Graduate Studies
Mathematical Sciences

2

University of Twente
The Netherl

Time Series Models

Agus Sukmana

MSc Engineering Mathematics

Literature Study Report

November 1998

Time Series Models

Agus Sukmana

Contents

Chapter 1

Introduction

This literature study is a part of my study at University of Twente. The purpose of my literature study is to explore information in the area of time series model and some curve fitting, and also to learn how to search and to select the relevant literature. This literature report will be used as basic knowledge before I am doing my final project at *Centraal Bedieningssysteem Distributiepompstations(CBD) Gemeentewaterleidingen* Amsterdam .

The huge amount of measurement from 120 channel is available (more or less 3456000 measurement per day) at CBD, from this heap information we want to learn the way the measurements vary overtime and to construct the time patterns of past period without storing all the original measurements. My final project deal with storing the time patterns in an efficient way and controlling the whole process of the waterworks. Base on this fact, time series model and curve fitting are chosen as a topic of my literature study.

My report is organized in the following way . First part of my report deal with the literature search process that I do and its result. The second part deal with my literature survey on time series and curve fitting. In this part I try to give an overview on the time series models (in general), linear stationary models, model identification, parameter estimation, diagnostic checking, regression model with time series error, some smoothing and curve fitting methods.

Part I

Report on Literature Search

Chapter 2

Literature Search

2.1 Search Plan

In order to find references that are related to my topic, some aspect should be consider. This section deals with everything that we have to consider such as:

- where we have to start
- where to find
- what criteria are used
- how to select appropriate one.

2.1.1 Starting Point

When my supervisor formulated this literature study, he gave me a book [34] to study on subject of *univariate Box-Jenkins methods.* After studied this books, I could be more specific in searching for references which related to the topic time series model. I determine some keyword which related to time series after studied this book. By using these book as initial references, I started searching from various sources.

2.1.2 Important Sources

There are many important sources for searching literature in the field of mathematics. Some of them are listed below.

- On-line Public Catalogue of University Library Twente
- The Netherlands Central Catalogue
- On-line Contents
- WebCat

- Carl Uncover
- Math SciNet
- Science Citation Index

Since November 1998, Twente University Library introduces the new database, is called *Picarta*. Because of time constraint, I could not try this database.

There are also many possibilities to search literature from other sources such as :

- Alta vista (www.altavista.digital.com) and Yahoo (www.yahoo.com) search engine.
- References from the articles

2.1.3 Selection Criteria

I have to perform selection criteria for keeping that my searching is not become wide and still on the right track. The following keywords are used to find articles that related to my purpose:

- time series
- autoregressive
- moving average
- spline
- smoothing
- regression

In order to select which articles to be considered whether it is relevant to my literature study or not, I select some articles whose title relevant to my topic, after that I read their summaries to make sure that articles relevant. If I think that articles is relevant then I make a copy of it. The last step is classification, whether the articles is become my reference or not.

2.2 Search Process and Result

2.2.1 On-line Public Catalogue (OPC)

The On-line Public catalogue Twente University Library contains all books, journals and audiovisual media available in the collections of the library. This catalogue can be searched from University Library or accessed via internet (http:\\www.ub.utwente.nl). We can use term: title word, author name, corporation (publisher), author-name-conference, systematic code, key words (in

Dutch), ISBN (for books), ISSN (for periodicals) , year of publication or combination of those term.

Base on title word criteria, the following result were obtained by OPC:

Title word	Search Results	Bibliography
Time series	117	[5],[17],[18],[37], [38]
Autoregressive	3	-
Moving Average	0	-
Spline?	41	[7],[8]
Smoothing?	33	[8],
Regression?	1	[35],[37],[43]

Base on those result, 11 books are added to my bibliography. Next, I will search using author criteria that found in the previous search, because sometimes there are books that can not be covered by title criteria.

The result of searching base on the author criteria :

Author	Search Result	Author	Search Result
Pankratz, Andrew	2	Pesaran, M.	3
Eubank, Randall	2	Dierckx	1
Box, G.E.P.	5	Seber,G.	3
Harvey, Andrew	4	Pandit, S	1

Base on those result, [34] is added. This book related to my topic but did not find in the first searching by title criteria.

I also tried to find other books from other resources outside Twente University Library, such as NCC.

2.2.2 Netherlands Central Catalogue (NCC)

The Netherlands Central Catalogue NCC contains bibliographic references and the locations of approximately 12 million books and almost 500,000 periodicals in more than 400 libraries in The Netherlands. The database is updated directly and continually. The NCC database is connected to the Interlibrary Loan System IBL. So, it is possible to requests for photocopies of articles or requests for books from other library using IBL account.

The following result were generate by NCC:

Title word	Search Results
Time series	703
Autoregressive	87
Moving Average	29
Spline?	168
Smooth?	432
Regression?	795

Var = $\Sigma(x-\bar{x})^2$

standar dev = $\sqrt{var}$

semakin besar x dng $\bar{x}$ S akan besar.

From those result, there are no books added to my bibliography.

For searching the articles which related to my topic, *snowball method* and *science citation index* are easiest ones if we have found an article as starting point. If a starting article were not found, MathSciNet will be come more helpful than other resources that I mentioned in subsection (2.1.2), because MathSciNet provided summary of article.

2.2.3 MathSciNet

MathSciNet is the searchable Web database providing access to Mathematical Reviews and Current Mathematical Publications from 1940 to the present produced by the American Mathematical Society. Current Mathematical Publications is a subject index of bibliographic data for recent and forthcoming publications. Most items are later reviewed in Mathematical Reviews. All items in Mathematical Reviews appear first in Current Mathematical Publications.

Current Mathematical Publications data is added daily. Mathematical Reviews data is added each month when the printed issue is complete. The Mathematical Reviews record for an item with a review replaces the Current Mathematical Publications record for that item. The database is available on the internet site. The mirror sites are located in Bielefeld (Germany), Bonn (Germany), Strasbourg (France), Houston TX (USA), Providence RI (USA). This site *http://ams.mathematik.uni-bielefeld.de/* is the nearest mirror site to Enschede).

Because most of articles in MathSciNet have summary, I used this database to perform complete and specific searched. Most of articles in my bibliography was found by MathSciNet.

The following results were generated by MathSciNet:

Title word	Search Results
Time series*	2014
Time series* AND Identification*	63
Time series* AND Estimate*	348
Time series* AND Check*	11
Time series* AND Select*	27
Autoregressive*	1168
Autoregressive* AND Identification*	19
Autoregressive* AND Estimat*	413
Autoregressive* AND Check*	1
Autoregressive* AND Select*	25
Moving Average*	514
Moving Average* AND Identification*	16
Moving Average* AND Estimat*	135
Moving Average* AND check*	0
Moving Average* AND select*	0
Autocorrelat*	551
Autocorrelat* AND regression*	73
Spline*	5017
Smoothing* & Spline*	327
Spline* & Regression*	77
Dynamic* & Regression*	36

According to those result, some articles are added to my bibliography. [3], [4], [10], [14], [21], [24], [26], [28], [30], [31], [46] is relevant to time series model, [1], [2], [11], [13], [16], [19], [22], [32], [33], [36], [39], [41], [42], [47], [48], [49], [50] is relevant to regression with time series error and [6], [9], [23], [45] is related to spline smoothing.

2.2.4 On-line Contents (OLC)

On-line Contents OLC contains references to all articles that appear in over 12.500 current periodicals. The database contains mostly academic journals, but also general and non-specialist periodicals are included. These journals can be found in the collections of Dutch libraries. The database is built on the basis of the tables of contents of each separate issue. Since September 1992, the OLC updated daily and annually more than 2 million article references are added. The OLC-database is connected to the Interlibrary Loan System IBL in The Netherlands. So, it is possible to request a copy of articles from other Dutch libraries using IBL account. The OLC-database is connected to the Interlibrary Loan System IBL in The Netherlands. So, it is possible to request a copy of articles from other Dutch libraries using IBL account.

The following result were searched by OLC :

Title word	Search Results
Time series	2712
Autoregressive	645
Moving Average	189
Spline AND regression	24
Smoothing AND Spline	55
Regression AND smoothing	24

2.2.5 UnCover

UnCover is a database of current article information taken from over 17,000 multidisciplinary journals. UnCover contains brief descriptive information for over 7,000,000 articles which have appeared since Fall 1988. The database is available on the internet site *http://uncweb.carl.org* . We can choose search type: keyword, author or journal title and combine that with years of publication.

The following result were generated by UnCover :

Keyword	Search Results
Time series	3538
Autoregressive	732
Moving Average	249
Regression AND Spline	24
Smoothing AND Spline	53
Smoothing AND Regression	32

2.2.6 WebCAT

WebCAT is the central catalogue of the WebDOC-project, in which via Internet full text access is provided to a large collection electronic documents by libraries from The Netherlands, Germany and the US and international publishers. Several search keys can be used, including full word searching on the titles and the abstracts.

The following results were generated by WEbCAT:

Title word	Search Results
Time series*	12
Autoregressive*	1
Moving Average*	0
Spline*	9
Smoothing* AND Spline*	2

2.2.7 Science Citation Index (SCI)

The science citation index is provided by the Institute for Scientific Information. It indexes 5300 major journals across 164 disciplines and covering 2000 more journals. The information from the articles is sorted to the following register : Source Index, Permutation Index, Corporate Index, Citation Index. The citation index contains all cited articles in a specific period. With the SCI it is possible to search for articles which have specific articles as reference. In the TW library there is SCI in CD-ROM format which contains articles from January 1990 up to June 1998.

The following result were generate by SCI:

Title word	Search Results
Time series*	1274
Autoregressive*	353
Moving Average*	103
Spline*	537
Smoothing* AND Spline*	15

2.2.8 Snowball Methods

As I have mention before that *snowball method* and *science citation index* are easiest ones if we have found an article as starting point. I started from [48], then from the reference of that article, I got 8 relevant articles.

2.2.9 Search Engine

Sometimes from internet search engine we can find preprint or technical report. I found articles [12], [27], [29], [51], [52], [53], [54], [55], [56] by altavista internet search engine.

2.3 Selection Result

The list of selected books and articles can be found at the end of this report in reference.

Part II

Report on Literature Survey

Chapter 3

Linear Stationary Time Series Model

3.1 Introduction-Box Jenkins Model

we consider about time series data. *Time series data* refers to observations on a variable that occur in time series sequence. *Time series analysis* refers to any kind of analysis involving time series data, or it used to explain behavior of time series data using only past observations on the variable in question. I only discuss time series analysis with emphasis on univariate analysis and Box-Jenkins model.

3.1.1 Requirement for Box-Jenkins model

- **Short-term forecasting**

  The Box-Jenkins model are suited to short-term forecasting model and to forecasting of series containing seasonal variation and shifting seasonal pattern.

- **Data type**

  This model deals only with data measured at equally spaced and discrete time intervals.

- **Sample size**

  Construction of an adequate ARIMA model requires a minimum 50 sample size. A large sample size is desirable when seasonal variation is present [5].

- **Stationary series**

  This method applies only to stationary time series. A stationary time series has mean, variance and autocorrelation function that are essentially

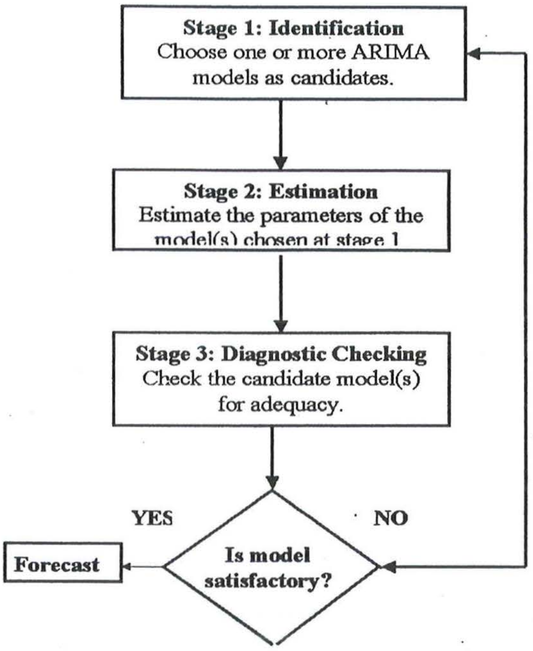
Stage 1: Identification
Choose one or more ARIMA models as candidates.
Stage 2: Estimation
Estimate the parameters of the model(s) chosen at stage 1
Stage 3: Diagnostic Checking
Check the candidate model(s) for adequacy.
YES
NO
Is model satisfactory?
Forecast

Figure 3.1: Box-Jenkins iterative approach

constant through time. Often, a nonstationary mean can be made stationary series with appropriate trasformations. The most common type of nonstationarity occurs when the mean of the realization change over time, for this type can frequently be rendered stationary by differencing.

3.1.2 Modelling procedure

Box-Jenkin in [34] propose a practical procedure for finding a good model. The procedure is summarized in figure 3.1.

Stage 1 : Identification

At this stage we tentatively select one or more ARIMA[1] models by looking at two graphs derived from the available data. This graph are called an estimated autocorrelation function (acf) and an estimated partial autocorrelation function (pacf). We choose a model whose associated theoretical acf and pacf look like the estimated acf and pacf calculated from data.

[1] ARIMA stands for Autoregressive Integrated Moving Average.

Stage 2 : Estimation

At this stage we obtain estimates of the parameters for the ARIMA model tentatively chosen at identification stage.

Stage 3 : Diagnostic checking

At the diagnostic-checking stage we perform tests to see if the estimated model is statistically adequate. If it is not satisfactory then we return to identification stage to select another models.

3.1.3 The Advantages of Box Jenkins Model

The Box-Jenkins approach has advantages compare to other traditional single-series [34].

- the concept associated with Box-Jenkins model are derived from a solid foundation of mathematical theory
- Box and Jenkins have developed a strategy for choosing one or more appropriate models out of larger family ARIMA models.
- An appropriate ARIMA model produces optimal univariate fore case because this model has smaller mean squares forecast error.

3.2 Linear Stationary Models

This section deals with definition and properties of stationary: autoregressive models, moving average model, and mixed autoregressive-moving average model.

3.2.1 Autoregressive Process

Auto Regressive Model (AR(p)).

Definition 1 *[25]An autoregressive process of order p is a process that satisfies a difference equation*

$$\tilde{Z}_t = \phi_1 \tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-2} + ... + \phi_p \tilde{Z}_{t-p} + a_t \tag{3.1}$$

where $\tilde{Z}_t = Z_t - \mu$ is the deviation of the process from some origin, $\phi_1, \phi_2, ..., \phi_p \in R$ and a_t is stationary white noise with mean μ and variance σ^2. (The a_t term in an ARIMA process usually assumed to be normally, identically and independently distributed random variables with a mean of zero and a constant variance).

The process is called autoregressive because the value of the process at time t beside on a pure random component depends on the p intermediate past values of the process itself.

Equation (3.1) can be written in backshift operator $(B\tilde{Z}_t = \tilde{Z}_{t-1})$ as

$$\begin{aligned} a_t &= \tilde{Z}_t - \phi_1\tilde{Z}_{t-1} - \phi_2\tilde{Z}_{t-2} - \ldots - \phi_p\tilde{Z}_{t-p} \\ &= (1 - \phi_1 B - \phi_2 B^2 - \ldots - \phi_p B^{t-p})\tilde{Z}_t \\ &= \phi(B)\tilde{Z}_t \qquad (3.2) \end{aligned}$$

and also in term of previous a's

$$\tilde{Z}_t = \phi^{-1}(B)a_t. \qquad (3.3)$$

Stationary Conditions

The Stationary requirement ensures that we can obtain useful estimates of the mean, variance and acf from sample. If the process mean were different each time period, we could not obtain useful estimates since we typically have only one observation available per time period.

The set of parameter $\phi_1, \phi_2, \ldots, \phi_p$ of AR(p) process (3.1) must satisfy certain conditions for the process to be stationary. These conditions are summarized in following table.

Summary of stationary conditions for AR coefficients

Model Type	Stationary Condition
AR(1)	$\lvert\phi_1\rvert < 1$
AR(2)	$\lvert\phi_2\rvert < 1,\ \phi_1 + \phi_2 < 1,\ \phi_2 - \phi_1 < 1$

For illustration, the AR(1) process

$$(1 - \phi_1 B)\tilde{Z}_t = a_t$$

may be written

$$\begin{aligned} \tilde{Z}_t &= (1 - \phi_1 B)^{-1}a_t \\ &= \sum_{j=0}^{\infty}\phi_1^j B^j a_t \qquad (3.4) \\ &= \sum_{j=0}^{\infty}\psi(B)a_t \qquad (3.5) \end{aligned}$$

where

$$\psi(B) = (1 - \phi_1 B)^{-1} = \sum_{j=0}^{\infty}\phi_1^j B^j \qquad (3.6)$$

If $|\phi_1| < 1$ then $(1 - \phi_1 B)^{-1}$ is equivalent to convergent geometric series. So, $|\phi_1| < 1$ is a condition for the process to be stationary .

The stationary condition become complicated when $p > 2$. When $p > 2$ we can at least check the necessary (but not sufficient) stationarity condition:

$$\phi_1 + \phi_2 + ... + \phi_p < 1$$

Theoretical acf and pacf for AR process

The theoretical autocorrelation function and partial autocorrelation function for AR(p) process will be discussed in this subsection. The idea in autocorrelation analysis is to calculate a correlation coefficient for each set of ordered pair $(z_t z_{t+k})$. Because we are finding the correlation between sets of numbers that are part of the same series, the resulting statistic is called an autocorrelation coefficient.

Consider the AR(p) process (3.1) and multiply by $\tilde{Z}_{t-k}$, to obtain

$$\tilde{Z}_{t-k}\tilde{Z}_t = \phi_1 \tilde{Z}_{t-k}\tilde{Z}_{t-1} + \phi_2 \tilde{Z}_{t-k}\tilde{Z}_{t-2} + ... + \phi_p \tilde{Z}_{t-k}\tilde{Z}_{t-p} + \tilde{Z}_{t-k} a_t \qquad (3.7)$$

on taking expected values in (3.7), we obtain autocovariances function

$$\begin{aligned} \gamma_k &= \phi_1 \gamma_{k-1} + \phi_2 \gamma_{k-2} + ... + \phi_p \gamma_{k-p} \ , \ k > 0 \qquad (3.8) \\ \gamma_0 &= \phi_1 \gamma_{-1} + \phi_2 \gamma_{-2} + ... + \phi_p \gamma_{-p} + \sigma_a^2 \qquad (3.9) \end{aligned}$$

on dividing (3.8) by γ_0, we obtain autocorrelation function

$$\rho_k = \phi_1 \rho_{k-1} + \phi_2 \rho_{k-1} + ... + \phi_p \rho_{k-1} \cdot \ , \ k > 0 \qquad (3.10)$$

and by dividing(3.8) by $\gamma_0 = \sigma_z^2$ and substituting $\gamma_k = \gamma_{-k}$, we obtain variance in following form

$$\sigma_z^2 = \frac{\sigma_a^2}{1 - \rho_1 \phi_1 - ... - \rho_p \phi_p} \qquad (3.11)$$

For example AR(1) process :

$$\begin{aligned} \rho_k &= \phi_1 \rho_{k-1} \ , k > 0 \\ \rho_0 &= 1 \end{aligned}$$

or

$$\rho_k = \phi_1^k \ , k \geq 0$$

the autocorrelation function decays exponentially to zero when ϕ_1 is positive, but decays exponentially to zero and oscillates in sign when ϕ_1 negative.

The partial autocorrelation function (pacf) is a tool which exploit the fact that whereas an AR(p) process has an autocorrelation function which is infinite in extend. Denote that ϕ_{kj} is the jth coefficient in an autoregressive process of order k. From (3.10) , the ϕ_{kj} satisfy the set equations

$$\rho_j = \phi_{k1}\rho_{j-1} + ... + \phi_{k(k-1)}\rho_{j-k+1} + \phi_{kk}\rho_{j-k} \text{ , } j = 1, 2, ..., k \tag{3.12}$$

or write it in Yule-Walker equation [5]

$$\mathbf{P}_k\boldsymbol{\phi}_k = \boldsymbol{\rho}_k \tag{3.13}$$

where

$$\mathbf{P}_k = \begin{bmatrix} 1 & \rho_1 & \cdots & \rho_{k-1} \\ \rho_1 & 1 & \cdots & \rho_{k-2} \\ \cdots & \cdots & \cdots & \cdots \\ \rho_{k-1} & \rho_{k-2} & \cdots & 1 \end{bmatrix} \qquad \phi_k = \begin{bmatrix} \phi_{k1} \\ \phi_{k2} \\ \cdots \\ \phi_{kk} \end{bmatrix} \qquad \rho_k = \begin{bmatrix} \rho_1 \\ \rho_2 \\ \cdots \\ \rho_k \end{bmatrix}.$$

Estimated acf and pacf for AR process

Estimated acf [34]:

$$r_k = \frac{\sum\limits_{t=1}^{n-k} \tilde{z}_t\tilde{z}_{t+k}}{\sum\limits_{t=1}^{n} (\tilde{z}_t)^2} \tag{3.14}$$

[5] suggest that the maximum number (k) of useful estimated autocorrelation is roughly $n/4$, where n is number of observation.

Estimated pacf ([34],p.40;[5],p.82-84):

$$\begin{aligned} \hat{\phi}_{11} &= r_1 \\ \hat{\phi}_{kk} &= \frac{r_k - \sum\limits_{j=1}^{k-1} \hat{\phi}_{k-1,j}r_{k-j}}{1 - \sum\limits_{j=1}^{k-1} \hat{\phi}_{k-1,j}r_j} \quad , \quad k = 2, 3, ... \end{aligned} \tag{3.15}$$

where

$$\hat{\phi}_{kj} = \hat{\phi}_{k-1,j} - \hat{\phi}_{kk}\hat{\phi}_{k-1,k-j} \text{ , } k = 3, 4, ... \text{ , } j = 1, 2, ..., k-1. \tag{3.16}$$

That recursive equations gives fairly good estimates of the pacf as long as a stationary series.

3.2.2 Moving Average Process

Moving Average Model (MA(q)).

Definition 2 *[25] A moving average of order q is a process Z_t that may be describes by the equation*

$$\begin{aligned} Z_t &= a_t - \theta_1 a_{t-1} - ... - \theta_q a_{t-q} \qquad (3.17) \\ t &= q, q+1, ... \end{aligned}$$

The coefficient $\theta_0, \theta_1, ..., \theta_k \in \mathbf{R}$ and a_t is white noise with mean μ and standard deviation σ. The value Z_t of the process at time t is weighted sum of the k+1 immediately preceding value of the white process a_t .

Invertibility Condition

The Invertibility requirement ensures that larger weight should be attached to more recent observations. If a model is non invertibility than the weight placed on the past observations do not decline as we move further to the past.

The set of parameter $\theta_1, \theta_2, ..., \theta_p$ of MA(q) process (3.17) must satisfy certain conditions for the process to be invertibility These conditions are summarized in following table.

Summary of invertibility conditions for MA coefficients

Model Type	Invertibility Condition
MA(0)	Always invertible
MA(1)	$\lvert\theta_1\rvert < 1$
MA(2)	$\lvert\theta_2\rvert < 1$, $\theta_1 + \theta_2 < 1$, $\theta_2 - \theta_1 < 1$

The invertibility condition become complicated when $q > 2$. When$q > 2$, we can at least check the necessary condition for invertibility

$$\theta_1 + \theta_2 + ... + \theta_q < 1.$$

Theoretical autocorrelation function (acf)

The theoretical autocorrelation function for MA(q) process will be discussed in this subsection.

Consider the MA(q) process .

$$\begin{aligned} \tilde{Z}_t &= \left(1 - \theta_1 B - ... - \theta_q B^q\right) a_t \\ &= \theta\left(B\right) a_t \qquad (3.18) \end{aligned}$$

autocovariance function

$$\begin{aligned} \gamma_k &= E\left(\tilde{z}_t \tilde{z}_{t-k}\right) \\ &= \begin{cases} \left(1 + \theta_1^2 + ... + \theta_q^2\right) \sigma_a^2 & k = 0 \\ \left(-\theta_k + \theta_1 \theta_{k+1} + ... + \theta_{q-k}\theta_q\right) \sigma_a^2 & k = 1, 2, ..., q \\ 0 & k > q \end{cases} \qquad (3.19) \end{aligned}$$

and the autocorrelation function

$$\rho_k = \begin{cases} \frac{-\theta_k+\theta_1\theta_{k+1}+...+\theta_{q-k}\theta_q}{1+\theta_1^2+...+\theta_q^2} & k=1,2,...,q \\ 0 & k>q \end{cases} \tag{3.20}$$

The autocorrelation function of moving average process has cut-off at lag q.

Example, MA(1) process.

Variance

$$\sigma_z^2 = \gamma_0 = \left(1+\theta_1^2\right)\sigma_a^2. \tag{3.21}$$

and autocorrelation function

$$\rho_k = \begin{cases} 1 & k=0 \\ \frac{-\theta_1}{1+\theta_1^2} & k=1 \\ 0 & k>1 \end{cases} \tag{3.22}$$

The autocorrelation at lag zero is always 1, at lag 1 is non zero and all other autocorrelation are zero.

In *a stationary* AR(p), a_t can be represented as a *finite* weighted sum of previous $\tilde{z}_t$'s or $\tilde{z}_t$ as an *infinite* weighted sum of previous a_t's. In *a invertable* MA(q) , $\tilde{z}_t$ can be represented as a *finite* weighted sum of previous a_t's or a_t can be represented as a *infinite* weighted sum of previous $\tilde{z}_t$'s

3.2.3 ARMA Processes

Autoregressive-Moving Average models (ARMA (p,q)).

Definition 3 *[25]The process Z_t is a mixed auto-regressive order p and moving average order q if it satisfies the difference equation*

$$Z_t = \phi_1 Z_{t-1} + \phi_2 Z_{t-2} + ... + \phi_p Z_{t-p} + \theta_t + \theta_1 a_{t-1} + ... + \theta_p a_{t-p} \tag{3.23}$$

for$t = p, p+1, ...$. We denote this process as an ARMA(p,q) process.

Stationarity and Invertibility Condition

Consider the ARMA (p,q) model

$$\phi(\mathbf{B})\,\tilde{z}_t = \theta(\mathbf{B})\,a_t. \tag{3.24}$$

Stationary condition for (3.24) is the roots of characteristic equation $\phi(\mathbf{B}) = 0$

must lies outside unit circle and invertible condition is the roots of $\theta(\mathbf{B})$ lies outside unit circle.

3.3 Model Identification

Identification is a critical stage in model building, and a knowledge of the theoretical acf's and pacf's is required for effective identification. Knowing the association between the common theoretical acf's and pacf's and their corresponding process does not guarantee that we will identify the best model, but by familiarity with the acf's and pacf's improve our changes of finding good model quickly.

At this stage we compare the estimated acf and pacf with various theoretical acf's and pacf to find a match. A model whose theoretical acf and pacf best match the estimated acf and pacf is chosen as a tentative model. The characteristics of theoretical acf and pacf for some common stationary process is listed at following table :

Characteristics of common stationary processes

Process	**acf**	**pacf**
AR(1)	Exponential decay: (i) on the positive side if $\phi_1 > 0$; (ii) alternating in sign starting on the negative side if $\phi_1 < 0$.	spike at lag 1, then cuts off to zero;(i) spike is positive if $\phi_1 > 0; (ii)$ spike is negatives if $\phi_2 < 0$.
AR(2)	A mixture of exponential decays or damped sine wave. The exact pattern depends on the signs and sizes of ϕ_1 and ϕ_2	Spike at lags 1 and 2, then cuts off to zero.
MA(1)	Spike at lag 1, then cut off to zero: (i) spike is positive if $\theta_1 < 0; (ii)$ spike is negative if $\theta_1 > 0$	Damps out exponentially:(i) alternating in sign, starting on the positive side, if $\theta_1 < 0$; (ii) on the negative side, if $\theta_1 > 0$
MA(2)	Spike at lags 1 and 2, then cuts off to zero	A mixture of exponential decays or a damped sine wave. The ecaxt pattern depends on the signs and sizes of θ_1 and θ_2.
ARMA(1,1)	Exponential decay from lag 1: (i) sign of $\rho_1 =$ sign of $(\phi_1 - \theta_1)$; (ii) all one sign if $\phi_1 > 0$; (iii) alternating in sign if $\phi_1 < 0$	Exponential decay from lag1: (i) $\phi_{11} = \rho_1$; (ii) all one sign if $\theta_1 > 0$; (iii) alternating in sign if $\theta_1 < 0$.

3.4 Estimation

From the previous stage we obtain some rough estimates of many autocorrelation and partial auto correlation as a guide to find an appropriate model. In the estimation stage we use available data for estimating parameter in efficient way.

This chapter deal with parameters estimation of Autoregressive, Moving average models .

3.4.1 Autoregressive

Substitute $\tilde{Z}_t = Z_t - \mu$ into(3.1) than we get

$$Z_t = \mu + \phi_1^2 Z_{t-1} + \phi_2^2 Z_{t-2} + ... + \phi_p^2 Z_{t-p} + a_t \tag{3.25}$$

$t = p, p+1, ...$, where a_t is white noise with mean 0 and variance σ^2. Now, the problem is : how to estimate the parameter μ, ϕ_i and σ , given N observations.

We will discuss two methods for solving that problem, viz least square and maximum likelihood method.

Least Square (LS) Estimator

By looking Z_t as a linear regression of the p previous value $Z_{t-1}, Z_{t-2}, ..., Z_{t-p}$, we applied LS method to this problem.

Recall equation (3.1), and write it in the form:

$$a_t = Z_t - \mu - \phi_1 Z_{t-1} - ... - \phi_p Z_{t-p}. \tag{3.26}$$

we look for the value of parameter μ, ϕ_i and σ for which $\sum a_t^2$ minimum. From [25] , we have normal equations

$$\hat{\mu} + \sum_{j=1}^{p} \hat{\phi}_j \bar{z}_j = \frac{1}{N-p} \sum_{t=n}^{N-1} z_t \tag{3.27}$$

$$\bar{z}_i \hat{\mu} + \sum_{j=1}^{p} \hat{\phi}_j c(i,j) = c(i,0) \tag{3.28}$$

where

$$\begin{aligned} z_i &= \frac{1}{N-p} \sum_{t=p}^{N-1} z_{t-i} \\ i &= 1, 2, ..., p \end{aligned} \tag{3.29}$$

and

$$\begin{aligned} c(i,j) &= \frac{1}{N-p} \sum_{t=p}^{N-1} z_{t-j} z_{t-1}. \\ i &= 1, 2, ..., p \\ j &= 0, 1, ..., p \end{aligned} \tag{3.30}$$

These normal equations yield $p+1$ estimator of μ, ϕ_1 , ... , ϕ_p. The variance σ^2 can be estimated by

$$\hat{\sigma}^2 = \frac{1}{N-p} \sum_{t=p}^{N-1} \left(z_t - \hat{\mu} - \hat{\phi}_i z_{t-1} - \ldots - \hat{\phi}_n z_{t-p} \right)^2 . \tag{3.31}$$

Maximum Likelihood (ML) Estimation

The assumption which are needed in maximum likelihood approach are :

a. the joint probability density function of the initial conditions $z_0, z_1, ..., z_{p-1}$ is known ,and

b. the probability density function of white noise also is known.(in this discussion only normally distributed will be considered).

We look for the value of parameter μ, ϕ_1 , ... , ϕ_p which are maximize

$$\sum_{t=p}^{N-p} \left(z_t - \mu - \sum_{j=1}^{p} \phi_j z_{t-j} \right)^2 . \tag{3.32}$$

The variance σ^2 can be estimated by

$$\hat{\sigma}^2 = \frac{1}{N-p} \sum_{t=p}^{N-1} \left(z_t - \hat{\mu} - \sum_{j=1}^{p} \hat{\phi}_j z_{t-j} \right)^2 . \tag{3.33}$$

The LS and ML estimation for μ and ϕ_j are identical if the white noise are normally distributed (as we suppose they are) [34].

[5] prefer to use the ML method than LS, because under assumption that the model is correct, the estimate derived from ML criterion reflect all useful information about the parameter contained in the data, but the computation for finding the exact ML estimate is rather difficult, except the white noise are normally distributed ([5],[18],[34]).

The LS and ML are asymptotically unbiased, consistent and asymptotically efficient under assumption the process is normally distributed [25].

3.4.2 Moving Average

Beside maximum likelihood method, [25] propose nonlinear least square method for estimating parameter of moving average model. The most commonly used nonlinear least squares method is marquardt's compromise, a combination of Gauss-Newton linearization and the gradient method. This method converges quickly to least squares value in most cases [34].

3.5 Diagnostic Checking

At this stage we determine whether a model is statistically adequate or not, in particular for independent white noises. Our goal is to build a model that completely explains any autocorrelation in the original series. If the assumption is not satisfies, there is an autocorrelation in the original series that has not been explained by the time series model. At this stage we use the residuals to test hypothesis about the independent of white noise.

The basic analytic tool for testing the hypothesis is the residual acf

$$r_k\left(a_t\right)=\frac{\sum_{t=1}^{n-k}\left(\hat{a}_t-\bar{a}\right)\left(\hat{a}_{t+k}-\bar{a}\right)}{\sum_{t=1}^{n}\left(\hat{a}_t-\bar{a}\right)^2}$$

The hypothesis $H_0 : \rho_k\left(a\right) = 0$ for each residual autocorrelation will be tested. [34] propose t-test for testing that hypothesis. Hypothesis will be rejected if the absolute value of approximate t-value

$$t=\frac{r_k\left(\hat{a}\right)}{s\left[r_k\left(\hat{a}\right)\right]}$$

where

$$s\left[r_k\left(\hat{a}\right)\right]=\left(1+2\sum_{j=1}^{k-1}r_j\left(\hat{a}\right)^2\right)^{1/2}. \quad n^{-1/2}$$

is larger than 1.25 at lag 1,2 and 3 and larger than 1.6 at larger lag. Other method for testing independency of white noise are chi-square test [26], residual plot, overfitting etc.[34].

Chapter 4

Regression Model with Time Series Error

4.1 Introduction

This chapter deals with regression model which errors described by time series model. [1] gave a review of earlier work on regression analysis when autocorrelation exist. [13] proved consistency properties of weighted least squares estimates of parameter β when $f(x_t, \beta)$ is nonlinear and e_t has a continuous spectrum. [39] consider the least squares estimation when the model is linear and errors follow an autoregressive moving average model. [11] provided a procedure for estimating the unknown parameter β in nonlinear regression settings. [16] consider maximum likelihood estimation under same conditions as those of [39], they used Kalman filter techniques. The same work has been done by [42] under the condition the root lies on the unit circle. [48] consider the regression model with time series errors but allowed the time process to be nonstationary.

4.2 Model

Consider the regression model

$$y_t = \mathbf{f}(x_t, \beta) + e_t \tag{4.1}$$

where the error e_t are not independent.

To represent the correlational structure of the error, we assume that they form a stationary time series, so that their mean, assumed zero, and their intercorrelation do not change over time.

4.3 Nonlinear Regression with Autocorrelated Errors

This section discuss estimation of the unknown parameter β of the nonlinear time series regression model $\mathbf{f}(x_t, \beta)$ based on [11]. The $\{e_t\}$ is assumed to be a covariance stationary time series. This means that the covariance $cov(e_t, e_{t+k})$ of time series depend only on the gap k and not on the position t in time.

One would estimate β^* by the value of β that minimizes

$$[\mathbf{y} - \mathbf{f}(\beta)]' \Gamma_n^{-1} [\mathbf{y} - \mathbf{f}(\beta)] \tag{4.2}$$

where Γ_n (were known) is the variance-covariance matrix of the disturbance vector $\mathbf{e}$. When Γ_n is not known, the obvious approach is to substitute an estimator of Γ_n in formula (4.2).

Assume that disturbance $\mathbf{e}$ follow autoregressive process of order p.

Estimation Procedure

1. to compute the ordinary least square estimator $\hat{\theta}$ which minimizes

$$[\mathbf{y} - \mathbf{f}(\beta)]' [\mathbf{y} - \mathbf{f}(\beta)]$$

using such as modified Gauss-Newton Method or Marquardt's algorithm.

2. to compute the residual $\hat{\mathbf{e}} = \mathbf{y} - \mathbf{f}(\boldsymbol{\beta})$ and estimate the autocovariances up to lag p of the disturbance using

$$\begin{aligned} \gamma(h) &= \frac{1}{n} \sum_{t=1}^{n-|h|} u_t u_{t+|h|} \\ h &= 0, 1, ..., p \end{aligned} \tag{4.3}$$

3. Let

$$\hat{\Gamma}_p = \begin{bmatrix} \hat{\gamma}(0) & \hat{\gamma}(1) & ... & \hat{\gamma}(p-1) \\ \hat{\gamma}(1) & \hat{\gamma}(0) & ... & \hat{\gamma}(p-2) \\ ... & ... & ... & ... \\ \hat{\gamma}(p-1) & \hat{\gamma}(p-2) & ... & \hat{\gamma}(0) \end{bmatrix} \tag{4.4}$$

and

$$\hat{\gamma}_p = \begin{bmatrix} \hat{\gamma}(1) & \hat{\gamma}(2) & ... & \hat{\gamma}(p) \end{bmatrix}'. \tag{4.5}$$

compute $\hat{\phi}$ using the Yule-Walker equations

$$\hat{\phi} = -\hat{\Gamma}_p \hat{\gamma}_p$$

and

$$\hat{\sigma}^2 = \hat{\gamma}(0) + \hat{\phi}' \hat{\gamma}_p.$$

Factor $\hat{\Gamma}_p^{-1} = \hat{P}_p'\hat{P}_p$ and set matrix $\hat{\mathbf{P}}$ (see [11] p.963).

4. Define

$$\nabla' f(x,\theta) = \left[\begin{array}{ccc} \frac{\partial}{\partial\theta_1} f(x,\theta) & \ldots & \frac{\partial}{\partial\theta_p} f(x,\theta) \end{array} \right]$$

and $F(\theta)$ = the n by p matrix whose the row of $\nabla' f(x,\theta)$.
Compute $\tilde{\theta}$ by minimizing

$$Q_n(\theta) = \frac{1}{n}\left[\hat{P}y - \hat{P}f(\theta)\right]'\left[\hat{P}y - \hat{P}f(\theta)\right] \tag{4.6}$$

and from this value obtain

$$\tilde{\sigma}^2 = \left[\hat{P}y - \hat{P}f(\tilde{\theta})\right]'\left[\hat{P}y - \hat{P}f(\tilde{\theta})\right] / (n-q)$$

and

$$\tilde{C} = \left[F'(\tilde{\theta})\hat{P}'\hat{P}F\left(\tilde{\theta}\right)\right]^{-1}.$$

$\sqrt{n\left(\ \tilde{\theta} - \theta^*\right)}$ is asymptotically normal distributed with a variance covariance matrix for which is strongly consistent estimator under appropriate regularity conditions. Marquard algorithm used for solving (4.6). [11] reported that ordinary least squares estimation $\hat{\theta}$ is a good start value for computing $\tilde{\theta}$. The estimation procedure may be iterated by returning the second step with $\tilde{\theta}$ replacing $\hat{\theta}$.

4.4 Regression Models with ARMA Errors.

A regression model which the errors follow a stationary autoregressive moving average will be discussed in this section based on ([39], [40],[16]). Maximum likelihood estimation and simultaneous least squares estimation in regression and the linear time series parameter is discussed.

Consider regression model (4.1) with

$$f(x_t,\beta) = \sum_{i=1}^{m} \beta_i x_{it} \tag{4.7}$$

and the errors e_t are assumed to follow a stationary mixed autoregressive-moving average process.

The ARMA(p,q) model for errors :

$$e_t = \sum_{j=1}^{p} \phi_j e_{t-j} - \sum_{k=1}^{q} \theta_k a_{t-k} + a_t \tag{4.8}$$

where is $\{a_t\}$ a set of independent random deviates with zero mean and variance σ^2. Define the backshift operator B by $Bw_t = w_{t-1}$ for any sequence $\{w_t\}$. If $\phi(B) = 1 - \sum \phi_j B^j$ and $\theta(B) = 1 - \sum \theta_k B^k$ are polynomials in B of degree p and q, then (4.8) may be written as

$$e_t = \phi^{-1}(B)\theta(B)a_t = \frac{\theta(B)}{\phi(B)}a_t \tag{4.9}$$

and (4.1) as

$$y_t = \sum_{i=1}^{m} \beta_i x_{it} + \frac{\theta(B)}{\phi(B)}a_t. \tag{4.10}$$

The assumptions for the model are :

1. the a_t are independent and identically distributed with zero mean, variance σ^2 and finite kurtosis γ_2 .
2. the root of polynomial $\phi(z) = 0$ and $\theta(z) = 0$ lies outside the unit circle, with no single root common to both polynomial
3. the constants x_{it} are bounded, for fixed $i,\ j,\ k\ and\ l$

$$\lim_{n\to\infty} \frac{1}{n} \sum x_{i,t-k} x_{j,t-l} \tag{4.11}$$

exists, and mxm matrix $\{\lim(1/n) \sum x_{it}x_{jt}\}$ is positive definite.

4.4.1 Least Squares Estimation

The analysis of this section refers to [39]. It is an extension of conditional sum of square approach employed by [5].

Suppose that $\{x_{it}; 1 \le i \le m\}$ and $\{y_t\}$ are series generated by (4.10). If we knew the true parameter value $\boldsymbol{\lambda} = (\beta, \phi, \theta)$ then the random deviates $\{a_t\}$ could be determined from the relation

$$a_t = \phi(B)\theta^{-1}(B)y_t - \sum \beta_i \left\{\phi(B)\theta^{-1}(B)x_{it}\right\} \tag{4.12}$$

obtain by solving (4.10) for a_t.

The true parameters are unknown, but for any vector of values $\dot{\boldsymbol{\lambda}} = \left(\dot\beta_1, ..., \dot\beta_m, \dot\phi_1, ..., \dot\phi_p, \dot\theta_1, ..., \dot\theta_q\right)'$ such that $\dot\phi$ and $\dot\theta$ satisfy 2^{nd} assumption. We define

$$\dot a_t = \dot\phi(B)\dot\theta^{-1}(B)y_t - \sum \dot\beta_i \left\{\dot\phi(B)\dot\theta^{-1}(B)x_{it}\right\} \tag{4.13}$$

or (by setting $\phi_0 = -1$),

$$\dot{a}_t = \sum_{k=1}^{q} \dot{\theta}_k \dot{a}_{t-k} - \sum_{j=0}^{p} \dot{\phi}_j y_{t-j} + \sum_{i=1}^{m} \sum_{j=0}^{p} \dot{\beta}_i \dot{\phi}_j x_{i,t-j}. \tag{4.14}$$

Thus, with the errors $\dot{a}_t$ as in (4.14) and with appropriate starting point, the least square estimate of $\boldsymbol{\lambda} = (\boldsymbol{\beta}, \boldsymbol{\phi}, \boldsymbol{\theta})$ are those value of $\hat{\boldsymbol{\lambda}}$ which minimize the sum of square

$$S(\dot{\lambda}) = \sum \dot{a}_t^2$$

as a function of $\dot{\boldsymbol{\lambda}}$ .Since $\dot{a}_t$ is not a linear function of $\dot{\boldsymbol{\lambda}}$, these estimates can be computed in practice by nonlinear estimation methods. The discussion of time series case is provided by [5].

If the $\{a_t\}$ were normally distributed , the log likelihood function of the parameter $\left(\dot{\lambda}, \dot{\sigma}^2\right)$ would be

$$\log \dot{L} = cons\tan t - \frac{1}{2} n \log \dot{\sigma}^2 - \frac{1}{2\dot{\sigma}^2} \sum \dot{a}_t^2 \tag{4.15}$$

and the maximization of (4.15) yields the same estimates of $\boldsymbol{\lambda}$ as the least squares procedure. An estimate of error variance

$$\hat{\sigma}^2 = \frac{1}{n} \sum \hat{a}_t^2 \tag{4.16}$$

where the $\hat{a}_t$ are the residual obtained by replacing $\boldsymbol{\lambda}$ in (4.13) of (4.14) by the estimates $\hat{\lambda}$.

Distribution of Estimates

The regression coefficient $\hat{\beta}$ possess a multivariate normal distribution with mean β and covariance matrix $\frac{\sigma^2}{n}\mathbf{B}^{-1}$, where

$$\mathbf{B} = \left\{ \lim_{n\to\infty} \frac{1}{n} \sum_{t=1}^{n} b_{it} b_{jt} \right\} \tag{4.17}$$

with the $b_{it}(i = 1, ..., m)$ define by relation

$$\phi(B) x_{ij} = \theta(B) b_{it} \tag{4.18}$$

The estimates $\hat{\eta} = (\hat{\phi}, \hat{\theta})$ are independent of $\hat{\beta}$ and are normally distributed with mean $\eta = (\phi, \theta)$ and covariance matrix

$$\frac{\sigma^2}{n}\begin{bmatrix} \mathbf{C} & \mathbf{E} \\ \mathbf{E}' & \mathbf{D} \end{bmatrix}^{-1} \tag{4.19}$$

where

$$\begin{aligned} \mathbf{C} &= \left\{\gamma_{|i-j|}\right\} \\ \mathbf{D} &= \left\{\delta_{|i-j|}\right\} \\ \mathbf{E} &= \left\{\omega_{i-j}\right\} \end{aligned}$$

$\gamma_k = E(u_t u_{t+k})$ and $\delta_k = E(v_t v_{t+k})$ are the lag k autocovariances of the autoregressive processes $\phi(B)u_t = a_t$ and $\theta(B)v_t = a_t$. $\omega_k = E(u_t v_{t+k})$ is the lag k cross covariance between these two processes.

The estimate $\hat{\sigma}^2$ is normal with mean σ^2 and variance $\frac{1}{n}\sigma^4(1 + \frac{1}{2}\gamma_2)$, independent of $(\hat{\phi}, \hat{\theta}\ )$ and independent of $\hat{\beta}$.

[39] has shown that :

- the least squares estimates of the regression parameter $\boldsymbol{\beta}$ and the parameter ϕ and θ is a quite good approximation to the first and second moment of the exact sampling distributions if the sample size is moderate.
- the estimates $\hat{\phi}$ and $\hat{\theta}$ are asymptotically uncorrelated with the regression estimates $\hat{\beta}$.
- the Monte Carlo investigation has indicate that the asymptotic result obtained can be misleading, if applied when only few observations are available.

Checking Adequacy of Fit

At previous section the estimation of the regression parameter β and the error parameters ϕ , θ was discussed. The validity of any methods concerning parameter estimation is predicted on the appropriateness of the assumed form of model to begin with, thus an important problem is that of examining model adequacy and test of fit. For an adequately fitting model the residuals $\{a_t\}$ should resemble the random deviates $\{a_t\}$ (nearly uncorrelated), so the large value of the residual autocorrelation would place the model under suspicion. That problem is further discussed in [40].

4.4.2 Maximum Likelihood Estimation

In this subsection I would like to discuss maximum likelihood estimator for regression models with correlated disturbances. The discussion of this problem is based on [16], which use Kalman filter. A number of author, for example [37], have recently stressed the desirability of computing estimator of ARMA parameter using exact likelihood function. [16] were demonstrated that this approach had computational, as well as theoretical, advantages over other methods and showed how the Kalman filter could be used to calculate exact maximum likelihood estimator of ARMA time series models.

Consider model (4.1), with the disturbance term is assumed to be generated by an autoregressive-moving average (p,q) process. The likelihood function for this model is

$$\begin{aligned} \log L\left(y; \phi, \theta, \beta, \sigma^2\right) &= -\frac{1}{2}n\log\left(2\pi\right) - \frac{1}{2}n\log\sigma^2 - \frac{1}{2}\log\left|V\right| \qquad (4.20) \\ & -\frac{1}{2}\sigma^{-2}\left(y - x\beta\right)' \mathbf{V}^{-1}\left(y - x\beta\right) \end{aligned}$$

matrix V is define by $E(ee') = \sigma^2 V$. The assumption needed to ensure the validity of kalman filter formulation of the generalized least squares estimator of β is that the autoregressive moving average process generating the $e's$ be stationary.

[16] has shown the advantages of Kalman approach such as:

- predictions of futures value of the dependent variable may be mad very easily when Kalman filter adopted.
- the recursions produce a set$n - k$ prediction errors which normally and independently distributed with zero mean and constant variance when the model is correctly specified and ϕ, θ must usually be estimated.

An alternative maximum likelihood procedure which incorporates the first observation and the stationary condition of the error process is proposed in [2]. This estimator is superior to conventional ones on theoretical grounds, and sampling experiments suggest that it may yield substantially better estimates in some circumstances.

Some econometrics book also discuss this problem, for example: [35], [17] and [37].

Chapter 5

Curve fitting and Smoothing

5.1 Introduction

The curve fitting problem can be formulated as follow: given value y_r , $r = 1, ..., m$ of the dependent variable y, corresponding to value x_r , $r = 1, ..., m$ of independent variable x, fit to the y_r a function $y(x) := y(x; \theta)$ of known form but containing a vector θ of n disposable parameter, to be determined such that $y(x_r) \simeq y_r$. As far as the form of the function $y(x)$ is concerned, it is most common to use polynomial or spline function.

The objectives of curve fitting [7] :

- Parameter Estimation

  The form of $y(x)$ may be dedicated by the context of the application in which case the parameter θ_i have specific physical meaning. The primary goal is then to estimated these parameters as accurately as possible from the given data.

- Data Smoothing

  If the given value y_r are accurate enough, it may be sufficient to determine an interpolating function $y(x)$. However, in most application the value y_r will be subjected to measurement errors. We hope that with this curve fitting process, these error will more or less be smoothed out and the graph of $y(x)$ looks smooth enough.

- Functional representation

  The representation of a discrete set of the data point (x_r, y_r) by a function $y(x)$ may have a number of advantages. First of all, value of any point x in the range of representation are now readily obtain. Further more the approximation can be used for determining derivative value, definite integral etc.

- Data reduction

  If the results have to be store for later use, it may be important that the number of parameter θ_i is less than number of data point, in which case we speak of data reduction.

5.2 Curve fitting with Spline

The fact that we choose splines as approximating functions means that the parameter θ_i to be determined (or to be fixed) are:

- the degree k of the spline
- the number and position of the knot

Approximation Criterion

The least-squares criterion

The least-squares criterion is very well known and general approximation criterion. Applied to spline function it means that we have to determine the spline $s(x)$ for which expression

$$\delta := \sum_{r=1}^{m}(w_r(y_r - s(x_r)))^2 \tag{5.1}$$

is minimized. The numbers w_r are weights.

The natural smoothing spline criterion

Find the function $y(x)$ for which

$$\eta := \int_{x_1}^{x_m} \left(y^{(l)}(x)\right)^2 dx$$

is minimal, subject to the condition

$$\delta := \sum_{r=1}^{m}(w_r(y_r - s(x_r)))^2 \leq S$$

where S is a specified number.

The criterion of Powell

Determine a cubic spline $s(x)$ that minimized

$$\varphi := \delta + \sum_{i=1}^{g}(\varpi_i d_i)^2$$

where δ is the result of (5.1). This criteria take care of smoothness of $s(x)$.

5.3 Nonparametric Regression Smoothing

Nonparametric regression attempts to uncover functional relationships without making sweeping assumptions on the type of functional dependence. For this reason, nonparametric regression is the smoothing method of choice when there is no theoretical basis or a priori reason for choosing a particular functional form ([8],[9]).

Cubic spline smoothing is a particularly flexible form of nonparametric regression based on strictly interpolating splines. The smoothing that nonparametric regression performs can be thought of as a process where each data point is replaced by a local average of the surrounding data points. Different nonparametric regression techniques define and calculate this local average in different ways.

The smoothing spline's determination of what is 'local' is based on the data itself [45], making it a particularly flexible smoother. With the underlying mathematical form of the interpolation spline, the smoothing spline has the ability to model a wide range of functional forms while the flexibility of the smoothing procedure makes smoothing splines especially robust.

Like most non-parametric regression techniques, the smoothing spline is itself a function of a smoothing parameter. This parameter determines the balance between fidelity to the data and the smoothness of the curve. Consequently,the successful use of smoothing splines to separate the signal from the noise depends on the choice of the "optimal" smoothing parameter.

In application, we often find that the error are correlated; for example: time series data. We know that correlation affect the selection of smoothing parameter, which are critical to the performance of smoothing spline estimates. That problem has been discussed in [55],[23], [51].

Appendix A

Literature Study

Supervisor :

- dr. K. Poortema
- dr.J.F. Fankena

For my final project two subjects are important:

(1). how to deal with dependent measurements?

(2). methods to do some curve fitting.

For subject (1) time series models have to be studied. Not only the respective model (autoregressive models, moving average models and combinations of these two type of models) have to be studied. Methods for choosing a model and checking whether a model suits the data should be studied as well.

Curve fitting may be done by means of regression models. If e.g. the kth predictor variable, x_k , is chosen to be $x_k = t^k$ the polynomial curves are fitted. Perhaps splines functions may be better choice for the predictor variables. At any case regression models may applied for subject (2) and these model have to extended/generalized in order to deal with dependent measurements.

Regression models with errors described by a time series model are the extensions/generalizations which have to be studied. Search for theory developed for this kind of regression models.

Bibliography

[1] Anderson, R.L. , The Problem of Autocorrelation in Regression Analysis, *Journal of The American Statistical Association 49* (1954), p.113-129.

[2] Beach, C.M., A Maximum Likelihood Procedure for Regression with Autocorrelated Error, *Econometrica* **46** (1978), 1, p.51-54.

[3] Beran, J. , Maximum Likelihood Estimation of the Differencing Parameter for Invertible Short and Long Memory Autoregressive Integrated Moving average Models, *Journal of Royal Statistical Society Serie B* **57** (1995), 4, p. 659-672.

[4] Bosq, D. and Shen, J., Estimation of an Autoregressive Semiparametric Model with Exogenous Variables, *Journal of Statistical Planning and Inference* **19**(1998), 2, p.105-127.

[5] Box, G.E.P and Jenkins, G.M, *Time Series Analysis : forecasting and control*, revised edition,Holden Day, San Francisco, 1976.

[6] Chen, Z., Fitting Multivariate Regression Functions by Interaction Spline Models, *Journal of The Royal Statistical Society serie B* **55**(1993), 2, p. 473-471.

[7] Dierckx, P. , Curve and Surface Fitting with Splines, Oxford Science Publications, Oxford, 1993

[8] Eubank, R.L., *Spline Smoothing and Nonparametric Regression,* Marcel Dekker, New York, 1988.

[9] Eubank, R.L. , A Simple Smoothing Spline, *The American Statistician* **48**(1994), 2, p.103-106.

[10] Denby, L., and Martin, R.D. , Robust Estimation of the First-Order Autoregressive Parameter, *Journal of The American Statistical Association* **74**(1979), 365, p.140-146.

[11] Gallant, A.R. and Goebel, J.J., Nonlinear Regression with Autocorrelated Error, *Journal of The American Statistical Association* **71** (1976), 356, p.961-967.

[12] Goutte, C. , Lag Space Estimation in time Series Modelling, preprint.

[13] Hannan, E.J. , Non-linear Time Series Regression, *Journal of Applied Probability*, **8** (1971), p.767-780.

[14] Hannan, E.J. and Nicholls, D.F., The Estimation of Mixed Regression, Autoregression, Moving Average, and Distributed Lag Models, *Econometrica* **40** (1972), p.529-547.

[15] Hallin, M. et. al., Characterization of Error Distributions in Time Series Regression Models, *Statistics and Probability Letters* **38**(1998), 4, p.335-345.

[16] Harvey, A.C. and Philips, G.D.A, Maximum Likelihood Estimation of Regression Models with Autoregressive-moving Average Disturbance, *Biometrika* **66** (1979), 1, 49-58.

[17] Harvey, A.C., *The Econometric Analysis of Time Series*, Philip Allan, Oxford, 1981.

[18] Harvey, A.C., *Time Series Model*, Philip Allan, New York, 1981.

[19] Haugh, L.D and Box, G.E.P, Identification of Dynamic Regression (Distributed Lag) Models Connecting Two Time Series, *Journal of The American Statistical Association* **72** (1977), 357, p.121-130.

[20] He, X., Linear Regression after Spline Transformation, *Biometrika* **84** (1997), 2, p.474-481.

[21] Hokstad, P. , A Method for Diagnostic Checking of Time Series Model, *Journal of Time Series Analysis* **4** (1983), 3, p.177-183.

[22] King, M., Testing for Autoregressive Against Moving Average Errors in the Regression Models with Autoregressive-moving Average Errors, *Journal of Econometrics* **21** (1971), p. 299-312.

[23] Kohn, R. et.al., Nonparametric Spline Regression with Autoregressive Moving Average Errors, *Biometrika* **79** (1992), 2, p.335-346.

[24] Koul, H. and Schick, A. , Efficient Estimation in Nonlinear Autoregressive Time-series Models, *Bernoulli* **3** (1997), 3, p.247-277.

[25] Kwakernak, H., Time series analysis and System Identification, Lecture Notes, University Twente, Enschede, 1998.

[26] Ljung, G.M., Diagnostic Testing of Univariete Time Series Models, *Biometrika* **73** (1986), 3, p. 725-730.

[27] Luo, Z. and Wahba,G. , Hybrid Adaptive Spline, *Technical Report no. 947 (1995)*, Department of Statistics University of Wisconsin at Madison.

[28] McLeod, A.I. , Diagnostic Checking ARMA Time Series Model Using Squared-residual Autocorrelation, *Journal of Time Series Analysis* **4** (1983), 4, p.269-273.

[29] Mammen, E. , et.al. A general Framework for constrained smoothing, *Technical Report* (1998), Department of Biostatistics Harvard School of Public Health.

[30] Mentz, R.P. et.al., On Residual Variance Estimation in Autoregressive Model, *Journal of Time Series Analysis* **19**(1998), 2, p.187-208.

[31] Neath, A.A.and Cavanaugh, J.E., Regression and Time Series Model Selection using Variants of The Schwarz Information Criterion, *Communications in Statistics-Theory and Methods **26** (1997),* p.559-580.

[32] Pena, D. , Measuring influence in Dynamic Regression Models, *Technometrics* **33** (1991),1, p.93-102.

[33] Palm, F.C. and Nijman , Missing Observations in the Dynamic Regression Model, *Econometrica **52** (1984), 6,* p.1415-1435.

[34] Pankratz, A., *Forecasting with Univariate Box-Jenkins Models : Concept and Cases*, John Wiley & Sons, New York, 1983.

[35] Pankratz, A., *Forecasting with 'Dynamic Regression Models*, John Wiley & Sons, New York, 1991.

[36] Pankratz, A., Detecting and Treating Outliers in Dynamic Regression Models, *Biometrika* **80** (1993), 4, p.847-856.

[37] Pesaran, M.H. and Slater, L.J., *Dynamic regression : Theory and Algorithms,* Ellis Horwood Ltd, 1980.

[38] Pandit, S.M. and Wu, S.M., *Time Series and System Analysis with Application*, John Wiley and Sons,New York, 1983

[39] Pierce, D.A. , Least Square Estimation in the Regression Model with Autoregressive-Moving Average Error, *Biometrika* **58** (1971), 2, p.299-312.

[40] Pierce, D.A. , Distribution of Residual Autocorrelation in the Regression Model with Autoregressive-Moving Average Errors, *Journal of the Royal Statistical Society serie B* **33** (1971), 1, p.140-146.

[41] Pierce, D.A. , Residual Correlations and Diagnostic Checking in Dynamic-disturbance Time Series Models, *Journal of The American Statistical Association* **67** (1972), p.636-640.

[42] Sargan, J.D. and Bhargava, A., Maximum Likelihood Estimation of Regression Models with First Order Moving Average Errors when The Root Lies on The Unit Circle, *Econometrica* **51** (1983), 3, p.799-820.

[43] Seber, G.A.F., and Wild, C.J., *Non Linear Regression*, John Wiley & Sons, New York, 1989.

[44] Shi, P. , M-type Regression Splines Involving Time Series, *Journal of Statistical Planning and Inference* **61**(1997), 1, p.17-38.

[45] Silverman, B.W., Some Aspects of the Spline Smoothing Approach to Nonparametric Regression Curve Fitting, *Journal of the Royal Statistical Society Series B* (1985), 47,p.1-52

[46] Teixeira, J.C. and Rodrigues, A.J., An Applied Study on Recursive Estimation Methods, Neural Networks and Forecasting, *European Journal of Operational Research* **100**(1997), p. 406-417.

[47] Tran, L. et.al. , Fixed Design Regression for Linear Time Series, *The Annals of Statistics 24 (1996), 3, p. 975-991..*

[48] Tsay, R.S., Regression Models with Time Series Errors, *Journal of The American Statistical Association* **79** (1984), 385, p.118-124.

[49] Turkington, D., Efficient Estimation in The Linear Simultaneous Equations Model with Vector Autoregressive Disturbances, *Journal of Econometrics* **85** (1998), 1, p.51-74.

[50] Venkataraman, K.N., Some Limit Theorem on Regression Models with a Nonstationary and Possible Nonlinear Time Series Errors, J*our. Math. Phys. Sci.* ***16*** *(1982), 4,* p.383-404.

[51] Wand, M.P., A Comparison of Regression Spline Smoothing Procedures, *Technical Report (1997)*, Australian Graduate School of Management University of New South Wales.

[52] Wand, M.P., On the Optimal Amount of Smoothing in Penalized Spline Regression, *Technical Report (1998)*, Department of Biostatistics, Harvard School of Public Health.

[53] Wang, Y. GRKPACK: Fitting Smoothing Spline ANOVA Models for Exponential Families, *Technical Report no. 942 (1995)*, Department of Statistics University of Wisconsin at Madison.

[54] Wang, Y. , Mixed-effect Smoothing Spline ANOVA, *Technical Report no. 967 (1996)*, Department of Statistics University of Wisconsin at Madison.

[55] Wang, Y. , Smoothing Spline Models with Correlated Random Errors, *Journal of The American Statistical Association 93*(1998), 441, p.341-348.

[56] Xiang, D. and Wahba, G. , Approximate Smoothing Spline Methods for Large Datasets in The Binary Case, *Technical Report no. 982 (1997)*, Department of Statistics University of Wisconsin at Madison.

References

[1] Anderson, R.L. , The Problem of Autocorrelation in Regression Analysis, *Journal of The American Statistical Association 49* (1954), p.113-129.

[2] Beach, C.M., A Maximum Likelihood Procedure for Regression with Autocorrelated Error, *Econometrica* **46** (1978), 1, p.51-54.

[3] Box, G.E.P and Jenkins, G.M, *Time Series Analysis : forecasting and control*, revised edition,Holden Day, San Francisco, 1976.

[4] Dierckx, P. , Curve and Surface Fitting with Splines, Oxford Science Publications, Oxford, 1993

[5] Eubank, R.L., *Spline Smoothing and Nonparametric Regression,* Marcel Dekker, New York, 1988.

[6] Gallant, A.R. and Goebel, J.J., Nonlinear Regression with Autocorrelated Error, *Journal of The American Statistical Association* **71** (1976), 356, p.961-967.

[7] Hannan, E.J. , Non-linear Time Series Regression, *Journal of Applied Probability,* **8** (1971), p.767-780.

[8] Harvey, A.C. and Philips, G.D.A, Maximum Likelihood Estimation of Regression Models with Autoregressive-moving Average Disturbance, *Biometrika* **66** (1979), 1, 49-58.

[9] Kwakernak, H., Time series analysis and System Identification, Lecture Notes, University Twente, Enschede, 1998.

[10] Pankratz, A., *Forecasting with Univariate Box-Jenkins Models : Concept and Cases*, John Wiley & Sons, New York, 1983.

[11] Pierce, D.A. , Least Square Estimation in the Regression Model with Autoregressive-Moving Average Error, *Biometrika* **58** (1971), 2, p.299-312.

[12] Pierce, D.A. , Distribution of Residual Autocorrelation in the Regression Model with Autoregressive-Moving Average Errors, *Journal of the Royal Statistical Society series B* **33** (1971), 1, p.140-146.

[13] Silverman, B.W., Some Aspects of the Spline Smoothing Approach to Nonparametric Regression Curve Fitting, *Journal of the Royal Statistical Society Series B* (1985), 47,p.1-52

[14] Tsay, R.S., Regression Models with Time Series Errors, *Journal of The American Statistical Association* **79** (1984), 385, p.118-124.