma}_o(\underline{\eta})$ and $\sigma^2\,\Sigma_{11.2}$ are the parameters of the conditional distribution of $\underline{\gamma}$ given $\underline{\eta}$ and σ^2, and $\sigma^2\Sigma_{22}$ is the variance-covariance matrix of the marginal distribution of $\underline{\eta}$ given σ^2.

Formula (3.19) is equivalent to

$$\Pi(\underline{\gamma},\underline{\eta},\sigma^2) \; \alpha \; \frac{1}{(\sigma^2)^{\frac{p+q+r+s+2k+1}{2}+1}}$$

$$\exp\left\{-\frac{1}{2\sigma^2}\left[(\underline{\eta}-\underline{\eta}_o(\underline{\gamma}))'\,\Sigma_{22.1}^{-1}\,(\underline{\eta}-\underline{\eta}_o(\underline{\gamma})) + (\underline{\gamma}-\underline{\gamma}_o)'\,\Sigma_{11}^{-1}\,(\underline{\gamma}-\underline{\gamma}_o) + \alpha\right]\right\} \tag{3.20}$$

where $\underline{\eta}_o(\gamma)$ and $\sigma^2\,\Sigma_{22.1}$ are the parameters of the conditional distribution of $\underline{\eta}$ given $\underline{\gamma}$, and $\sigma^2\,\Sigma_{11}$ is the variance-covariance matrix of the marginal distribution of $\underline{\gamma}$ given σ^2.

Equations (3.11) and (3.19) will be used to obtain the conditional posterior distribution of $\underline{\gamma}$ given $\underline{\eta}$ and the marginal distribution of $\underline{\eta}$. Equations (3.12) and (3.20) will be used later on to obtain the conditional posterior distribution of $\underline{\eta}$ given $\underline{\gamma}$ and the marginal distribution of $\underline{\gamma}$.

The prior distributions given by (3.19) or (3.20) constitute a reasonably general family and will be shown to produce conditional posterior distributions that correspond to members of standard families and are easy to work with.

The posterior and predictive distributions will also be given for the "noninformative" prior

$$\Pi(\underline{\eta}, \underline{\gamma}, \sigma^2) \alpha \frac{1}{\sigma 2} \tag{3.21}$$

Posterior Distributions

In this section we will first obtain the conditional posterior distribution of $\underline{\gamma}$ given $\underline{\eta}$, and the marginal distribution of $\underline{\eta}$. The formulas for the conditional distribution of $\underline{\eta}$ given $\underline{\gamma}$, and the marginal of $\underline{\gamma}$ are analogous and their derivation will not be given in detail.

From (3.11) and (3.19), the joint posterior distribution of $\underline{\gamma}$, $\underline{\eta}$ and σ^2 given the vector of observations and the prior hyperparameters is given by

$$\Pi(\underline{\gamma},\underline{\eta},\sigma^2|\underline{Y}) \ \alpha \ \frac{1}{(\sigma^2)^{(N+q+r+2k+3)/2}}$$

$$\exp\left\{-\frac{1}{2\sigma^2}\left[(\underline{z} - X\underline{\gamma})'(\underline{z}-X\underline{\gamma}) + (\gamma-\gamma_{\sigma}(\underline{\eta}))' \Sigma_{11.2}^{-1} (\underline{\gamma}-\underline{\gamma}_o(\underline{\eta})) + (\underline{\eta}-\underline{\eta}_o)' \Sigma_{22}^{-1} (\underline{\eta}-\underline{\eta}_o) + \alpha\right]\right\} \tag{3.22}$$

which can be written as

$$\Pi(\underline{\gamma},\underline{\eta},\sigma^2 \ \underline{Y}) \ \alpha \ \frac{1}{(\sigma^2)^{(N+q+r+2k+3)/2}}$$

$$\exp\left\{-\frac{1}{2\sigma^2}\left[(\underline{\gamma}-\hat{\underline{\gamma}})X'X(\underline{\gamma}-\hat{\underline{\gamma}}) + (\underline{z}-X\hat{\underline{\gamma}})'(\underline{z}-X\hat{\underline{\gamma}}) + (\underline{\gamma}-\underline{\gamma}_o(\underline{\eta}))'\Sigma_{11.2}^{-1} (\underline{\gamma}-\underline{\gamma}_o(\underline{\eta})) + (\underline{\eta}-\underline{\eta}_o)' \Sigma_{22}^{-1} (\underline{\eta}-\underline{\eta}_o) + \alpha\right]\right\} \tag{3.23}$$

where

$$\hat{\underline{\gamma}} = (X'X)^{-1} X'\underline{z} \tag{3.24}$$

After some transformations in the exponent, the posterior distribution corresponds to

$$\Pi(\underline{\gamma},\underline{\eta},\sigma^2|\underline{Y}) \;\alpha\; \frac{1}{(\sigma^2)^{(N+q+r+2k+3)/2}}$$

$$\exp\left\{-\frac{1}{2\sigma^2}\left[(\underline{\gamma}-\underline{\gamma}_*)'(X'X+\Sigma_{11.2}^{-1})\;(\underline{\gamma}-\underline{\gamma}_*)+R\right]\right\} \qquad (3.25)$$

where

$$\underline{\gamma}_* = (X'X+\Sigma_{11.2}^{-1})^{-1}\;(X'X\underline{\gamma}+\Sigma_{11.2}^{-1}\;\underline{\gamma}_o(\underline{\eta})) \qquad (3.26)$$

and

$$R = \hat{\underline{\gamma}}'(X'X)\hat{\underline{\gamma}} + \underline{\gamma}_o(\underline{\eta})'\Sigma_{11.2}\;\underline{\gamma}_o(\underline{\eta}) - \underline{\gamma}_*'(X'X+\Sigma_{11.2}^{-1})\;\underline{\gamma}_*$$

$$+(\underline{z}-X\hat{\underline{\gamma}})'(\underline{z}-X\hat{\gamma}) + (\underline{\eta}-\underline{\eta}_o)'\Sigma_{22}^{-1}\;(\underline{\eta}-\underline{\eta}_o) + \alpha \qquad (3.27)$$

Integrating (3.25) with respect to σ^2 we obtain

$$\Pi(\underline{\gamma},\underline{\eta}|\underline{Y})\alpha\left[1+\frac{(\underline{\gamma}-\underline{\gamma}_*)'(X'X+\Sigma_{11.2}^{-1})(\underline{\gamma}-\underline{\gamma}_*)}{R}\right]^{-\frac{\nu+\;r+s+1}{2}} \qquad (3.28)$$

where $\nu = (N+q+2k-s)$.

Since the conditional distribution of $\underline{\gamma}$ given $\underline{\eta}$ is proportional to the joint, then

$$\Pi(\underline{\gamma}|\underline{\eta},\underline{Y})\alpha\quad\Pi(\underline{\gamma},\underline{\eta}|\underline{Y}) \qquad (3.29)$$

Notice that, for fixed $\underline{\eta}$, (3.29) corresponds to a multivariate t-distribution with ν degrees of freedom (See Zellner (1971), for a detailed discussion on this distribution).

Now, from (3.25), integrating out $\underline{\gamma}$ and σ^2 we obtain the marginal posterior distribution of $\underline{\eta}$ as

$$\Pi(\underline{\eta}|\underline{Y}) \alpha \; |X'X + \Sigma_{11.2}^{-1}|^{-\frac{1}{2}} \; R^{-\frac{N+q+2k-s}{2}} \tag{3.30}$$

with R given by (3.27).

In general, (3.30) does not correspond to any standard distribution and its analysis has to be done using numerical procedures.

If the prior (3.21) is used, equation (3.29) becomes

$$\Pi(\underline{\gamma}|\underline{\eta},\underline{Y}) \alpha \left[1 + \frac{(\underline{\gamma}-\hat{\underline{\gamma}})X'X(\underline{\gamma}-\hat{\underline{\gamma}})}{R_o}\right]^{-(\nu_o+r+s+1)} \tag{3.31}$$

where

$$R_o = (\underline{z}-X\hat{\underline{\gamma}})'(\underline{z}-X\hat{\underline{\gamma}}) \tag{3.32}$$

$$\hat{\gamma} = (X'X)^{-1}X'\underline{z}$$

and $\nu_o = N-2s-p-r-1$. Also, equation (3.30) corresponds to

$$\Pi(\underline{\eta}|\underline{Y}) \alpha |X'X|^{-\frac{1}{2}} \; R_o^{-\frac{N-2s-p-r-1}{2}} \tag{3.33}$$

with R_o as in (3.32).

Following a similar procedure but starting from (3.15) and (3.20), the conditional posterior distribution of $\underline{\eta}$ given $\underline{\gamma}$ is

$$\Pi(\underline{\eta}|\underline{\gamma},\underline{Y}) \alpha \left[1 + \frac{(\underline{\eta}-\underline{\eta}_*)'(Q'Q + \Sigma_{22.1}^{-1})\;(\underline{\eta}-\underline{\eta}_*)}{R_*}\right]^{-\frac{\nu+p+q}{2}} \tag{3.34}$$

where $\nu = (N-p+r+sk+1)$

$$\underline{\eta}_* = (Q'Q + \Sigma_{22.1}^{-1})^{-1} \; (\underline{Q}'Q\hat{\underline{\eta}} + \Sigma_{22.1}^{-1}\;\underline{\eta}_o(\gamma)) \tag{3.35}$$

$$R_* = \underline{\eta}_o(\underline{\gamma})'\;\Sigma_{22.1}^{-1}\;\underline{\eta}_o(\underline{\gamma}) + \hat{\underline{\eta}}'Q'Q\underline{\eta}-\underline{\eta}_*'(\Sigma_{22.1}^{-1} + Q'Q)\underline{\eta}_*$$

$$+(\underline{U}-Q\hat{\underline{\eta}})'(\underline{U}-Q\hat{\underline{\eta}}) + (\underline{\gamma}-\underline{\gamma}_o)'\;\Sigma_{11}^{-1}\;(\underline{\gamma}-\underline{\gamma}_o) \tag{3.36}$$

and, the marginal posterior density of γ will be

$$\Pi(\underline{\gamma}|\underline{Y}) \alpha \; |Q'Q + \Sigma_{22.1}^{-1}|^{-\frac{1}{2}} \; R_*^{-\frac{N+r+2k+1-p}{2}} \tag{3.37}$$

Notice that, again, except probably for some particular cases (3.37) does not correspond to any standard distribution and its analysis has to be done using numerical methods. It is worth mentioning here that (3.34) does correspond to a multivariate t-distribution.

Here again, if the prior (3.21) is used, equation (3.34) becomes

$$\Pi(\underline{\eta}|\underline{\gamma},\underline{Y}) \alpha \left[1 + \frac{(\underline{\eta}-\hat{\underline{\eta}})'Q'Q(\underline{\eta}-\hat{\underline{\eta}})}{R_o^*} \right]^{-\frac{\nu_o^* + p+q}{2}} \tag{3.38}$$

where

$$R_o^* = (\underline{U}-Q\hat{\underline{\eta}})'(\underline{U}-Q\hat{\underline{\eta}}) \tag{3.39}$$

$$\hat{\underline{\eta}} = (Q'Q)^{-1} \; Q'\underline{U}$$

and $\nu_o^* = N-s-2p-q$. Also, equation (3.37) corresponds to

$$\Pi(\underline{\gamma}|\underline{Y}) \alpha \; |Q'Q|^{-\frac{1}{2}} \; (R_o^*)^{-\frac{N-2p-q-s}{2}} \quad . \tag{3.40}$$

It is important to mention that equations (3.31), (3.33), (3.38) and (3.40) can be obtained from the corresponding distributions through a limiting process.

The remaining part of this section will be devoted to construct the one step ahead predictive distribution, conditional on the noise parameters. To do this, equation (3.29) will be used and, the density of Y_{N+1} given the previous observations and of course the parameters of the distribution will be considered.

Predictive Distribution

Perhaps the main reason for the analysis of time series is to draw inferences on the behavior of future observations, this is done from the Bayesian point of view through the predictive distribution. We will consider here the derivation of the one step ahead predictive distribution, conditional to the values of the noise parameters. Since this distribution corresponds to a standard one, we will be able to use it to obtain conditional estimators and to construct regions for the values of the future observation of a given probability value.

Conditional to $\underline{\gamma}, \underline{\eta}, \sigma^2$, the vector of observations $\underline{Y}$, and given $a_t = \hat{a}_t$, t = 1, ..., N, the density function of Y_{N+1} is given by

$$f(Y_{N+1} | \underline{\gamma}, \underline{\eta}, \sigma^2, \underline{Y}) \alpha \frac{1}{(\sigma^2)^{\frac{1}{2}}} \exp\left\{-\frac{1}{2\sigma^2} (z_{N+1} + \underline{X}'_{N+1}\underline{\gamma})^2\right\} \tag{3.41}$$

where

$$z_{N+1} = \phi(B)Y_{N+1} - (\theta(B) - 1)a_{N+1}$$

$$= Y_{N+1} + (\phi(B) - 1)Y_{N+1} - (\theta(B) - 1)\hat{a}_{N+1}$$

$$X_{N+1,j} = \phi(B)\xi_{N+1-j+1} \qquad j = 1, 2, \ldots, r + 1$$

$$X_{N+1,j} = \theta(B)\hat{a}_{N+1-j+r+1} - \phi(B)Y_{N+1-j+r+1} \qquad j = r+2, \ldots, r+s+1$$

and $\underline{\gamma}' = (\underline{\omega}', \underline{\delta}')$.

For simplicity, since z_{N+1} depends on the observed values of Y and on the vector of noise parameters, we will derive its predictive distribution from which that of Y_{N+1} can be easily obtained.

Multiplying (3.26) and (3.41) after some algebraic transformations,

$$\Pi(\underline{\gamma},\sigma^2|\underline{\eta},\underline{Y})\; f(z_{N+1}|\underline{\gamma},\underline{\eta},\sigma^2,\underline{Y})$$

$$\alpha \frac{1}{(\sigma^2)^{(N+q+r+2k+4)/2}}$$

$$\exp\left\{\frac{-1}{2\sigma^2}\left[(\underline{\gamma}-\underline{\gamma}^+)'(X'X + \Sigma_{11.2}^{-1} + \underline{X}_{N+1}\,\underline{X}_{N+1}')(\underline{\gamma}-\underline{\gamma}^+) + R^+\right]\right\} \tag{3.42}$$

From (3.42), after integrating out, first $\underline{\gamma}$ and then σ^2, the predictive distribution of z_{N+1} given $\underline{\eta}$ and $\underline{Y}$ has the form

$$\Pi(z_{N+1}|\underline{\eta},\underline{Y})\alpha|X'X + \Sigma_{11.2}^{-1} + \underline{X}_{N+1}\,\underline{X}_{N+1}'|^{-\frac{1}{2}}$$

$$\left[R^+\right]^{-\frac{N+q-s-2k+1}{2}} \tag{3.43}$$

where,

$$R^+=R+\underline{\gamma}_*'(\Sigma_{11.2}^{-1} + X'X)\underline{\gamma}_*-\underline{\gamma}_*'(X'X+ \Sigma_{11.2}^{-1})(X'X + \Sigma_{11.2}^{-1} + \underline{X}_{N+1}\,\underline{X}_{N+1}')^{-1}$$

$$(X'X + \Sigma_{11.2}^{-1})\,\underline{\gamma}_* - \mu_z^2\left[1 - \underline{X}_{N+1}'(X'X + \Sigma_{11.2}^{-1} + \underline{X}_{N+1}\,\underline{X}_{N+1}')^{-1}\underline{X}_{N+1}\right]$$

$$+\left[1 - \underline{X}_{N+1}'(X'X + \Sigma_{11.2}^{-1} + \underline{X}_{N+1}\,\underline{X}_{N+1}')^{-1}\underline{X}_{N+1}\right]\left[z_{N+1} - \mu_z\right]^2 \tag{3.44}$$

and

$$\mu_z = (1 - \underline{X}_{N+1}'\,(X'X + \Sigma_{11.2}^{-1} + \underline{X}_{N+1}\,\underline{X}_{N+1}')^{-1}\,\underline{X}_{N+1})^{-1}$$

$$(\underline{X}_{N+1}'(X'X + \Sigma_{11.2}^{-1} + \underline{X}_{N+1}\,X_{N+1}')^{-1}\;(X'X + \Sigma_{11.2}^{-1})\underline{\gamma}_*\,. \tag{3.45}$$

Therefore,

$$\Pi(z_{N+1}|\underline{\eta},\underline{Y})\alpha\left[1 + \frac{(z_{N+1}-\mu_z)^2}{R_o^+}\right]^{-\frac{N+q-s+2k+1}{2}} \tag{3.46}$$

with

$$R_o^+ = \Big[R + \underline{\gamma}_*'(\Sigma_{11.2}^{-1} + X'X)\underline{\gamma}^*-\underline{\gamma}_*'(X'X + \Sigma_{11.2}^{-1})(X'X + \Sigma_{11.2}^{-1} + X_{N+1}X_{N+1}')^{-1}$$

$$(X'X + \Sigma_{11.2})\underline{\gamma}_* - \mu_{z}^{2}(1-\underline{X}_{N+1}'(X'X + \Sigma_{11.2}^{-1} + \underline{X}_{N+1}\,\underline{X}_{N+1}')\underline{X}_{N+1})\Big]\Big/$$

$$\left[1 - \underline{X}_{N+1}'(X'X + \Sigma_{11.2}^{-1} + \underline{X}_{N+1}\,\underline{X}_{N+1}')^{-1}\underline{X}_{N+1}\right] \qquad (3.47)$$

Notice that (3.46) corresponds to the density of a t-distribution with parameters μ_{z}, $R_o^{+}/(N+q-s+2k)$ and $N+q-s+2k$ degrees of freedom. From (3.46), the predictive distribution of Y_{N+1} can be easily obtained.

If the prior (3.21) is used, equation (3.46) becomes

$$\Pi(z_{N+1}|\underline{n},\underline{Y}) \alpha \left[1 + \frac{(z_{N+1}-\mu_{z})^2}{R_o^{+}}\right]^{-\frac{N-2s-r-p-1\,+1}{2}} \qquad (3.48)$$

where

$$\mu_{z} = \frac{\underline{X}_{N+1}'(X'X+\underline{X}_{N+1}X_{N+1}')^{-1}X'X\underline{\hat{\gamma}}}{\underline{1} - \underline{X}_{N+1}'(X'X+\underline{X}_{N+1}\underline{X}_{N+1}')^{-1}\underline{X}_{N+1}}$$

$$R_o^{+} = \Bigg[(\underline{z}-X\underline{\hat{\gamma}})'(\underline{z}-X\hat{\gamma}) + \underline{\hat{\gamma}}'X'X\underline{\hat{\gamma}} - \underline{\hat{\gamma}}'X'X(X'X + \underline{X}_{N+1}\,\underline{X}_{N+1}')^{-1}X'X\hat{\gamma}$$

$$- \frac{(X_{N+1}'(X'X+X_{N+1}\,\underline{X}_{N+1}')^{-1}X'X\,\underline{\hat{\gamma}}\,)^2}{(1- X_{N+1}'(X'X+\underline{X}_{N+1}\,\underline{X}_{N+1}')^{-1}\underline{X}_{N+1})}\Bigg] \; (1+\underline{X}_{N+1}'(X'X+\underline{X}_{N+1}\underline{X}_{N+1}')X_{N+1})^{-1}$$

and $\underline{\hat{\gamma}} = (X'X)^{-1}X'\underline{z}$.

In the next section, we will derive the posterior distribution for the linear intervention model. Although the mathematics involved is very similar to that used here, it presents situations which are interesting to mention. Due to the similarities, the presentation in section 3.2 will not be so detailed, and the predictive distribution will not be derived.

3.3 Multiple Linear Interventions Model

The Approximate Likelihood Function

The model to be analyzed in this section corresponds to

$$Y_t = \sum_{i=1}^{m} \omega_i(B)\xi_{t,i} + \frac{\theta(B)}{\phi(B)} a_t \qquad t = 1, 2, \ldots, N \tag{3.49}$$

where

$$\omega_i(B) = \omega_{i,0} + \omega_{i,1} B + \cdots + \omega_{i,r_i} B^{r_i} \qquad i = 1, 2, \ldots, m$$

$$\theta(B) = 1 + \theta_1 B + \cdots + \theta_q B^q$$

$$\phi(B) = 1 + \phi_1 B + \cdots + \phi_p B^p$$

and can be rewritten as

$$\phi(B)Y_t = \sum_{i=1}^{m} \omega_i(B)\ \phi(B)\xi_{t,i} + \theta(B)a_t\ . \tag{3.50}$$

It is important to mention here that the model described by (3.49) or (3.50), is linear in the intervention parameters. It will be evident later that, if the noise was generated by a pure moving average model, the posterior analysis can be done without the need of conditional distributions. Also, we should mention here that, if one has a single linear intervention (m=1) this model is a particular case of the one described by (3.2) and (3.3).

From (3.50), we can express the residuals a_t in terms of the observations and parameters as

$$a_t = \phi(B)Y_t - \sum_{i=1}^{m} \sum_{j=1}^{r_i} \omega_{i,j} B^j \phi(B)\xi_{t,i}$$

$$- \theta_1 Ba_t - \cdots - \theta_q B^q a_t \tag{3.51}$$

and again, since the Jacobian of the transformation, assuming the first p observations as fixed and the first p residuals with a value of zero is equal to one, the likelihood function can be written as

$$\ell(\underline{\gamma},\underline{\eta},\sigma^2|\underline{Y})\alpha \frac{1}{(\sigma^2)^{\frac{N-p}{2}}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=p}^{N} a_t^2\right\} \tag{3.52}$$

where $\underline{\gamma}' = (\underline{\omega}_1', \underline{\omega}_2', \ldots, \underline{\omega}_k', \underline{\theta}')$, $\underline{\theta}' = (\theta_1, \ldots, \theta_q)$, $\underline{\omega}_i' = (\omega_{i,o}, \ldots, \omega_{i,r_i})$ $i = 1, 2, \ldots, m$, $\underline{\phi}' = (\phi_1, \ldots, \phi_p)$ and, a_t, $t = p, \ldots N$, are given by (3.51).

Needless to say, here again, due to the nonlinearities in the model, the analysis becomes very complicated unless we use an approximation of the likelihood as in the previous section. We will again substitute the unobservable variables a_t, in the right hand side or (3.51), by the observable variable $\hat{a}_t$ calculated recursively from (3.51) using the maximum likelihood estimators of the parameters as the parameter values. The approximate likelihood function corresponds in this case to

$$\ell^*(\underline{\gamma},\underline{\phi},\sigma^2|\underline{Y})\alpha \frac{1}{(\sigma^2)^{\frac{N-p}{2}}} \exp\left\{-\frac{1}{2\sigma^2}(\underline{Z}-X\underline{\gamma})'(\underline{Z}-X\underline{\gamma})\right\} \tag{3.53}$$

where

$$Z_t = \phi(B)Y_t \tag{3.54}$$

and the rows of the matrix X are given by

$$\underline{X}_t' = (\phi(B)\xi_{t_1}, \ldots, B^{r_1}\phi(B)\xi_{t_1}, \ldots, \phi(B)\xi_{t,m}, \ldots, \phi(B)B^{r_m}\xi_{t,m},$$
$$\hat{a}_{t-1}, \ldots, \hat{a}_{t-q}). \tag{3.55}$$

Equation (3.53) will be used to obtain the conditional posterior distribution of $\underline{\gamma}$ given $\underline{\phi}$. Notice that $\underline{Z}$ and X are both functions of $\underline{\phi}$. Formula (3.53) is also equivalent to

$$\ell^*(\underline{\gamma},\underline{\eta},\sigma^2|\underline{Y}) = \frac{1}{(\sigma^2)^{\frac{N-p}{2}}} \exp\left\{-\frac{1}{2\sigma^2}(\underline{U}-Q\underline{\eta})'(\underline{U}-Q\underline{\eta})\right\} \quad (3.56)$$

where

$$U_t = Y_t - \sum_{i=1}^{m} \omega_i(B)\xi_{t,i} + \sum_{i=1}^{q} \theta_i a_{t-i} \quad (3.57)$$

and the rows of Q are

$$\underline{Q}_t' = (B(\sum_{i=1}^{m} \omega_i(B)\xi_{t,i} - Y_t), \ldots, B^p(\sum_{i=1}^{m} \omega_i(B)\xi_{t,i} - Y_t)) \quad (3.58)$$

for i = p, ..., N.

This last version of the likelihood will be used to derive the conditional posterior distribution of the autoregressive parameters given the intervention and moving average parameters.

<u>The Prior Distribution</u>

Using (as in section 3.1) a multivariate normal for the conditional distribution of the model parameters given σ^2, and an inverted-gamma for the marginal of σ^2 we have that

$$\Pi(\gamma,\phi,\sigma^2)\alpha \frac{1}{(\sigma^2)^{\frac{\Sigma r_i+m+p+q+2k}{2}+1}}$$

$$\exp\left\{-\frac{1}{2\sigma^2}\left[(\underline{\gamma}-\underline{\gamma}_o(\underline{\phi}))' \Sigma_{11.2}^{-1}(\underline{\gamma}-\underline{\gamma}_o(\phi)) + (\underline{\phi}-\underline{\phi}_o)' \Sigma_{22}^{-1}(\phi-\phi_o) + \alpha\right]\right\} \quad (3.59)$$

where $\sigma^2\ \Sigma_{11.2}$ and $\underline{\gamma}_o(\phi)$ are the parameters of the conditional distribution of $\underline{\gamma}$ given $\underline{\phi}$ and σ^2, and $\sigma^2\Sigma_{22}$ is the variance-covariance matrix of the marginal distribution of ϕ given σ^2.

Formula (3.59) is equivalent to

$$\Pi(\gamma\ \phi\ \sigma\ )\alpha\ (\sigma^2)^{-\frac{\Sigma r_i+m+p+q+2k}{2}+1}$$

$$\exp\left\{-\frac{1}{2\sigma^2}\left[(\underline{\phi}-\underline{\phi}_o(\underline{\gamma}))'\Sigma_{22.1}^{-1}(\underline{\phi}-\underline{\phi}_o(\underline{\gamma})) + (\underline{\gamma}-\underline{\gamma}_o)'\Sigma_{11}^{-1}(\gamma-\gamma_o) + \alpha\right]\right\} \tag{3.60}$$

where $\underline{\phi}_o(\underline{\gamma})$ and $\sigma^2\Sigma_{22.1}$ are the parameters of the conditional distribution of $\underline{\phi}$ given $\underline{\gamma}$ and σ^2, and $\sigma^2\Sigma_{11}$ is the variance-covariance matrix of the marginal distribution of $\underline{\gamma}$ given σ^2.

As in the previous section, (3.59) and (3.60) will be used to obtain the posterior distribution as needed.

The Posterior Distribution

Trying to avoid the details of a derivation which is analogous to the one presented in the previous section, the conditional posterior distribution of $\underline{\gamma}$ given $\underline{\phi}$, obtained from (3.53) and (3.59) corresponds to

$$\Pi(\underline{\gamma}|\underline{\phi},\underline{Y})\alpha\left[1 + \frac{(\underline{\gamma}-\underline{\gamma}_*)'(X'X+\Sigma_{11.2}^{-1})(\underline{\gamma}-\underline{\gamma}_*)}{R}\right]^{-\frac{N+2k+\sum_{i=1}^{m} r_i+m+q}{2}} \tag{3.61}$$

where

$$\underline{\gamma}_* = (X'X+\Sigma_{11.2}^{-1})^{-1}(X'X\hat{\underline{\gamma}}+\Sigma_{11.2}^{-1}\gamma_o(\phi)) \tag{3.62}$$

$$\hat{\underline{\gamma}} = (X'X)^{-1}X'\underline{z} \tag{3.63}$$

and

$$R = \hat{\underline{\gamma}}'(X'X)\hat{\underline{\gamma}} + \underline{\gamma}_o(\underline{\phi})'\Sigma_{11.2}^{-1}\underline{\gamma}_o(\underline{\phi}) - \underline{\gamma}_*'(X'X+\Sigma_{11.2}^{-1})\underline{\gamma}_*$$
$$+(\underline{z}-X\hat{\gamma})'(\underline{z}-X\hat{\gamma}) + \alpha + (\underline{\phi}-\underline{\phi}_o)'\Sigma_{22}^{-1}(\phi-\underline{\phi}_o). \tag{3.64}$$

The marginal posterior distribution of $\underline{\phi}$ is given by

$$\Pi(\underline{\phi}|\underline{Y})\alpha|X'X+\Sigma_{11.2}^{-1}|^{-\frac{1}{2}}R^{-(\sum_{i=1}^{m} r_i+m+N+q+2k)/2} \tag{3.65}$$

Also, using now (3.56) and (3.60), the conditional posterior distribution of $\underline{\phi}$ given $\underline{\gamma}$ corresponds to

$$\Pi(\underline{\phi}|\underline{\gamma},\underline{Y}) \propto \left[1 + \frac{(\underline{\phi}-\underline{\phi}^*)'(Q'Q + \Sigma_{22.1}^{-1})(\underline{\phi}-\underline{\phi}^*)}{R^*}\right]^{-\frac{N+\Sigma r_i+m+q+2k-p\ +p}{2}} \tag{3.66}$$

where

$$\underline{\phi}^* = (Q'Q + \Sigma_{22.1}^{-1})^{-1}(Q'Q\hat{\underline{\phi}} + \Sigma_{22.1}\,\underline{\phi}_o(\underline{\gamma})) \tag{3.67}$$

$$\hat{\underline{\phi}} = (Q'Q)^{-1}Q'\underline{U} \tag{3.68}$$

and

$$R^* = \hat{\underline{\phi}}'Q'Q\hat{\underline{\phi}} + \underline{\phi}_o(\gamma)'\,\Sigma_{22.1}^{-1}\,\underline{\phi}_o(\underline{\gamma}) + \underline{\phi}^{*\prime}(Q'Q + \Sigma_{22.1}^{-1})\,\underline{\phi}^*$$

$$+(\underline{U}-Q\hat{\underline{\phi}})'(\underline{U}-Q\hat{\underline{\phi}}) + \alpha + (\underline{\gamma}-\underline{\gamma}_o)'\,\Sigma_{11}^{-1}\,(\underline{\gamma}-\gamma_o) \tag{3.69}$$

The marginal posterior distribution of $\underline{\gamma}$ is then

$$\Pi(\underline{\gamma}|\underline{Y}) \propto |Q'Q + \Sigma_{22.1}^{-1}|^{-\frac{1}{2}}(R^*)^{-\frac{N+\sum_{i=1}^{m} r_i+m+q+2k}{2}} \tag{3.70}$$

Notice that (3.61) and (3.66) correspond to multivariate t-distributions and their properties are well known (see Zellner (1971)). In general the marginal posterior densities (3.65) and (3.70) do not correspond to standard distributions and have to be analyzed numerically. In (3.65) one usually does not have more than two variables but (3.70) may have more in which case the analysis becomes complicated if not impossible.

3.3 Comments

The formulas derived in the previous sections of this chapter will be used in Chapter IV to implement a time series analysis and also in

Chapter V where we will illustrate their use with several examples.

Notice that nothing was done to obtain the posterior distribution of the parameter σ^2 and this was because we are mainly interested in the modeling process and also because there is no way in which we can obtain a general closed form for the marginal posterior distribution. Of course, the conditional posterior distribution of σ^2 given all the other parameters can be easily obtained and is proportional to the joint posterior distribution.

CHAPTER IV

TIME SERIES ANALYSIS

The purpose of this chapter is to present procedures to model the intervention effect through the three stages used by Box and Jenkins (1970) to analyze a time series, namely, identification, estimation, and diagnostic checking. The predictive distribution derived in the previous chapter will be used to draw inferences about a future observation.

Throughout this chapter we will assume that the noise model has been identified using the observations prior to the intervention, and that reliable estimators are available for the corresponding parameters. Only in the estimation part we will propose a procedure to jointly estimate all the parameters of the intervention model using the joint posterior mode.

4.1 Identification

For a time series that was generated by an ARMA(p,q) process, Broemeling and Shaarawy (1984b) proposed an identification procedure in which, after obtaining the posterior distribution of the parameters using p=q=2, a series of tests are performed in an attempt to reduce the order of the model. The procedure proposed here to identify the intervention transfer function is analogous. Since the noise model does not change with the intervention, the observations obtained prior to it, will enable us to use the Bayesian identification procedure mentioned above and to estimate the noise parameters through their posterior distribution. The point estimators will be used as the parameter values for the conditional distribution of the intervention parameters give the noise ones. We will use equation (3.29) to start our identification process.

The discussion here will be centered on the single intervention model (3.3) but, the procedure can be easily extended to the other model analyzed in the previous chapter.

As in the procedure proposed by Broemeling and Shaarawy (1984b), we will start our identification process assuming r=s=2. Although this seems to be restrictive, we believe that this model is flexible enough to handle a wide variety of situations. See, for example, Box and Jenkins (1970) for a detailed description of the different effects that can be modeled using this transfer function when the intervention variable corresponds to a pulse or step function. The calculation of the conditional posterior distribution (3.29) with r=s=2 will be the starting point of a series of tests that will lead us to find a model that reasonably fits the observed data. To do this we propose the following steps:

Step 1

Test the null hypothesis H_o: $\omega_2 = 0$. This can be done by constructing a Highest Posterior Density (HPD) interval, for a given probability level, for ω_2 using its marginal distribution (univariate t). If the value of zero is an element of the interval we should accept the null hypothesis and proceed with the second step. Otherwise, we can assume that the degree of the polynomial $\omega(B)$ is equal to two and proceed to step 4 to identify the degree of $\delta(B)$.

<u>Step 2</u>

Test the null hypothesis H_o: $\omega_1 = 0$ given that $\omega_2 = 0$. Here we proceed as in step one except that we should use the conditional distribution of ω_1 given that $\omega_2 = 0$. If the null hypothesis is accepted one should proceed with step 3. If this is not the case, the degree of $\omega(B)$ can be assumed to be equal to one and start with step 4.

<u>Step 3</u>

Test the null hypothesis H_o: $\omega_o = 0$ given that $\omega_1 = 0$ and $\omega_2 = 0$. Now, if the null hypothesis is accepted we can conclude that the intervention does not produce any effect, and use an ARMA(p,q) model for the inferential process. On the contrary, if the hypothesis is rejected, $\omega(B)$ has only a constant term and we should proceed with step 4 to determine the value of s.

<u>Step 4</u>

Test the null hypothesis H_o: $\delta_2 = 0$ given that some of the parameters in the numerator are equal to zero. This depends upon the step from which we came. If we reject H_o, the degree of the polynomial $\delta(B)$ is equal to two and the identification is finished. Otherwise we should proceed with step 5.

<u>Step 5</u>

Test the hypothesis H_o: $\delta_1 = 0$ given that $\delta_2 = 0$ and also that the corresponding parameters of $\omega(B)$ are equal to zero as in step 4. Here, if H_o is rejected the degree of $\delta(B)$ is one and the identification procedure is terminated. Otherwise, $\delta(B)$ is to be taken as 1.

At this point we are ready to start the estimation process which will be described in section 4.2. Notice that in the above testing sequence we tested first the parameters in the numerator one at a time. The idea behind this is that if the intervention does not have any effect the identification will end earlier. Another possibility is to test the hypothesis that all the numerator parameter are jointly equal to zero before starting with step 1. To do this we should determine an HPD region for the vector $(\omega_o, \omega_1, \omega_2)$ and check if it contains the null vector. If this is the case we should conclude that there is no intervention effect and terminate the identification process. This is analogous to what is usually done in regression analysis with the overall test of the model.

4.2 Estimation

Under the assumption that the noise model was identified using the observations prior to the intervention and that, as a result of the identification procedure described in the previous section, we now have the values of p,q, r, and s that will be used to obtain estimators of the parameters of the model. We will discuss here three methods that can be used to obtain those estimators, depending on the type of inferences that we want to make and, perhaps on the complexity of the model to be used.

Method 1: Conditional Estimation

If the number of observations prior to the intervention is large and, if the posterior distribution obtained through these observations guarantees the reliability of the noise estimators, we can proceed to estimate the intervention parameters using the conditional posterior distribution (3.29).

If a quadratic loss funtion is used, the point estimators correspond to the conditional posterior mean vector which can be easily obtained from the parameters of the distribution.

Also, if HPD regions are desired, they can be obtained using the fact (see Box and Tiao (1973)) that the quadratic form

$$F = (\underline{\gamma}-\underline{\gamma}_*)'(X'X + \Sigma_{11.2}^{-1})(\underline{\gamma}-\underline{\gamma}_*)/((r+s+1)R/\nu) \tag{4.1}$$

has an F distribution with (r+s+1) and ν degrees of freedom in the numerator and denominator respectively and, that the multivariate t-distribution is monotone decreasing function of it. Therefore, a region described by

$$\{\underline{\gamma} \in R^{r+s+1} | F \leq F_{r+s+1,\nu,p}\} \tag{4.2}$$

where $F_{r+s+2,\nu,p}$ is the p-th quantile of the F distribution with the corresponding degrees of freedom is the mentioned HPD region. Notice that the region described by (4.2) corresponds to an r+s+1-dimensional elipsoid.

It is important to mention that the point estimators and HPD regions described before depend upon the conditioning values of the noise parameters and that in some cases it would be very important to study their sensitivity.

Method 2: Joint Posterior Mode

If only point estimators are needed, the joint mode of the posterior distribution can be obtained through the conditional posterior distributions (3.29) and (3.31). An iterative procedure has to be used since the modes of the conditional distributions $\underline{\gamma}_*$ and $\underline{\eta}_*$, see (326) and (3.35), depend upon the values used for the respective conditioning parameters.

To start the procedure, we can use the estimators of the noise parameters obtained prior to the intervention. This will produce an initial value for $\underline{\gamma}_*$, say $\underline{\gamma}_*^{1}$, which can be used to calculate $\underline{\eta}_*^{1}$. Now, $\underline{\eta}_*^{1}$ can be used to obtain $\underline{\gamma}_*^{2}$, and so on. This iterative procedure should be continued until certain stability is reached in the values obtained. Notice that a regression problem has to be solved in each of the iteration steps to obtain the values of $\underline{\gamma}$ (or $\underline{\eta}$) and that the calculation of the inverse of a matrix is also required.

Method 3: Joint Posterior Means

If we are only interested on the point estimators for $\underline{\gamma}$ and $\underline{\eta}$, the joint posterior mean can be used. Depending on the particular case this can be calculated in different ways.

For example, if the number of parameters of the intervention and noise models is small, their expectation can be obtained numerically using the marginal posterior densities (3.30) and (3.37). Now, if only the number of parameters in the noise model is small, one can obtain first the conditional expectation of $\underline{\gamma}$ given $\underline{\eta}$ from (3.29) and obtain its expectation numerically using the marginal distribution of $\underline{\eta}$ given by (3.30), from which the posterior expectation should be obtained.

Each of the methods described above has advantages and disadvantages over the others, and the selection of the one to be used depends on the particular problem that has to be analyzed.

Notice that methods 1 and 3 produce new estimators for the noise parameters that could be useful to check the validity of the identified model. This topic will be the subject of the next section.

4.3 Diagnostic Checking

After the identification of the model for the intervention effect, one would like to check its validity before using it for inferential purposes. We will propose here a procedure to be used when the assumptions about the noise model are believed to be correct. We will use therefore the conditional posterior distribution of the intervention parameters given the noise ones. Whether we should check the validity of the noise model or not, with all the information available depends on the particular situation but, to do this we can develop a procedure analogous to the one to be described below.

Let us assume that in the identified model, $r = r_o$ and $s = s_o$. To check the validity of the intervention model we propose the following steps:

<u>Step 1</u>

With $r = r_o + 1$ and $s = s_o$, test the null hypothesis H_o: $\omega_{r_o+1} = 0$ using an HPD interval constructed with the marginal posterior distribution of ω_{r_o+1} given the noise parameters as in section 4.1. Depending on whether the null hypothesis is accepted or not, use $r = r_o$ or $r = r_o + 1$ to proceed with step 2.

<u>Step 2</u>

With the value of r determinated in the previous step, assume that $s = s_o + 1$ and test the null hypothesis H_o: $\delta_{s_o+1} = 0$ using the marginal posterior distribution of δ_{s_o+1} given the noise parameters. Take $s = s_o$ or $s = s_o + 1$ depending on whether H_o is rejected or not, and use the resultant model in future inference processes.

Notice that what we propose is to increase the degree of the polynomials, first in the numerator and then in the denominator by one, and check if the observed data support the augmented model. Although one can think on other ways of checking the model, this seems to be a feasible one. A similar procedure can be used to check the noise model using all the observations, but in this case we should use the conditional distribution of the noise parameters given the intervention ones. Whether we should do this or not depends upon the particular application and as in any modeling procedure the experimenter has to make the final decision.

4.4 Forecasting

Although the study of the nature of the intervention effect is an important problem by itself, in many situations we would like to draw inferences about the behavior of future observations. This can be accomplished through the predictive distribution given by (3.46).

If a quadratic loss function seems to be reasonable, the estimator of the future observation is the mean of the predictive distribution. From (3.46) we can easily obtain the conditional mean and, if necessary, one can obtain its expectation using (3.30). Conditional HPD intervals can also be obtained from (3.46) for a given probability level.

4.5 Comments

Due to the complexity of the problem analyzed, we were forced to propose procedures that are based on conditional distributions. We cannot claim that these are optimal and we should stress the fact that at least, an analysis of the sensitivity of the conditional results should be made before jumping to conclusions in any given situation.

To avoid repetitions in the presentation, we did not present the case in which one possibly has more than one intervention. The procedures described before can be easily adapted to analyze cases in which a multiple linear intervention model should be used.

CHAPTER V

NUMERICAL STUDY

This chapter will be devoted to present the results obtained from the application of the formulas derived in Chapter III and the procedures proposed in Chapter IV to two examples.

The estimation procedure will be illustrated in section 1, through the analysis of a data set of the monthly averages of the oxidant (O_3) level in Downtown Los Angeles. In section 2, artificially generated data will be used to illustrate the whole time series analysis procedure presented in Chapter IV. Some comments on the results will be given in section 3.

5.1 Bayesian Estimation

In this section we will obtain the Bayesian estimators for the parameters of the model proposed by Box and Tiao (1975), to analyze the series of monthly averages of oxidant (O_3) level in Downtown Los Angeles from January 1955 to December 1972. The reader should refer to that paper for a detailed justification of the model used and a description of the nature of the interventions involved. We will assume here that no prior information is available and we will use the improper prior

$$\Pi(\omega_o, \omega_1, \omega_2, \delta_1, \delta_2, \sigma^2) = 1/\sigma^2 \tag{5.1}$$

The model proposed by Box and Tiao corresponds to

$$Y_t = \omega_{01}\,\xi_{t1} + \frac{\omega_{02}\xi_{t2}}{1-B^{12}} + \frac{\omega_{03}\xi_{t3}}{1-B^{12}} + \frac{(1-\theta_1 B)(1-\theta_2 B^{12})}{1-B^{12}}\,a_t \tag{5.2}$$

where

$$\xi_{t1} = \begin{cases} 0, & t < \text{January, 1960} \\ 1, & t \leq \text{January, 1960} \end{cases}$$

$$\xi_{t2} = \begin{cases} 1, & \text{"summer" months June-October beginning 1966} \\ 0, & \text{otherwise} \end{cases}$$

$$\xi_{t3} = \begin{cases} 1, & \text{"winter" months November-May beginning 1966} \\ 0, & \text{otherwise} \end{cases}$$

and a_t, $t = 1, \ldots, 216$ are assumed to be independent identically distributed random variables with mean zero and variance σ^2.

Obviously, the model (5.1) corresponds to what we called a multiple linear intervention model, except that the noise does not correspond to the one described by (3.49) since a seasonal effect is considered here. Nevertheless, we were able to work out the posterior analysis using the conditional distribution of the intervention parameters (ω_{01}, ω_{02} and ω_{03}) given θ_1 and θ_2, and the marginal distribution of the noise parameters.

For the prior (5.1), the conditional distribution of the intervention parameters given θ_1 and θ_2, corresponds in this case to

$$\Pi(\omega_{01}, \omega_{02}, \omega_{03} | \theta_1, \theta_2, \underline{Y})$$

$$\alpha \left[1 + \frac{(\underline{\gamma}-\hat{\underline{\gamma}})X'X(\underline{\gamma}-\hat{\underline{\gamma}})}{(s^2)\gamma}\right]^{-\frac{\nu+3}{2}} \tag{5.3}$$

where

$$\underline{\gamma}' = (\omega_{01}, \omega_{02}, \omega_{03})$$

$$X = \begin{bmatrix} \xi_{13,1} - \xi_{1,1} & \xi_{13,2} & \xi_{13,2} \\ \xi_{14,1} - \xi_{2,1} & \xi_{14,2} & \xi_{14,2} \\ \cdots\cdots & \cdots & \cdots \\ \xi_{N,1} - \xi_{N-12,1} & \xi_{N,2} & \xi_{N,3} \end{bmatrix}$$

$$z_t = Y_t - Y_{t-12} + \theta_1 \hat{a}_{t-1} + \theta_2 \hat{a}_{t-2} - \theta_1\theta_2 \hat{a}_{t-13} \quad t = 13, \ldots, N$$

$$\hat{\underline{\gamma}} = (X'X)^{-1} X'\underline{z}$$

$$s^2 = (\underline{z}-X\hat{\underline{\gamma}})'(\underline{z}-X\hat{\underline{\gamma}})/\nu$$

and $\nu = N - 15$.

Notice that the matrix X is independent of the noise parameters and, therefore the correlation structure of $\underline{\gamma}$ is independent of the conditioning values used. Also notice that $\hat{\underline{\gamma}}$ and s^2 depend upon θ_1 and θ_2 through $\underline{z}$.

Now, the marginal posterior distribution of the noise parameters is given by

$$\Pi(\theta_1, \theta_2) \alpha \left[\nu\ s^2\right]^{-\frac{\nu-2}{2}} \tag{5.4}$$

with s^2 and ν as in (5.3).

The values of the conditional maximum likelihood estimators for the data taken from the SAS/ETS User's Guide (page 107) correspond to

$\theta_1 = .2998$, $\theta_2 = .5923$, $\omega_{01} = -1.2624$, $\omega_{02} = -.2615$ and $\omega_{03} = -.08196$. These values were used to calculate the residuals estimators and to obtain the parameters for (5.3) and (5.4), which correspond to

$$X'X = \begin{pmatrix} 12 & 0 & 0 \\ 0 & 35 & 0 \\ 0 & 0 & 45 \end{pmatrix} \tag{5.5}$$

$$\underline{\hat{\gamma}}' = (-1,1297, -.2320, -.06456) \tag{5.6}$$

$\nu = 211$, and $s^2 = 0.69015$ if the values for θ_1 and θ_2 are $-.2998$ and $.5923$ respectively.

The conditional point estimator for $\underline{\gamma}$ is given by $\underline{\hat{\gamma}}$ if a quadratic loss function is considered. Also from (5.3), HPD regions can be obtained and used to test some hypotheses of interest. For example, to test the null hypothesis that $\omega_{03} = 0$, i.e. to test whether there is any effect of the winter intervention, we can use the fact that, given $\theta_1 = -.2998$ and $\theta_2 = .5923$,

$$\frac{\omega_{03} + .0646}{.11867}$$

has a t-distribution with 211 degrees of freedom and that the 95% HPD region for ω_{03} corresponds to

$$-.29721 < \omega_{03} < .16801 .$$

Since the value of zero is included in the region, we can conclude that there was no effect of the winter intervention. Of course, this is a conditional test and the conclusion depends upon the selected values of the noise parameters. To check the sensitivity of the test we calculated the limits of the interval for various values of the noise parameters in the rectangle

$$\{(\theta_1, \theta_2);\ -.5 < \theta_1 < .1 \text{ and } .3 < \theta_2 < 1\}$$

outside of which the volume of their posterior distribution is negligible, and in all the cases the value of zero was included.

So far, in terms of point estimation what we have done is to apply what we called method I in the previous chapter. In this example, since the posterior distribution of the noise parameters is not a standard one, we did not use method II, but we will present the results obtained through the application of method III.

Using the numerical integration subroutine DBLIN from the IMSL library, we calculated the first and second moments of the joint posterior distribution. The means, which are the point estimators of the parameters of the model if a quadratic loss function is adequate, are

$$\begin{aligned} E[\omega_{01}] &= -1.1343 \\ E[\omega_{02}] &= -0.2309 \\ E[\omega_{03}] &= -0.0667 \\ E[\theta_1] &= -0.2725 \\ E[\theta_2] &= 0.5824 \end{aligned} \tag{5.7}$$

The variance-covariance matrix is

$$\begin{bmatrix} .058495 & .000027 & .000063 & -.000302 & .001149 \\ .000027 & .020138 & -.000004 & .000338 & .000439 \\ .000063 & -.000004 & .015646 & -.000198 & .000178 \\ -.000302 & .000338 & -.000198 & .003879 & .000545 \\ .001149 & .000439 & .000178 & .000545 & .004882 \end{bmatrix} \tag{5.8}$$

and the correlation matrix corresponds to

$$\begin{bmatrix} 1.0000 & .0008 & .0021 & 0.0201 & .0679 \\ .0008 & 1.0000 & -.0002 & .0382 & .0443 \\ .0021 & -.0002 & 1.0000 & -.0255 & .0203 \\ -.0201 & .0382 & -.0255 & 1.0000 & .1252 \\ .0679 & .0443 & .0203 & .1252 & 1.0000 \end{bmatrix} \tag{5.9}$$

The values obtained for the point estimators through the approximate Bayesian procedure do not differ much from the conditional maximum likelihood estimators and, from (5.6) and (5.7) we can observe that the conditional and marginal estimators of the intervention parameters are almost the same. The variances shown in (5.8) have the same magnitude as the ones reported by Box and Tiao (1975). Notice that the correlations in (5.9) are all very small, most of them negligible.

5.2 Time Series Analysis

In this section we will present the results obtained from the application of the time series analysis procedures proposed in the previous chapter, to an artificially generated data set.

For this example, 200 observations were generated using the RANNOR subroutine described in the SAS User's Guide: Basics (1982), and the model

$$Y_t = \frac{.3}{1-.9B} \xi_t + (1-.25B)\, a_t \quad t = 1, \ldots, 200 \tag{5.10}$$

where

$$\xi_t = \begin{cases} 1, & t \le 100 \\ 0, & \text{otherwise} \end{cases} \tag{5.11}$$

and a_t, $t = 1, \ldots, 200$, are independent, normally distributed with mean zero and variance $\frac{1}{4}$.

Throughout this section, we will use the Bayesian estimator of the parameter of the noise model, obtained from the first 100 observations, using the procedure proposed by Shaarawy (1984) after following the identification procedure of Shaarawy and Broemeling (1984b) for the conditioning value of θ_1. The value obtained was $\theta_1 = -.355723$.

Identification

We started the identification stage assuming that r = s = 2 as proposed in section 4.1, i.e. we used an initial model

$$Y_t = \frac{\omega_0+\omega_1 B+\omega_2 B^2}{1+\delta_1 B+\delta_2 B^2} \xi_t + (1+\theta_1 B)a_t \qquad (5.12)$$

with a_t as in (5.11). The conditional posterior distribution of $\underline{\gamma}' = (\omega_0, \delta_1, \delta_2, \omega_1, \omega_2)$ given $\theta_1 = -.355723$ is given by equation (3.31). The parameters calculated for this particular series are

$$(X'X)= \begin{bmatrix} 100.000 & -276.930 & -272.975 & 99.000 & 98.000 \\ -276.930 & 797.948 & 792.458 & -276.026 & -275.270 \\ -272.975 & 792.458 & 789.033 & -272.978 & -272.992 \\ 99.000 & -276.026 & -272.978 & 99.000 & 98.000 \\ 98.000 & -275.270 & -272.992 & 98.000 & 98.000 \end{bmatrix}$$

$$\hat{\underline{\gamma}} = \begin{bmatrix} .92892 \\ -.53442 \\ -.32034 \\ -.66333 \\ .17305 \end{bmatrix}$$

and $s^2 = R_o/\nu_o = .248359$.

The different steps of the identification procedure produced the following results:

Step 1

Test: H_o: $\omega_2 = 0$

95% HPD intervals: $-1.0219 < \omega_2 < 1.3590$

Result: $\omega_2 = 0$

Step 2

Test: H_o: $\omega_1 = 0$ given $\omega_2 = 0$

95% HPD interval: $-1.6430 < \omega_1 < 0.6265$

Result: $\omega_1 = 0$

Step 3

Test: H_o: $\omega_0 = 0$ given $\omega_1 = \omega_2 = 0$

95% HPD interval: $0.05828 < \omega_0 < 1.04635$

Result: $\omega_0 \neq 0$

Step 4

Test: H_o: $\delta_2 = 0$ given $\omega_1 = \omega_2 = 0$

95% HPD interval: $-1.1505 < \delta_2 < 0.24675$

Result: $\delta_2 = 0$

Step 5

Test: H_o: $\delta_1 = 0$ given $\delta_2 = \omega_1 = \omega_2 = 0$

95% HPD interval: $-1.01469 < \delta_1 < -.69021$

Result: $\delta_1 \neq 0$

The resulting model is therefore

$$Y_t = \frac{\omega_0}{1+\delta_1 B} \xi_t + (1+\theta_1 B)a_t \qquad (5.13)$$

which corresponds to the one used to generate the series.

Estimation

For the model (5.13), the conditional posterior distribution of ω_0 and δ_1 given $\theta_1 = -.355723$ corresponds to (3.31) with the following parameters:

$$X'X = \begin{pmatrix} 100.000 & -275.493 \\ -275.493 & 797.743 \end{pmatrix}$$

$$\underline{\hat{\gamma}} = \begin{pmatrix} .52717 \\ -.82294 \end{pmatrix}$$

and $s^2 = R_o/\nu_o = 1.239026$.

Figures 1 and 2 show the behavior of the series and the values taken by the identified intervention function using the conditional mean $\hat{\gamma}$ to estimate ω_0 and ω_1, and the maximum likelihood estimates of γ.

Diagnostic Checking

Following the procedure described in section 4.3, the number of parameters was increased by one, first in the numerator and then in the denominator, and the hypothesis that the added parameter is equal to zero tested. The results obtained are shown below.

Case 1: $r = 1$ and $s = 2$

Test: $\delta_2 = 0$

95% HPD interval: $-1.7244 < \delta_2 < 1.1984$

Result: $\delta_2 = 0$

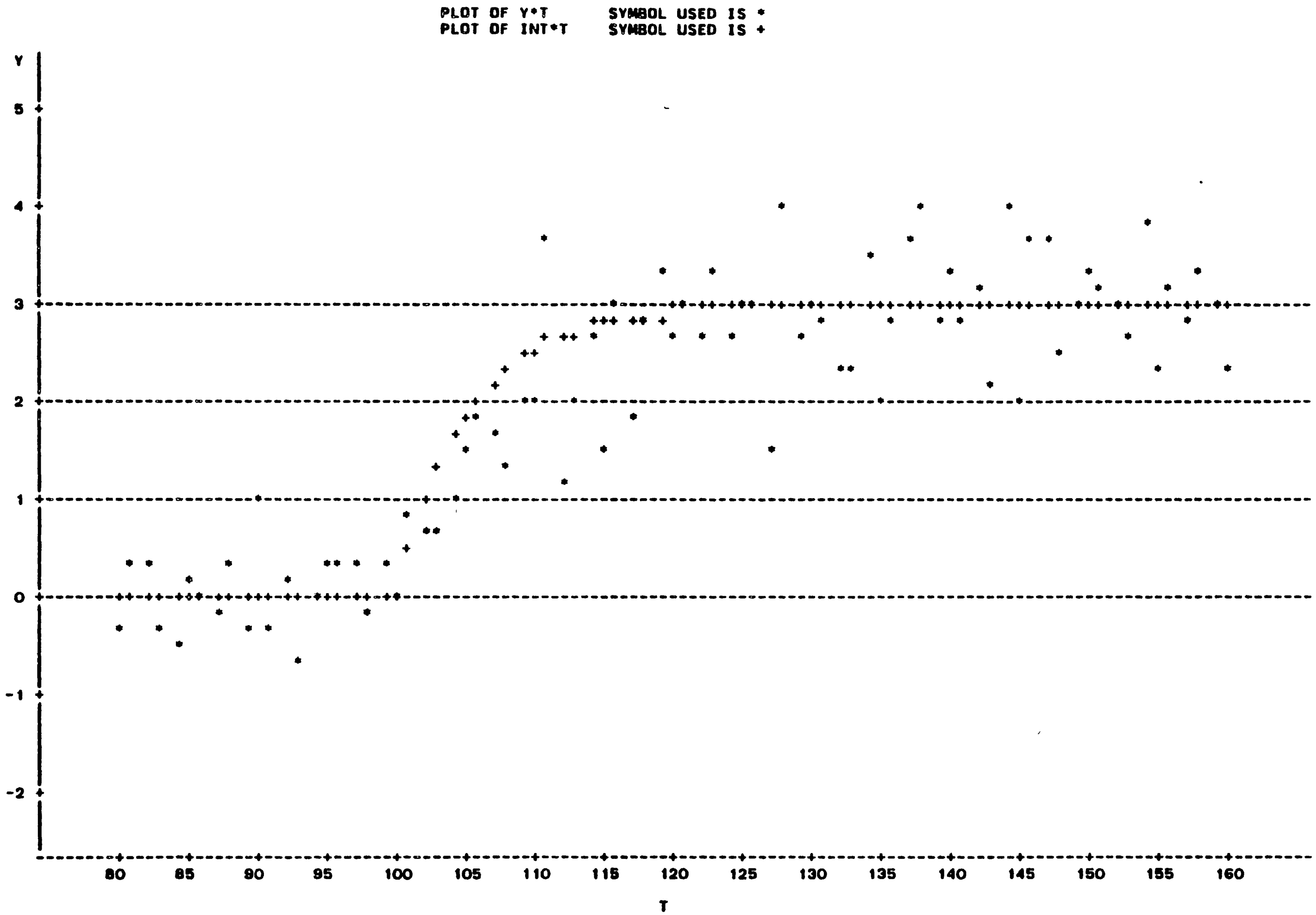

Figure 1. Bayesian Estimators

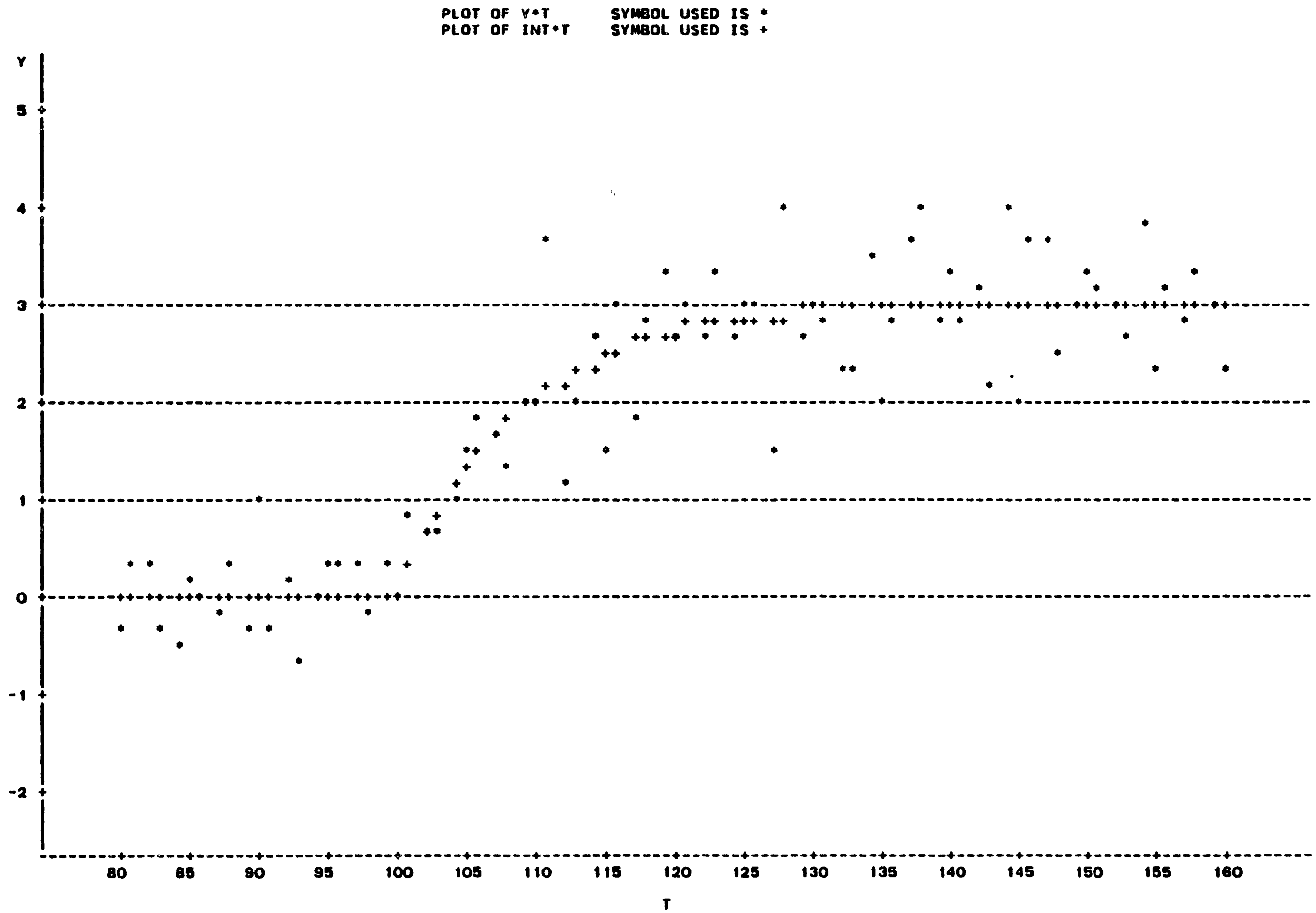

Figure 2. ML Estimators

Case 2: $r = 2$ and $s = 1$

Test: $\omega_1 = 0$

95% HPD interval: $-2.1467 < \omega_1 < .77284$

Result: $\omega_1 = 0$

Therefore, the identified model seems to be the appropriate one.

Forecasting

Using equation (3.48), for this example, the predictive distribution, conditional to θ_1, is such that

$$\frac{Y_{N+1} - 2.9624}{.48896}$$

has a t-distribution with 196 degrees of freedom. If a quadratic loss function is used, the point estimator of the future observation is equal to

$$\hat{Y}_{N+1} = 2.9624$$

and the 95% HPD interval corresponds to

$$2.004 < Y_{N+1} < 3.921.$$

The actual value of this observation in the generated series was equal to 2.4623.

In this example, as in the previous one, we checked the sensitivity of the results to changes in the value of the conditioning parameter and we did not observe appreciable changes. We also generated series using a larger variance for the errors, and we observed that as the variance increases, the intervention transfer function becomes closer to a constant of value 3. To check the sensitivity to changes in the number of observations, we used the first m elements, for m = 110, 120, 150, and 175, of the generated series. The identification procedure produced the correct model in all the cases, while poor estimations and forecastings were obtained fror m = 110 and m = 120. It seems that, at least in this example, for the estimation and forecasting procedures to be reasonably effective, one should keep observing the series until it reaches the steady state. Further work has to be done to test the performance of these procedures in cases where the transient stage of the intervention effect takes much longer, so that the number of observation before reaching the steady state is large. Notice that in the example presented, the series reaches its final level about 20 time units after the starting point of the intervention.

5.3 Comments

The examples presented before illustrate the use of the procedures proposed in this report and show that they produce reasonable results. Other examples were analized and the results obtained support the adequacy of the method.

Due to the complexity of the model, we had to use ad-hoc computer programs for each case but we hope to build in the future a package which could be used by the experimenter to analyze a wide variety of situations. One of the advantages is the fact that it is possible to include prior information when available, and also the fact that the procedure becomes very similar to the modeling process used in linear regression.

CHAPTER VI

SUMMARY

In this paper, a Bayesian procedure was proposed to analyze time series with interventions. The procedure relies on the use of an approximation to the likelihood function which eliminates some of the non-linearities involved.

Chapter III included the derivation of the posterior and predictive distributions for the case of a single intervention model, and the posterior distribution for a model with multiple linear interventions. In both cases, the noise was assumed to be generated by an ARMA(p,q) process. Normal-inverted gamma and "non-informative" prior distributions were considered.

Chapter IV described a new Bayesian approach to analyze these series through the conditional posterior distributions. Procedures were proposed to implement the different stages of the time series analysis, namely; identification, estimation, diagnostic checking, and forecasting. The discussion was restricted to the case in which there is only one intervention and the noise model does not include seasonal factors, but the procedure can be easily extended to include them.

The estimation procedure was applied in Chapter V, with some modifications, to a problem which includes several interventions and the noise contains a seasonal factor. The results obtained were very satisfactory. The complete time series analysis procedure was also used in this chapter to analyze an artificially generated data set with successful results.

Although further work has to be done from the theoretical point of view, the author believes from the results obtained that the procedures can be successfully applied in a wide variety of situations. An interesting case would be that of a regression model where the errors are generated by an ARMA process. A more detailed study has to be done on the adequacy of the approximation to the likelihood and this is more important in cases when the transient part of the intervention occurs in a short period of time. Also, since the procedure relies on the use of conditional posterior distributions, an analysis of the sensitivity of the results to changes in the conditioning values has to be done to assure the validity of the conclusions. When possible, the marginal posterior distributions should be used.

BIBLIOGRAPHY

Abraham, B. (1980). "Intervention analysis and multiple time series." Biometrika, 67, 70-78.

Box, G.E.P. and Jenkins, G.M. (1970). Time Series Analysis, Forecasting and Control, San Francisco: Holden Day, Inc.

Box, G.E.P. and Tiao, G.C. (1965). "A change in level of a non-stationary time series." Biometrika, 52, 181-192.

Box, G.E.P. and Tiao, G.C. (1975). "Intervention analysis with applications to economic and environmental problems." Journal of the American Statistical Association, 70, 70-79.

Box, G.E.P. and Tiao, G.C. (1976). "Comparison of forecast and actuality." Applied Statistics, 25, 195-200.

Broemeling, L.D. and Land, M. (1984). "On forecasting with univariate autoregressive processes: A Bayesian approach." Communications in Statistics, 13, 1305-1320.

Broemeling, L.D. and Shaarawy, S.M. (1984a). "A Bayesian analysis of Time Series." To appear in Bayesian Inference Decisions Techniques with Applications: Essays in Honor of Bruno de Finetti, edited by Goel and Zellner.

Broemeling, L.D. and Shaarawy, S.M. (1984b). "Bayesian Identification of Time Series." Unpublished manuscript.

Cook, P.J. (1983). "A Bayesian analysis of autoregressive processes: time and frequency domain." Ph.D. dissertation Oklahoma State University. Unpublished manuscript.

Glass, G.V., Willson, V.L. and Gottman, J.M. (1975). Design and Analysis of Time-Series Experiments, Boulder, Colorado: Colorado Associated University Press.

Lasarre, S. and Tau, S.H. (1982). "Evaluation of safety measures on the frequency and gravity of traffic accidents in France by means of intervention analysis." Time Series Analysis: Theory and Practice 1, edited by O.D. Anderson. Amsterdam: North Holland Publishing.

Liu, L.M. and Tiao, G.C. (1980). "Parameter estimation in dynamic models." Communications in Statistics, Series A, 9, 501-517.

McLeod, A.I. (1978). "On the distribution of residual autocorrelations in Box-Jenkins models." Journal of the Royal Statistical Society, Series B, 40, 296-302.

Newbold, P. (1973). "Bayesian estimation of Box-Jenkins transfer function-noise models." Journal of Royal Statistical Society, Series B, 35, 333-336.

Priestley, M.B. (1982). Spectral Analysis and Time Series. New York: Academic Press

Salazar, D. (1982). "Structural changes in time series models." Journal of Econometrics, 19, 147-163.

SAS User's Guide: Basics (1982). North Carolina: SAS Institute, Inc.

SAS/ETS User's Guide: Econometric and Time-Series Library (1982). North Carolina: SAS Institute, Inc.

Shahabudin, S. (1980). "Analysis of stock prices using time series with intervention analysis." American Statistical Association Proceedings on Economic Statistics Section, 311-314.

Shaarawy, S.M. (1984). "Bayesian analysis of Moving Average Processes." Ph.D. dissertation. Unpublished manuscript.

Shaarawy, S.M. and Broemeling, L.D. (1984a). "Bayesian inferences and forecasts with moving averages processes." Communications in Statistics, Series A, 13, 1871-1888.

Shaarawy, S.M. and Broemeling, L.D. (1984b). "Inference and prediction with ARMA processes." Unpublished manuscript.

Shaarawy, S.M. and Broemeling, L.D. (1985). "Bayesian identification of multiple ARMA models." Unpublished manuscript.

Smith, J.Q. (1979). "A generalization of the Bayesian steady forecasting model." Journal of the Royal Statistical Society, Series B, 41, 375-387.

Stoline, M.R. and Huitema, B.E. (1978). "Aspects of the intervention time series analysis of the pilot programs of Michigan public act 339. An act decriminalizing drunkenness." American Statistical Association Proceedings of Social Statistics, 397-402.

Spliid, H. (1983). "A fast estimation method for the vector autoregressive moving average model with exogenous variables." Journal of the American Statistical Association, 78, 843-849.

Tiao, G.C., Box, G.E.P. and Hamming, W.J. (1975). "Analysis of Los Angeles photochemical smog data: A statistical overview." _Journal of Air Pollution Control Association_, 25, 260-268.

Zellner, A. (1971). _An Introduction to Bayesian Inference in Econometrics_. New York: John Wiley.

APPENDIX

CONSIDERATIONS ON THE APPROXIMATE LIKELIHOOD

The purpose of this appendix is to present a preliminary analysis on the theoretical justification of the approximation to the likelihood function described by (3.11) and (3.15), for the case in which no interaction is present. This approximation was first proposed by Shaarawy (1984) to analyze a time series generated by a moving average process, but no mathematical proof has been given to guarantee its adequacy. In this regard, much more work is yet to be done.

The behavior of the estimated residuals, calculated through the maximum likelihood estimators of the parameters of the model, depends entirely on the large sample properties of those estimators. Even though, the author did not find in the literature a rigorous proof specially for the case of mixed (ARMA) models, if the estimators are assumed to be consistent and asymptotically normal, which is what most authors do, it is not difficult to show that the estimated residuals are asymptotically independent and normally distributed with mean zero and variance σ^2. From this it seems reasonable to use them in the construction of the approximate likelihood function. Nevertheless, it is not clear how to prove that the function obtained converges somehow, to the exact likelihood as the sample size goes to infinity. The numerical results obtained from the application of this approach to a large variety of examples show that this must be the case.

It is important to mention that, the approximate and exact likelihoods coincide when the values of the parameters are equal to the maximum likelihood estimators and that, as the sample size increases both function become more concentrated. Due to the consistency properties, the exact likelihood will be concentrated around the real values of the parameters and the approximate one appears to behave in the same way. From those considerations it is possible to conjucture that at least in a neighborhood of the maximum likelihood estimators, the values of the two functions become closer as the number of observations increases.

Also, it is important to notice that the approximate likelihood function corresponds to that of a linear regression model where some of the independent variables are lagged series of the estimated residuals, and the noise terms are assumed to be independent and normally distributed with mean zero and variance σ^2. The points at which the maximum of the approximated likelihood is attained, which coincides with the mean of the posterior distribution for non-informative priors, is therefore a "good" estimator of the point at which the exact likelihood attains its maximum.

Certainly, the previous analysis does not constitute a mathematical proof, but those statements are supported by the extensive numerical study done by Shaarawy to produce the papers mentioned in Chapter II of this work and also by this author before the procedure was applied to the intervention models.

VITA 2

Joaquin Diaz-Saiz

Candidate for the Degree of

Doctor of Philosophy

Thesis: TIME SERIES ANALYSIS WITH INTERVENTIONS: A BAYESIAN APPROACH

Major Field: Statistics

Biographical:

Personal Data: Born in Xalapa, Ver., Mexico, September 20, 1944, the son of Francisco Diaz Ponte and Alicia Saiz de Diaz.

Education: Received Bachelor of Science degree in Physical and mathematical Sciences from the Instituto Tecnologico y de Estudios Superiores de Monterrey, 1966; received Master of Science degree in Statistics from the Centro Interamericano de Ensenanza de Estadistica, Santiago, Chile, 1968; completed the requirements for the Doctor of Philosophy degree at Oklahoma State University in July, 1985.

Professional Experience: Assistant Professor, Department of Mathematics, Instituto Tecnologico y de Estudios Superiores de Monterrey, Mexico, February, 1969 to December, 1970; Teaching Assistant, University of Minnesota, January, 1971, to December 1971; Assistant Professor, University of the Americas, Mexico, February, 1972, to June 1976; Associate Professor, Universidad Nacional Autonoma de Mexico, June 1976, to July 1985; Research Associate, Oklahoma State University, August, 1983, to July, 1985.

Professional Organizations: Member of the American Statistical Association and the Statistical Society of Canada.