Statistical-Grey Consistent Grey Differential Equation Modelling

Yan Hong Cui

MSc Thesis Submitted on 21 August 2007

Word Count: 27 358

MSc Thesis Presented for the Degree of Master of Science, in the Department of Statistical Sciences, University of Cape Town, South Africa.

Supervisor: Professor Renkuan Guo

Declaration

I, Yan Hong Cui, hereby,

1. grant the University of Cape Town free licence to reproduce the above thesis in whole or in part, for the purpose of research;
2. declare that:
   i. the above thesis is my own unaided work, both in concept and execution,
   ii. neither the substance nor any part of the above thesis has been submitted in the past, or is being, or is to be submitted for a degree at this University or at any other university.

I am now presenting the thesis for examination for the degree of Master.

Yan Hong Cui

Abstract

The thesis title is “Statistical-Grey Consistent Grey Differential Equation Modelling”. Generally, we defined uncertainty as the events that can not be described by precise deterministic laws. The research purpose of this thesis is to build the Statistical-Grey Consistent Grey Differential Equation Model based on Grey mathematics to solve the small sample data uncertainty problems.

Grey mathematics was initiated by Professor Ju-long Deng in 1982 and quickly obtained wide attention in the mainland of China, Taiwan, Hong Kong, Japan, and others. Grey mathematics is small-sample oriented and in nature it is a subset of modern approximation theory. In grey mathematics many efficient and accessible features were developed and are applicable in research, business and industries.

However, we note that linear regression models are used in grey modelling as a medium to obtain the coefficients for grey models but the standard grey mathematics throws away the statistical information while just using the coefficient estimates to facilitate a trend analysis. We further notice that even the statistical insignificant coefficient estimators are still inserted into grey models for prediction purposes. We call this phenomenon as statistical-grey inconsistency in grey modelling.

The research objective of this thesis is from examining Deng’s standard GM (1, 1) model, to point out that the nature of Grey Differential Equation model is a pair of differential equation model and a regression model translated from the differential equation by discreteness .The phenomenon of the organically packed two models, we call it as Coupling Principle in GDE modelling. Then we realized that the problem of addressing high statistical efficiency demand on GDE data modelling, which in nature a problem whether or not the model engaged could capture the pattern carried in the sample data. Just like the standard regression modelling exercise, different sample data may contain different trends or patterns and thus different forms of regression models should be explored. Based on such believe, we extended GM (1, 1) model and thus establish a larger modelling possessing diversified pattern-catching capacity. In this thesis we also propose a distance measure between optimally data-fitted spline functional and constraint functional of GM (1, 1) and in terms of the distance measure for seeking the optimal constraint functional for GM (1, 1) modelling. We also explored the possibility to relax the strictly positive

discrete data sequence assumption extend to arbitrary discrete data sequence assumption for grey differential equation modelling in terms of optimally data-fitted spline functional.

In chapter 2, we first reviewed some fundamental features of grey mathematics, particularly, GDE models. Then we start to introduce the meaning of grey uncertainty and the procedure of GDE modelling, especially, the first-order, one-variable GDE model, abbreviated as GM (1, 1) model.

In chapter 3, we pointed out the importance of the goodness-of-fit information of the regression model itself. During the modelling procedure, if the data assimilated parameter pair has a great efficiency, while the GDE model with the significant coefficient also has reasonable model efficiency. After that, we gave the definition of statistical-grey consistency and statistical-grey inconsistency. And we also mentioned about a new ratio idea to improve efficiency of Statistical-Grey Consistency model.

In chapter 4, we reviewed a couple of class of spline functions under certain optimal conditions. Then we re-examined the nature and the optimality of the two critical data operations proposed by Deng (2002), accumulative generating operation (abbreviated as AGO) and the inverse accumulative generating operation (abbreviated as IAGO) and then the exchangeability between AGO, IAGO and the integration of certain spline function, the derivative of certain spline function respectively. We further explored the roles of spline functions in GM(1,1) model, Particularly, we proposed a distance measure between optimally data-fitted spline functional and constraint functional of GM(1,1) and in terms of the distance measured for seeking the optimal constraint functional for GM(1,1) modelling. Finally, we explored the possibility to relax the strictly positive discrete data sequence assumption extend to arbitrary discrete data sequence assumption for grey differential equation modelling in terms of optimally data-fitted spline functional.

In chapter 5, we explored the optimality nature of GM(1,1) model from various angles and finally we identified the coupling feature in GM(1,1) model, which we call it as Coupling Principle in GDE modelling or simply Coupling Principle. At this new standing point we recognized GM(1,1) model is a pair of differential equation model and a regression model translated from the differential equation by discretization. Therefore, we explored the fundamental features of the organically packed two models – the motivated differential equation model and the coupled regression model.

In chapter 6, we explored the problem of addressing high statistical efficiency demand in GDE data modelling, which in nature a problem whether or not the model engaged could capture the pattern carried in the sample data. Just like the standard regression modelling exercise, different sample data may contain different trends or patterns and thus different forms of regression models should be explored. Based on such believe, we will extend GM(1,1) model and thus established a larger modelling family possessing diversified pattern-catching capacity.

In chapter 7, we list all the Matlab toolbox program and visual C++ program uses for this thesis.

Yan Hong Cui 21 August 2007

Acknowledgements

First, I sincerely thank my supervisor, Professor Renkuan Guo, for his research-lead supervision. Professor Guo patiently trained me to read and learn relevant grey mathematics literature, brought me into his problem-addressing research on grey differential equation theory, assigned me from time to time certain programming and computing tasks, and via these tasks further solidifying my understanding on his research ideas. I can say that without my supervisor's invaluable guidance I would not have today's achievements in research exploration and knowledge expansions.

Secondly, heartfelt hanks are given to my parents for their far-sightedness and insistence on further education for me, and for their constant encouragement and support.

I would like to thank Dr. Danni Guo for her many helps on research, thesis writings and the thesis template created by her. Many thanks to Professors Dunne, Thiart and Troskie for their supports during my MSc study in the Department of Statisticsl Sciences.

I must thank my roommate, Mr. Li, for his smart and competent management in our daily life during our MSc studies.

At last, I need to thank my parents for their encouragements and support.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

1.1 Problem Addressing

In today's rapid competitive business world, we have to make lot of complicated decisions as quickly as possible. In business world, it is often face to decision without sufficient data. Therefore small sample methodology is very critical. The one-variable grey differential equation model (GDE), proposed by Deng, (1985) requires very small sample size, as little as four sample points, but the model may often provide reasonable predictability and model accuracy. The grey differential equation modelling procedure is a simple 3-step approach.**1. We need utilize accumulative generation operation to operate a data transformation, and utilize inverse accumulative generation operation transfer data back after the estimation. 2. Fit a simple regression model for seeking parameters estimation in GDE model. 3. Simulate the final solution of whitening differential equation for filtering or prediction.**

Grey approach is small-sample oriented and in nature it is a subset of approximation theory in general which offers many efficient and accessible features and therefore enjoys a very wide industrial applications.

However, we note that linear regression models are used in GDE modelling as a medium to obtain the coefficients for grey models but the standard grey mathematics throws away the statistical information while just using the coefficient estimates to facilitate the data distribution tendency. We further notice that even the statistical insignificant coefficient estimators are still inserted into GDE models for prediction purposes. We call this phenomenon "statistical-grey inconsistency" in grey modelling.

In aim of the thesis

Part A: (1) To paid data into GM model, during the modelling procedure, if the data assimilated parameter pair has a great efficiency, while the GDE model with the significant coefficient also has reasonable model efficiency. (2) Define the modelling is statistical-grey consistency or statistical-grey inconsistency. (3) Introduce a new ratio idea to improve efficiency of Statistical-Grey Consistency model.

Part B: We will explore the roles of spline functions in GM (1, 1) model, to show the exchangeability between accumulative generating operations (AGO), the inverse accumulative generating operation (IAGO) and the integration of certain spline function, the derivative of certain spline function respectively. We further explore the roles of spline functions in GM(1,1) model, particularly, we propose a distance measure between optimally data-fitted spline functional and constraint functional of GM(1,1) and in terms of the distance measure for seeking the optimal constraint functional for GM(1,1) modelling. Finally, we explore the possibility to relax the strictly positive discrete data sequence assumption extend to arbitrary discrete data sequence assumption for grey differential equation modelling in terms of optimally data-fitted spline functional.

Part C: After we explore the fundamental features of the organically packed two models – the motivated differential equation model and the coupled regression model, we will show the problem of addressing in this thesis, high statistical efficiency demand in GDE data modelling, which in nature a problem whether or not the model engaged could capture the pattern carried in the sample data. Just like the standard regression modelling exercise, different sample data may contain different trends or patterns and thus different forms of regression models should be explored. Based on such believe, we will extend GM (1, 1) model and thus establish a larger modelling family possessing diversified pattern-catching capacity. We also created some statistical-grey consistent grey differential equation model computer programs for this thesis and attached all of them in appendix.

1.2 Thesis Structure

We structure the remaining paper as follows:

In chapter 2, we first review some fundamental features of grey mathematics, particularly, GDE models. Then we start to introduce the meaning of grey uncertainty and the procedure of GDE modelling, especially, the first-order, one-variable GDE model, abbreviated as GM (1, 1) model.

In chapter 3, we point out the importance of the goodness-of-fit information of the regression model itself. During the modelling procedure, if the data assimilated parameter pair has a great efficiency, while the GDE model with the significant coefficient also has reasonable model efficiency. After that, we give the definition of statistical-grey consistency and statistical-grey inconsistency. And we also mentioned about a new ratio idea to improve efficiency of Statistical-Grey Consistency model.

In chapter 4, we review a couple of class of spline functions under certain optimal conditions. Then we re-examine the nature and the optimality of the two critical data operations proposed by

Deng (2002), accumulative generating operation (abbreviated as AGO) and the inverse accumulative generating operation (abbreviated as IAGO) and then the exchangeability between AGO, IAGO and the integration of certain spline function, the derivative of certain spline function respectively. We further explore the roles of spline functions in GM(1,1) model, Particularly, we propose a distance measure between optimally data-fitted spline functional and constraint functional of GM(1,1) and in terms of the distance measure for seeking the optimal constraint functional for GM(1,1) modelling. Finally, we explore the possibility to relax the strictly positive discrete data sequence assumption extend to arbitrary discrete data sequence assumption for grey differential equation modelling in terms of optimally data-fitted spline functional.

In chapter 5, we are exploring the optimality nature of GM(1,1) model from various angles and finally we identify the coupling feature in GM(1,1) model, which we call it as Coupling Principle in GDE modelling or simply Coupling Principle. At this new standing point we recognize GM(1,1) model is a pair of differential equation model and a regression model translated from the differential equation by discretization. Therefore, we explore the fundamental features of the organically packed two models – the motivated differential equation model and the coupled regression model.

In chapter 6, we explore the problem of addressing high statistical efficiency demand in GDE data modelling, which in nature a problem whether or not the model engaged could capture the pattern carried in the sample data. Just like the standard regression modelling exercise, different sample data may contain different trends or patterns and thus different forms of regression models should be explored. Based on such believe, we will extend GM(1,1) model and thus establish a larger modelling family possessing diversified pattern-catching capacity.

In chapter 7, we list all the Matlab toolbox program and visual C++ program uses for this thesis.

Chapter 8 is the final chapter for conclusions.

1.3 Software Used for Study

The softwares for this thesis are developed in MatLab and Microsoft visual C++.

We list all the MatLab toolbox program and Visual C++ program in Chapter 7. The algorithm we use for optimize the function value is genetic algorithm. In this thesis we utilize genetic algorithm methodology to estimate the parameters fitting for coupled regression model.

The program's structure in this thesis (for details please check chapter 7):

1. Based on GM (1, 1) model procedure, utilize visual C++ and Matlab language program to build software for GM (1, 1) modelling.

Program structure:

Import the sample size of the modelling data sequence (denotes as K) from keyboard

↓

Assume $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$ is the modelling data sequence, then $\sigma(k) = \dfrac{x^{(0)}(k)}{x^{(0)}(k+1)}, k = 1,2,\cdots,n$, are called stepwise ratios of the sequence $X^{(0)}$ and $\exp(-2/(k+1)) \le \sigma(x) \le \exp(2/(k+1))$

↓

Import the modelling data sequence $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$ from keyboard

↓

Take 1-time accumulating generation operation of $X^{(0)}$, to generate the data sequence. $X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right)$

↓

Check whether the modelling data sequence $X^{(0)}$ obeys the stepwise ratio law. $\exp(-2/(k+1)) \le \sigma(x) \le \exp(2/(k+1))$

↓

Generate $z^{(1)}$ data sequence from $z^{(1)} = \dfrac{1}{2}\left(x^{(1)}(k) + x^{(1)}(k-1)\right)$

↓

Utilizing least squares estimate the parameter sequence $\hat{a} = [a,b]^T = \left[B^T B\right]^{-1} B^T Y$, $B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}$, $Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}$

Generate the grey model and the time response sequence from $\hat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a}\right]e^{-ak} + \frac{b}{a}$

↓

Simulate value of $X^{(1)}$ and estimation value of $X^{(0)}$ from equation $\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$

↓

Calculate the Errors and Relative Errors

↓

Type 'Y' from keyboard when finished typing keys

2. Based on GM (1, 1) prediction model procedure, utilize visual C++ language program to build software for GM (1, 1) model prediction.

Import the sample size of the modelling data sequence (denotes as K) from keyboard

↓

Assume $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$ is the modelling data sequence, then $\sigma(k) = \frac{x^{(0)}(k)}{x^{(0)}(k+1)}, k = 1,2,\cdots,n$, are called stepwise ratios of the sequence $X^{(0)}$ and $\exp(-2/(k+1)) \leq \sigma(x) \leq \exp(2/(k+1))$

↓

Import the modelling data sequence $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$ from keyboard

↓

Take 1-time accumulating generation operation of $X^{(0)}$, to generate the data sequence. $X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right)$

↓

Check whether the modelling data sequence $X^{(0)}$ obeys the stepwise ratio law. $\exp(-2/(k+1)) \leq \sigma(x) \leq \exp(2/(k+1))$

↓

Generate $z^{(1)}$ data sequence from $z^{(1)} = \frac{1}{2}\left(x^{(1)}(k) + x^{(1)}(k-1)\right)$

↓

Utilizing least squares estimate the parameter sequence $\hat{a} = [a,b]^T = \left[B^T B\right]^{-1} B^T Y$, $B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}$, $Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}$

↓

Enter how many data you want to predict from keyboard

↓

Generate the grey model and the time response sequence from $\hat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a}\right] e^{-ak} + \frac{b}{a}$

↓

Simulate value of $X^{(1)}$ and estimation value of $X^{(0)}$ from equation $\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$

↓

Simulate value of predict data

↓

Calculate the Errors and Relative Errors

↓

Type 'Y' from keyboard when finished typing keys

3. Based on Statistical-Grey Consistency GM (1, 1) model procedure, utilize Matlab program to build Matlab toolbox.

Program structure:

The program structure of Statistical-Grey Consistency GM (1, 1) model base on the program structure of GM (1, 1) modelling procedure and combined with lot of necessary information of modelling export, such as: 1-AGO data sequence value $X^{(1)}$; the parameters of regression model α, β; the estimation value of 1-AGO data sequence value $\hat{X}^{(1)}$; $z^{(1)}$ data sequence value from $z^{(1)} = \frac{1}{2}\left(x^{(1)}(k) + x^{(1)}(k-1)\right)$; R^2 value of regression model; R^2 value of Statistical-Grey Consistency model; error term of $X^{(1)}$ level; error term of Statistical-Grey Consistency model; average error of Statistical-Grey Consistency model; Confidence of Statistical-Grey Consistency model.

In this dissertation, we also utilize genetic algorithm to optimize our program. Genetic algorithm (abbreviated as GA) was formally introduced in University of Michigan, United States, by John Holland in 1970s .GA inspired by Charles Darwin's Evolution theory, transforming the equation solution searching space into biological evolution space, made all the possible solution into vectors- chromosome, each element of the vector we call it gene. Then the GA selects random chromosome from the population of "parents" and born the "children" population for next generation, after repeat modifies and over successive generations, the population "Evolves" toward an optimal solution.

Chapter 2. A Review on Grey Differential Equation Models

Uncertainty problems exist commonly in every spectrum of scientific research fields and technical developments. In general, there are many types of uncertainties. The two essential ones are random uncertainty and fuzzy uncertainty.

Fuzzy mathematics initiated by Zadeh, L. A. (1965) and axiomatically founded by Liu, B. and Liu, Y. K. (2002), mainly deal with problem of the fuzzy phenomena, featured by cognitive uncertainty by experience with the help of affiliations. An example is the degree of tall people. Most would like agree that one person taller than 180cm is a tall guy, but what about a person whose height is 180cm or 179cm? So it's difficult for us to say one people tall or not when we face a real person. Obviously, given a well-defined realm seems like absurd. Uncertainty reasoning becomes valuable accurately when we're facing how people really recognize the concept tall as opposed to a simple-minded classification useful for accounting purposes only.

Probability theory deals with random phenomena. The random uncertainty law, typically represented the probability distribution, is established in terms of the sample drawn from the population under study. Hypothesis testing and statistical estimation procedures are all relying on the sampling. The inference quality depends upon the sample size.

In today's rapid competitive business environments, people have to make decisions based on information available, including sample data. Data collection is expensive and often impossible during to time constraint. This implies that people may often face the circumstances of sparse data available.

Thus the grey mathematics was developed for addressing the decision making under sparse data. Professor Deng, J. L. (1982), first proposed the grey mathematics in 1982, his first paper was published in 1982, which is ukontrol problems of Grey Systems".

Then the initiative idea of 80's quickly obtained attentions in the mainland of China, Taiwan, ,mHong Kong, Japan, and other areas of the world. Over hundreds of university offered course work and practicum of grey mathematics. Many illustrious academic scholars provide their great support of it, such as: Chinese academicians Song Jia, Qian Xue-shen, and Zhang Zhong-jun. A former editor-in-chief of the journal Systems and Control Letters, Professor Roger W. Brockett of Harvard University, Who offered a comment as: "Grey system is an initiative work" and "all the results are new". "Journal of Grey System" under Professor Deng Ju-long's supervision established in 1989 now becomes major Communication terrace for grey system mathematics scholars. Thousands of articles in relation to grey system mathematics published in various international journals and lots of international conferences listing grey systems as a special topic. Less than 20 years, grey mathematics has been successfully applied to widespread fields: industry technology, agriculture, medicine, robot, image processing, etc.

2.1 System

In our daily social life, education, economics, even scientific research activities, everywhere, we always involved with a term "system", but lots of times it is hard for us to give an exactly meaning to say what it is.

The term "system" originated from very old ancient Greek language at the earliest stage, its original meaning is the common characteristic of the events and the position of each thing should occupy. Based on the background of modern science technical development and the practice of mankind social practice, the connotation of the system concept already extends and has its new meaning.

There have appeared various concepts of system in history, Webster point out a system should be described as "an organic or organized whole; a synthesis of various notions and principles forming the whole; an aggregate of elements that are regular, rely on and react on each other." Another scholar Bertalanffy, K.L.V. (1951) defined the system as "the whole of the elements that react on each other".

In today's modern scientific technology, a system often means a group of interacting, interrelated, or interdependent elements forming a complex whole organism, especially with regard to its vital processes or functions. But the definition of the system aren't quite common in today's academic area, in here, we try to enumerate 2 popular ones.

1. Describe a system as a set contains various identifiable, independence elements.

$$A = \{a_i \in A | i = 1,2,\cdots,n\} \tag{2.1}$$

Where a_i refer as the system's all elements.

2. Describe a system as a collection of 5 dissimilar elements.

$$A = \{X, Y, Z, \alpha, \beta\} \tag{2.2}$$

Where X refers as a input space;

Y refers as output space;

Z refers as state space;

α refers as transition function;

β refers as output function. (Modified from Li, F., 1992)

2.2 Information and Uncertainty Information

Before our discussion with uncertainty information in a grey system, we should better know what are the information and the definition of it. By tracing the development of modern science and social history, there has not been given an exactly definition of information accepted by the majority information science scholars.

First let us review some definitions of information already been given by some knowledgeable scholars' literature: Douglas Raber & John M. Budd (2003) point out "from the perspective of semiotics, "information" is an ambiguous theoretical concept because the word is used to represent both signifier and signified, both text and content". Shannon, C. E. (1948) defined information as "information is eliminated random uncertainty". Ashby, W.R. (1964) described it as "information is the variation of a thing". In Buckland, M. K.'s, (1991), published book "Information and Information System" mentions "Many problems come immediately out, provided that you discuss the intention of information. The concept of information is significant only when people obtain the information and understand it. It also gives us much food for thought. The word information itself is equivocal and its application is diversified."

From the above points of view of information, we divide the concept of information into four aspects: Information as a representation of knowledge, Information as data in the environment, Information as part of communication, Information as a resource or commodity.

The rapid development of modern science and the progress of human society's demands provided the environments and conditions for us to collect enough information to make a decision, so we can easily collect information from internet, newspapers, magazines, television broadcasting, from anywhere we could. If we consider various environments and conditions (in real world they are internet, newspapers, magazines ...) as different systems, we may say that systems involving without people including become less and less impotent. Mostly, Uncertainty information may appear if the system interest consider about the factors of human beings.

The types of Uncertainties we already known based on published studies listing as below:

1. Grey uncertainty.
2. Stochastic uncertainty,
3. Unascertainty,
4. Fuzzy uncertainty,
5. Rough uncertainty,
6. Soros reflexive uncertainty. (modified from Liu, S., and Lin, Y. (2006))

2.2.1 Grey Uncertainties

Suppose A is a piece of grey information defined as follows. Let x be an unknown, $S \neq \phi$ a set, S' a subset of S, $u =$"x belongs to S" and $A =$"x belongs to S'". Then the so-called grey uncertainty stands for the uncertainty of which specific value of the unknown x should take. For example, suppose we are given that $u =$"x belongs to S", $S =$"R is the set of all real numbers", S' = the interval [2, 3], and $A =$"x belongs to S'". Then the piece of grey information A brings about the following uncertainty: we know that x is a number between 2 and 3 inclusive. However, we do not know which value x really assumes. (Modified from Liu, S., and Lin, Y. (2006))

2.2.2 Stochastic Uncertainty

If x is unknown, S a nonempty set, $u =$"x belongs to S" and $A =$"x belongs to S and the possibility for $x = e \in S$ is α_e, where $0 \leq \alpha_e \leq 1$ and $\sum_{e \in S} \alpha_e = 1$." In this case A is called a piece of stochastic information. When a piece of stochastic uncertainty. Such uncertainty is created because the piece of stochastic information A can only spell out how likely the unknown x equals a special element $e \in S$. This implies that the probability α_e can be very close to 1 or equal to 1, however, the large probability does not guarantee that $x = e$ will definitely be true. (Modified from Liu, S., and Lin, Y. (2006))

2.2.3 Unascertainty

If in the definition of a piece of stochastic information A, we replace the condition that

$$\sum_{e\in S}\alpha_e = 1 \tag{2.3}$$

by

$$\sum_{e\in S}\alpha_e \leq 1 \tag{2.4}$$

then A is called a piece of unascertained information.

The main difference between stochastic and unascertained information is that the former concept is developed on the assumption that all possible outcomes of an experiment are known, whereas for unascertained information, we assume that only some possible outcomes of the experiment are known to the researcher. (Modified from Liu, S., and Lin, Y. (2006))

2.2.4 Fuzzy Uncertainty

A piece A of tidings is called a piece of fuzzy information, if A satisfies: x is an unknown, S a nonempty set, the position tiding $u =$"x belongs to S" and $A =$"x belongs to S and the degree of the membership for $x = e \in S$ is $\alpha_e, 0 \leq \alpha_e \leq 1$."(Modified from Liu, S., and Lin, Y. (2006))

2.2.5 Rough Uncertainty

Let u be a set of elements, And, a subset $r \subseteq p(u)$, the power set of u, is called a partition of u, if the following conditions hold true:

1. $\bigcup r = \bigcup\{x : x \in r\} = U$,
2. $\forall A, B \in r$, If $A \neq B$, then $A \bigcap B = \phi$.

Let $K = (u, R)$ be knowledge base over u, where u is the universal set of all objects involved in a study, and R is a given set of partitions of the set u, called a knowledge base over u, A subset $X \subseteq u$ is called exact in K, if there exists a $P \subseteq R$ such that X is the union of some elements in $\bigcap P$. Otherwise, X is said to be rough in K. (Modified from Liu, S., and Lin, Y. (2006))

2.2.6 Soros Reflexive Uncertainty

Let x be the unknown path a true historic process will eventually take and S = all possible outcomes of this historical process. Then, a Soros reflexive uncertain information is defined as follows: $u = "x \in S"$ and $A \subseteq S$ is a piece of information regarding the position of x in S defined by A = "if it is expected $x = e \in S$ with a degree of credence α_e, $0 \le \alpha_e \le 1$, then $x = e \in S$ has a degree of credence $1 - \alpha_e$." Now the uncertainty associated with a piece of Soros reflexive uncertain information is that the more accurate a prediction about a true historical process is, the more uncertain the expected future will become. (Modified from Liu, S., and Lin, Y. (2006))

2.3 Fundamental Concepts of Grey System Mathematics

Grey system is different from other systems' naming, was chosen based on dissimilar colours under investigation. According to Deng's (1985, 2002) description, in modern control theory, "black box" habitually represent a system with internal relations was totally unknown. Here, we provide a comparable definition by using word "black" standard for the system with the information was totally unknown. "White" standard for the system with the information was complete known. " Grey" is the colour, degree of clarity lies between "White" and "black", which represent a system information is partially known and partially unknown. Greyness's basic characteristic in grey mathematics is incomplete information, but "grey" also can be extended if we view it from different angles and make use of it in diverse conditions. (Table 2.3.1) for more details:

Table 2.3.1 Comparison between black, grey and white systems

	Black	Grey	White
Information	Unknown	Incomplete	Know
Appearance	Dark	Grey	Bright
Process	New	Replace old with new	Old
Property	Chaos	Complexity	Order
Methodology	Negative	Transition	Positive
Attitude	Indulgence	Tolerance	Seriousness
Conclusion	No solution	Multiple solution	Unique solution

(Modified from Liu, S., and Lin, Y. (2006))

It is necessary to point out that the grey system concept engaged by grey research community is merely a descriptive terminology used in modern control theory. This description is not rigorous enough to facilitate a definition of the grey system as the starting point to establish a mathematical foundation. The reason why we can not start with the descriptive terminology of the grey system used in modern control theory to define a grey system is that it accepts a system, where "the information is partially known and partially unknown" as a grey system. This description is too obscure to be a mathematical definition to establish a solid mathematical for grey theory. As a matter of fact there are many systems where "the information is partially known and partially unknown". For example, Hidden Markov Models in stochastic process theory. A Hidden Markov system has its true state following a Markov model but is unobservable. The observational data is not the true state information but contains noises. Therefore, Hidden Markov system is a system in which "the information is partially known and partially unknown" but it belongs to the family of stochastic systems. In other words, a Hidden Markov system is not a grey system. Therefore a system with partially known and partially unknown information can not be as a definition of grey system.

Without any doubt, over past 26 years' developments on grey theory, a critical feature of grey theory lies in small sample inferences. The information incompleteness of a grey system is caused by sparse data availability extracted from the system under study. In other words, using the feature of an uncertainty rooted from small sample size to characterize a grey system is logical. Explicitly, a grey system has only small sample information available for system dynamic investigation even though this is not still rigorous enough as a rigorous definition for a grey system.

The school of grey system theory clearly identifies itself as a mathematical branch to deal with small sample inference problems. For example, the pioneer of the grey system theory, Professor Deng wrote: "..., the ones in myriad sample can be solved by probability and statistics ways, ..., uncertainty in less data little sample, incomplete information and devoid of experience, which is just suitable to be deal with by grey system theory" . This statement is essentially implies that small sample inference is an advantage aspect of grey system theory which is able to address the inherent defects of conventional, statistical methods. Almost all grey system related books and articles accepted this statement including our own early publications and quoted it as the necessary reason for the usages of grey system theory.

Furthermore, it is necessary to point out that the fuzzy mathematics can be able also used to deal with small sample inferences.

In recent years, Neural-Network Computation theory (Caudill, M., (1989) and applications have become mature, with which the small sample inferences can be also performed with high accuracy.

Table 2.3.2 Branches in statistics with small sample inference ability

Branches	bayesian	bootstrap	small sample asymptotic
Small size	small-large	small	small
feature	prior information	random number generation	saddle point approximation
strength	intuitive	implementation	solid
weakness	complicated	not timeouts	complicated
distribution	not needed	not needed	not needed

Now it is necessary to discuss the concept of grey uncertainty. As a general research practice, grey researchers have being tried to investigate the uncertainty associated with grey system which is caused by small sample size. Grey researchers identified such uncertainty as grey uncertainty and regarded it as being fundamental and parallel to random uncertainty etc. It is inevitable to raise a question here: is the grey uncertainty an essential uncertainty or a phenomenal uncertainty? Here, an essential uncertainty is defined as the one which can not be decomposed further and should be characterized rigorously by a set of axioms. For example, random uncertainty is an essential uncertainty and it is characterized by a set of axioms of probability measure. Another essential uncertainty is fuzzy uncertainty, which is characterized by five axioms of fuzzy credibility measure (Liu, B. and Liu, Y. K. (2002)).

The classification of an essential uncertainty or a phenomenal uncertainty is again a question based on a mathematical rigor. The current well-accepted essential uncertainties with rigorous mathematical foundations are random uncertainty, which is established on Cantor Set theory, and fuzzy uncertainty, which is established on fuzzy set theory. Any other forms of uncertainties should be treated by one of the two essential uncertainties or both.

Since the proposal of grey system theory term by Deng (1985), grey researchers always believed a set-theoretical foundation was established by Deng in terms of haze set concept.

If we start with the haze set theory which Deng (1985) claimed as the set theoretic foundation for grey system theory. Then the theoretical structure should be similar to that of probabilistic system or the fuzzy system. **In other words, the haze set should be able to use for constructing a set theoretical foundation, on which the whole grey system theory is based on it. Fundamentally, the following elements of the haze set foundation should be defined: (1) The set class composed of some subsets of a given haze set H, denoted as H and may call it as**

"haze-algebra", (if it exists), contrasting to a σ-algebra in probability theory, or the power set 2^{Θ} in fuzzy credibility measure theory. The class of subsets of a haze set, haze-algebra, should be closed under to a group of haze set operations. (2) A normed set mapping from set class "Haze-algebra" to unit interval [0, 1]. The set function is denoted as ϖ, if exists. In other words, set function $\varpi : H \to [0,1]$ is supposed to be the grey measure for quantifying grey event uncertainty. (3) The triple, denoted as (H, H, ϖ) should be definable and called as grey measure space. (4) A grey variable concept should be defined. A grey variable is a mapping from grey measure space to real line, i.e. $\varsigma(\otimes) : (H, \mathrm{H}, \varpi) \to \Re$ (5) The grey distribution for a grey variable should be defined and therefore denoted as $F_{\varsigma(\otimes)}(x) = \varpi\{\varsigma(\otimes \le x)\}, x \in \Re$ (6) Fundamental concepts and laws on grey variable, for example, expectation, variance, entropy, etc, etc. (Guo, R. and Guo, D. (2007))

However, up to date, we did not see any solid establishments of the above-mentioned elements based on haze set settled in grey theory literature.

Therefore, grey uncertainty is a phenomenal one and may be to be treated by the theories of essential uncertainties.

Furthermore, the reason why we do not use term "grey system theory" in my thesis rather we use "grey mathematics" is that the so-called grey system theory is not qualified to be a mathematical theory with rigorous mathematical (particularly, a set-theoretical) foundation yet. However, we accept the fact that grey methodological development is a branch of approximation theory and therefore it should be called as grey mathematics.

2.4 Grey differential equation, GM (1, 1) and GM (2, 1) model

It is well-known fact that numerical methods for solving differential equation in terms of finite difference methods have been long developed. But Deng, J. L. (2002) approach is very different from the traditional numerical analysis. Professor Deng's differential equation modelling is engaging the estimation of the unknown parameters in the differential equation in terms of easily handled regression approach, and then supplies a solution (in theoretical form) to the differential equation, which is determined and fully specified by the estimated parameters, for describing the dynamic and further system analysis.

Although we no longer use the term "grey differential equation" and change to a term of "differential equation motivated regression model", abbreviated as DEMR model for revealing its intrinsic feature of this mathematical model. However, it is still worth to pay an effort to review what Professor Deng has contributed to GDE modelling. (R. Guo. D. Guo & C. Thiart (2007))

2.4.1 Grey Sequences Generation

Because there are only few typical distributions employed in probability theory and statistic processes, many times statistical laws will become inefficient when dealing with small sample estimation or a large quantity data with non-typical distribution. In grey mathematics, people apply a method to find out realistic governing laws from available data to reflect the state of whole data sequence. The method is called a generation of grey sequence.

Definition 2.4.1: Assume that $X^{(0)}$ is a raw data sequence, G is an operator for modify the sequence satisfied that:

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right) \tag{2.3}$$

And

$$X^{(0)}G = \left(x^{(0)}(1)g, x^{(0)}(2)g, \cdots, x^{(0)}(n)g\right) \tag{2.4}$$

Where

$$x^{(0)}(k)g = \sum_{i=1}^{k} x^{(0)}(i), \tag{2.5}$$

For k=1, 2...n. The operator g of the data sequence is called a (first-order) accumulating generator of $X^{(0)}$, denoted 1-AGO.

The inverse accumulating generation is often used after 1-AGO and grey prediction, plays a role as returning data, from the accumulated one to its original condition.

Example 2.4.1: The following sequence

$X^{(0)} = (1.25, 2.2, 1.65, 3.14, 2.87) = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, x_4^{(0)}, x_5^{(0)})$

(See Figure 2.4.1)

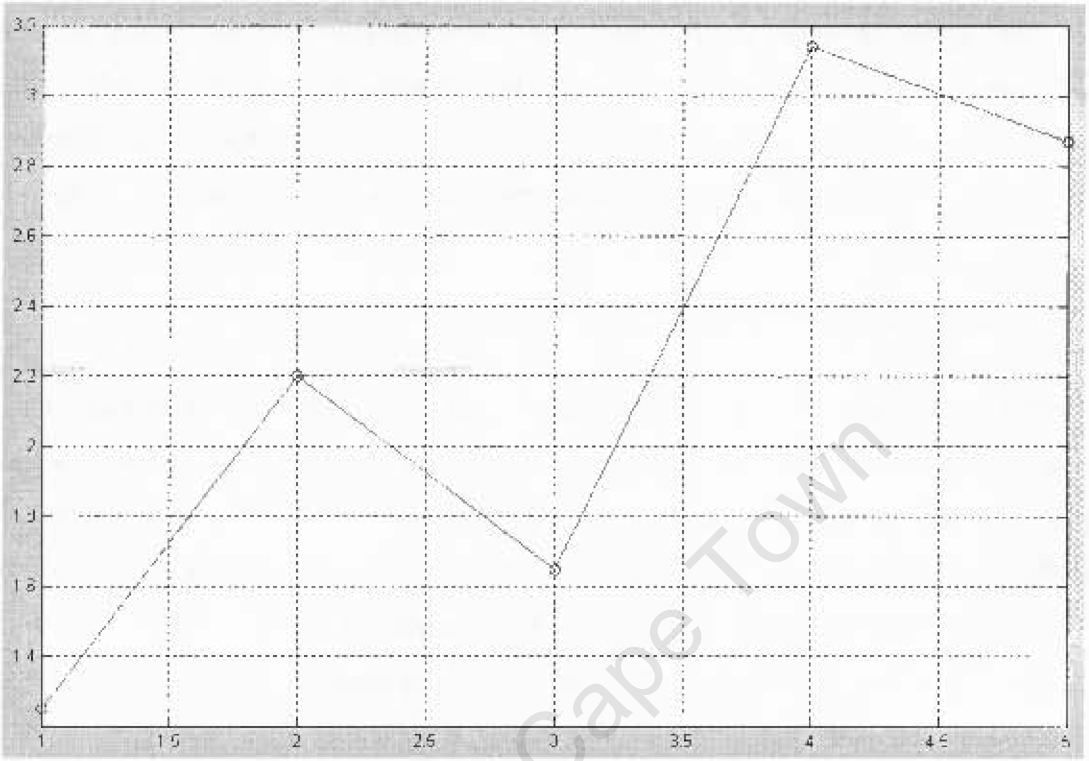

Figure 2.4.1 Illustration of the data sequence $X^{(0)}$

It doesn't show any regularity or pattern.

If we utilize accumulating generation once to the raw data sequence $X^{(0)}$, then it represent as:

$X^{(1)} = (1.25, 3.45, 5.1, 8.24, 11.11) = (x_1^{(0)}, x_1^{(0)} + x_2^{(0)}, x_1^{(0)} + x_2^{(0)} + x_3^{(0)}, x_1^{(0)} + x_2^{(0)} + x_3^{(0)} + x_4^{(0)},$
$x_1^{(0)} + x_2^{(0)} + x_3^{(0)} + x_4^{(0)} + x_5^{(0)}) = (x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, x_4^{(1)}, x_5^{(1)})$

Most of raw data sequence after accumulating generation once may transform the pattern from no regularity condition to a tendency of growth. This method occupied important position in grey system mathematics, which clearly draws out characteristics hidden in the chaotic raw data sequence. (See Figure 2.4.2)

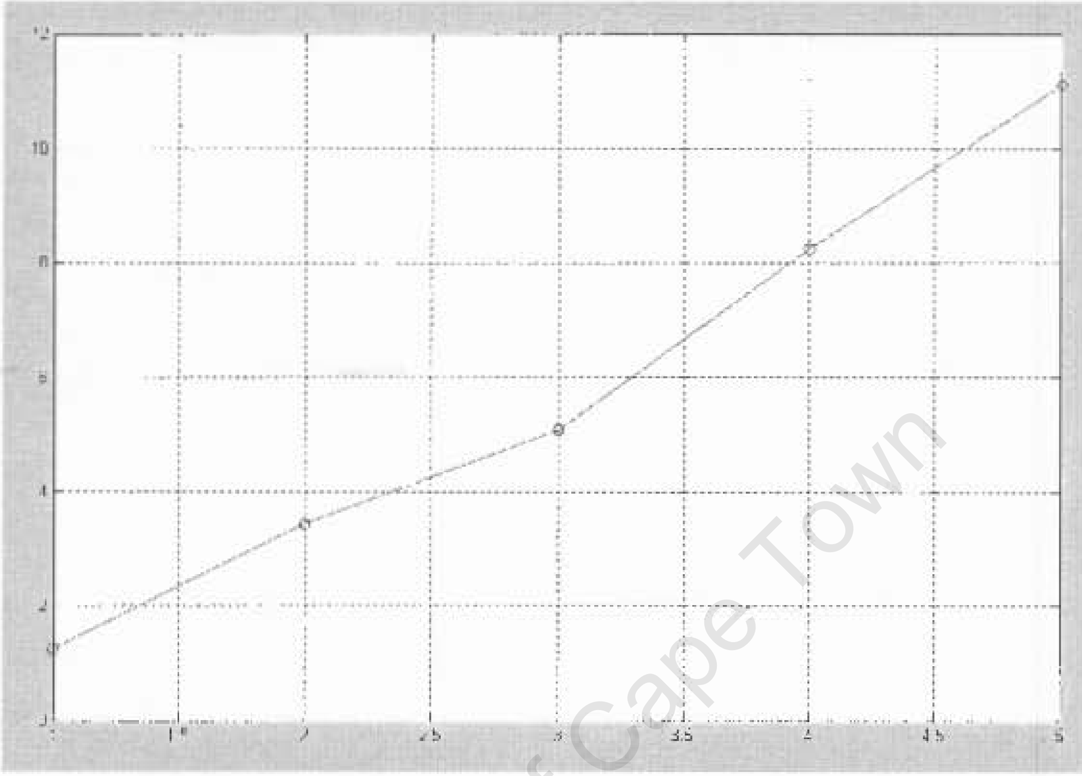

Figure 2.4.2 the curve represents an obvious tendency of growth

Definition 2.4.2: Assume that $X^{(0)}$ is a raw data sequence, G is an operator for modify the sequence satisfied that:

$$X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right) \tag{2.6}$$

And

$$X^{(0)}G = \left(x^{(0)}(1)g, x^{(0)}(2)g, \cdots, x^{(0)}(n)g\right) \tag{2.7}$$

Where

$$x^{(0)}(1)g = x^{(0)}(1) \tag{2.8}$$

$$x^{(0)}(k)g = x^{(0)}(k) - x^{(0)}(k-1), \tag{2.9}$$

For k=1, 2...n. The operator g of the data sequence is called a (first-order) Inverse accumulating generator of $X^{(0)}$, denoted 1-IAGO.

Without any doubts, AGO and IAGO operation played some active roles in GDE modelling because AGO will smoothen data and remove fluctuations in original data sequence and IAGO will provide the approximation to derivative. The success underlying AGO and IAGO roots in a mathematical fact that AGO is a partial sum which approximates an integral and the primitive function and also the IAGO is a difference which approximates a derivative of a continuous function.

As a matter of fact, the creation of AGO and IAGO did not truly help the development and acceptance of the grey mathematics, rather they confused many traditional mathematicians, statisticians and engineers. AGO is partial sum and IAGO is difference and these terms are used over hundred years already. Furthermore, the usage of AGO and IAGO hided their basic mathematical properties – They are linear transformations and the transformation will not change the model accuracy in its original scale of sample data.

2.4.2 Smooth discrete sequence

Grey differential equation model (abbreviate as GM) is usually dealing with discrete data sequence, but one major characteristic of smooth continuous functions is being differentiable everywhere, its means we can't use derivatives and related methods to study the smoothness of discrete data term in here. From the numerical approximation theory, we could use accumulating sum computation dealing with discrete data term to substitute the integral of the continue function, and then the result term could possess similar characteristics as the integral one. The smooth sequence idea will through accumulating sum of the discrete data term to draw out whether they possess the similar smooth characteristics as the continue functions. If they have, then we will be treated them as being smooth.

Definition 2.4.3: let a positive discrete data $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$ be a sequence, then

$$\sigma(k) = \frac{x(k)}{x(k-1)} \tag{2.10}$$

$k = 2,3,\cdots,n,$ is called stepwise ratios of the sequence $X^{(0)}$,and,

$$\rho(k) = \frac{x(k)}{\sum_{i=1}^{k-1} x(i)} \tag{2.11}$$

$k = 2,3,\cdots,n,$ is called smooth ratios of the sequence $X^{(0)}$.

When $k \geq 2$,

$$\frac{\rho(k+1)}{\rho(k)} < 1 \text{ and } \sigma(k) < 2 \tag{2.12}$$

When $k \geq 3$,

$$0 \leq \rho(k) < 0.5 \tag{2.13}$$

If k is infinite large, $\rho(k)$ will approximate to 0.

Then $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$ is said to be a quasi-smooth sequence. (Liu and Lin 1998)

From the example 2.4.1 we know, once the accumulating procedure is applied, the generated data sequence would show up an exponential tendency. And it is quite easy for people simulate an exponential function.

In general, if a discrete data sequence $X^{(0)}$ satisfied the quasi-smoothness conditions, we can easily following the grey mathematics modelling method to build a grey model step by step.

On the other hand, if the discrete data isn't quasi-smooth enough, we should take the accumulating generation procedure over and over again till the sequence satisfied the quasi-smooth data sequence condition.

2.4.3 Grey differential equation

Differential equation is quite useful in various systems research areas, because it can easily draw out the essence from the development things. Unfortunately, Grey system mathematics most deal with discrete data sequence, based on that condition, differential equation might be arduous to perform its working efficiency. Only the condition of differentiability is assumed, people could handle the grey mathematics problem make use of the differential equations. After the introduction of the concept of smooth discrete data sequence previously, we can easily establish models similar to differentiable equations for discrete sample sequence using.

Definition 2.4.4: If an equation follows the form

$$d^{(i)}(k_i) + ax^{(1)}(k_i) = b \tag{2.14}$$

It is called an equation of grey differential type.

Proposition 2.4.1: For the following equation of grey differential type

$$x^{(0)}(k) + \alpha x^{(1)}(k) = b \tag{2.15}$$

The grey derivative $x^{(0)}(k)$ and elements in the set of background values

$$\{x^{(1)}(k), x^{(1)}(k-1)\} \tag{2.16}$$

do not satisfy the horizontal mapping relation.

Proposition 2.4.2: If the background value is taken to be the mean of the entries in $X^{(1)}$, that is let

$z^{(1)}(k) = 0.5x^{(1)}(k) + 0.5x^{(1)}(k-1)$,

Then the background value $z^{(1)}(k)$ and the components $x^{(1)}(k)$ and $x^{(1)}(k-1)$ of the grey derivative satisfy the arithmetic horizontal mapping relation.

Definition 2.4.5: If an equation of grey differential type satisfies the following conditions,

7. the information density is infinitely large;
8. the sequence possesses the intension of grey differentiation; and
9. the mapping from the set of background values to the components of the grey derivative satisfy the horizontal mapping relation,

then this equation of grey differential type is called a grey differential equation.

Definition 2.4.6: Given a discrete positive real-valued data sequence $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$ satisfy the equation

$$x^{(0)}(k) + az^{(1)}(k) = b, k = 2,3,4,\cdots,n \quad (GM(1,1)) \tag{2.17}$$

Where

$$z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right] \tag{2.18}$$

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4,\cdots,n \tag{2.19}$$

Then equation 2.15 is called a grey differential equation and equation 2.17 is called GM (1, 1) model.

2.4.4 GM (n, h) model

GM (n, h) model standard for h different variables with construction of n^{th} order differential equations represent.

Definition 2.4.7: Consider about we have h different N-dimensional data sequence, denote as:

$$\left\{X_K^{(0)}(i) \middle| k=1,2,\cdots,h; i=1,2,\cdots,N\right\} \tag{2.20}$$

The 1-time Accumulating Generation Operation (abbreviate as 1-AGO) sequence of $\left\{X_K^{(0)}(i) \middle| k=1,2,\cdots,h; i=1,2,\cdots,N\right\}$ is:

$$\left\{X_K^{(1)}(i) \middle| k=1,2,\cdots,h; i=1,2,\cdots,N\right\} \tag{2.21}$$

Where:

$$X_K^{(1)}(i) = \sum_{i=1}^{N} X_K^{(0)}(i),\ i=1,2,\cdots,N; k=1,2,\cdots,h; \tag{2.22}$$

The Inverse Accumulating Generation Operation (abbreviate as IAGO) sequence is:

$$\left\{\alpha^{(j)}(X_k, i) \middle| k=1,2,\cdots,h; i=1,2,\cdots,N; j=1,2,\cdots,l\right\} \tag{2.23}$$

$$\alpha^{(1)}(X_k, i) = X^{(0)}(t_i) \tag{2.24}$$

$$\alpha^{(2)}(X_k, i) = X^{(0)}(t_i) - X^{(0)}(t_{i-1}) \tag{2.25}$$

$$\alpha^{(j)}(X_k, i) = \alpha^{(j-1)}(X_k, i) - \alpha^{(j-1)}(X_k, i-1) \tag{2.26}$$

Above all, the GM (n, h) model is:

$$\sum_{i=0}^{n} \alpha_i \frac{d^{n-1}\hat{X}_1^{(1)}}{dt^{n-1}} = \sum_{i=1}^{n-1} b_i \hat{X}_{i+1}^{(1)} \tag{2.27}$$

Where

$$X_1^{(0)}(0) = X_1^{(1)}(1) \tag{2.28}$$

$$X_k^{(1)}(i) = \sum_{i=1}^{N} X_k^{(0)}(i), i=1,2,\cdots,N; k=1,2,\cdots,h; \tag{2.29}$$

2.4.5 GM (1, 1) model

The firs- order and one-variable grey differential equation model, abbreviate as GM (1, 1) is a particular case of GM (n, h) when n=1 and h=1. In GM (n, h) model when $h \geq 2$, we can't using it to do predictions, only can using for relationship analysis between two elements. In general, we normally utilize GM (n, 1) model for data analysis and prediction, but in the real activities of science research, most of the applications about grey mathematics is based on the GM (1, 1) model, in other words, the GM (1, 1) model is the most important part of Grey modelling.

The most especially characteristic of the GM (1, 1) model is the high predictive power of small-sample analysis, requires very small sample size, as little as four sample points, but the result possibly will achieve very high predictability accuracy. Another chiefly characteristic of GM (1, 1) is easy for people modelling and manipulate. The modelling of GM (1, 1) is just an undemanding job need to estimate the parameters utilize a simple regression, which could easy been solved by using a Microsoft Excel spreadsheet with basic computations. The comprehension needed for GM (1, 1) modelling is just some fundamental knowledge of statistics, which is offered in the beginning of any statistics or data analysis course, easy for people to understanding and utilizing for applications.

Definition 2.4.8: Let $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$ be a non-negative discrete data sequence, and take 1-time accumulating generation generate data sequence $X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right)$ where

$$X^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4,\cdots,n \tag{2.30}$$

And $Z^{(1)}$ is the mean generated sequence of consecutive neighbours of $X^{(1)}$ given by

$$Z^{(1)} = \left(z^{(1)}(1), z^{(1)}(2), \cdots, z^{(1)}(n)\right), \tag{2.31}$$

Where

$$Z^{(1)}(k) = \frac{1}{2}\left(x^{(1)}(k) + x^{(1)}(k-1)\right), k = 1,2,\cdots,n. \tag{2.32}$$

Then the equation

$$x^{(0)}(k) + az^{(1)}(k) = b, k = 2,3,4,\cdots,n \tag{2.33}$$

Hinted by the whitening differential equation

$$\frac{d\hat{x}^{(1)}(t)}{dt} + a\hat{x}^{(1)} = b \tag{2.34}$$

is called a GM (1, 1) model with respect to a positive data

$$X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)) \tag{2.35}$$

Sequence.

Definition 2.4.9: Let $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$ be a non-negative discrete data sequence, and take 1-time accumulating generation generate data sequence $X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right)$ where

$$X^{(1)}(k)=\sum_{i=1}^{k}x^{(0)}(i), k=2,3,4,\cdots,n \tag{2.36}$$

And $Z^{(1)}$ is the mean generated sequence of consecutive neighbours of $X^{(1)}$ given by

$$Z^{(1)}=\left(z^{(1)}(1),z^{(1)}(2),\cdots,z^{(1)}(n)\right), \tag{2.37}$$

Where

$$Z^{(1)}(k)=\frac{1}{2}\left(x^{(1)}(k)+x^{(1)}(k-1)\right), k=1,2,\cdots,n. \tag{2.38}$$

If $\hat{a}=[a,b]^T$ is a sequence of parameters, and

$$Y=\begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} \qquad B=\begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} \tag{2.39}$$

Then the least squares estimate sequence of the grey differential equation

$$x^{(0)}(k)+az^{(1)}(k)=b \tag{2.40}$$

Satisfies

$$\hat{a}=\left[B^T B\right]^{-1}B^T Y \tag{2.41}$$

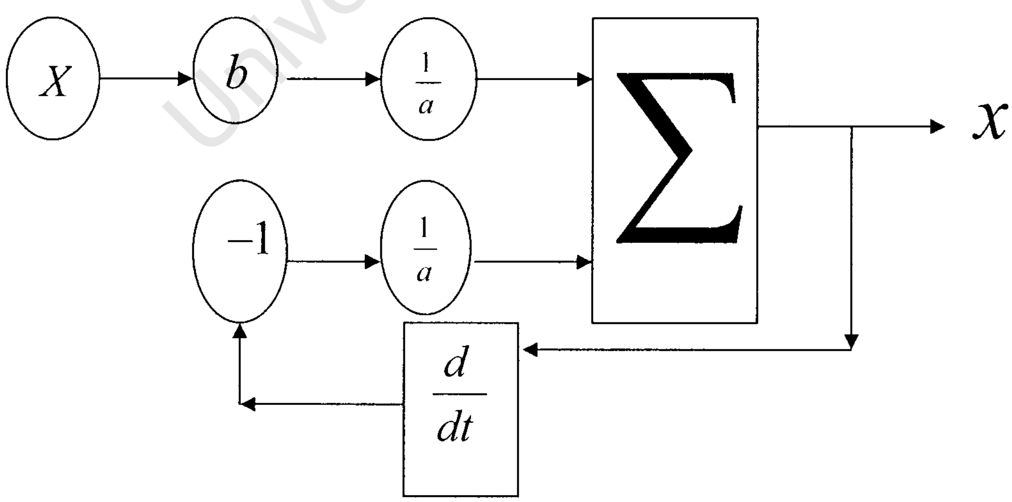

Figure 2.4.3 GM (1, 1) model concept

Definition 2.4.10: Let $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$ be positive discrete data sequences, $X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right)$ is 1-time accumulating generation operation generated from $X^{(0)}$, and $Z^{(1)}$ is the mean generated sequence of consecutive neighbours of $X^{(1)}$. If

$$[a, b]^T = \left[B^T B\right]^{-1} B^T Y \tag{2.42}$$

Then

$$\frac{dx^{(1)}}{dt} + ax^{(1)} = b \tag{2.43}$$

Is called the whitening differential equation or the shadow equation of the grey differential equation

$$\hat{x}^{(0)}(k) + az^{(1)}(k) = b. \tag{2.44}$$

The parameter a is called the developing coefficient, parameter b is the grey input, $x^{(0)}$ is called a grey derivative and $x^{(1)}(k)$ is called the k^{th} 1-time accumulating generation operation of $x^{(0)}$ (abbreviate as 1-AGO) value.

Theorem 2.4.1: The GM (1, 1) model

$$x^{(0)}(k) + az^{(1)}(k) = b \tag{2.45}$$

Can be transformed into

$$x^{(0)}(k) = \beta - \alpha x^{(1)}(k-1), \tag{2.46}$$

Where

$$\beta = \frac{b}{1+0.5a}, \alpha = \frac{a}{1+0.5a}. \tag{2.47}$$

Theorem: Assume that

$$\beta = \frac{b}{1+0.5a}, \alpha = \frac{a}{1+0.5a} \tag{2.48}$$

And

$$\hat{X}^{(1)} = \left(\hat{x}^{(1)}(1), \hat{x}^{(1)}(2), \cdots, \hat{x}^{(1)}(n)\right) \tag{2.49}$$

Is the time response sequence of the GM (1, 1) model, where

$$\hat{x}^{(1)}(k) = \left[x^{(0)}(1) - \frac{b}{a}\right] e^{-a(k-1)} + \frac{b}{a}, k = 1,2,\cdots,n. \tag{2.50}$$

Then

$$\hat{x}^{(0)}(k) = \left[\beta - \alpha x^{(0)}(1)\right]e^{-a(k-2)} \tag{2.51}$$

2.4.6 GM (1, 1) modelling procedure

In this section we will give an example following GM (1, 1) modelling procedure to introduce how to modelling a GM (1, 1) model step by step.

Definition 2.4.11: Let $X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$ be positive discrete data sequences, $\hat{X}^{(0)} = \left(\hat{x}^{(0)}(1), \hat{x}^{(0)}(2), \cdots, \hat{x}^{(0)}(n)\right)$ is estimation data sequence generated from $X^{(0)}$, Errors of GM (1, 1) model represents as $\varepsilon(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$. Relative Errors of GM (1, 1) notates as $\Delta_K = \dfrac{|\varepsilon(k)|}{x^{(0)}(k)}$.

Example 2.4.2: Let $X^{(0)} = (2.874, 3.278, 3.337, 3.390, 3.679) = (x_1^{(0)}, x_2^{(0)}, x_3^{(0)}, x_4^{(0)}, x_5^{(0)})$, we strictly obey the GM (1, 1) modelling procedure law to simulate $X^{(0)}$, and evaluate the accuracy of $X^{(0)}$. (Liu, S., and Lin, Y. (2006))

Step1. At the beginning we first take 1-time accumulating generation operation of $X^{(0)}$, to generate the data sequence $X^{(1)}$.

$$X^{(1)} = (x_1^{(1)}, x_2^{(1)}, x_3^{(1)}, x_4^{(1)}, x_5^{(1)}) = (2.874, 6.152, 9.489, 12.879, 16.558) \tag{2.52}$$

Step2. Check raw data sequence $X^{(0)}$ whether or not satisfy the quasi-smoothness condition $\dfrac{\rho(k+1)}{\rho(k)} < 1$ and when k>3, $0 \le \rho(k) < 0.5$; Check data sequence $X^{(1)}$ whether or not satisfy the quasi-smoothness condition $\sigma^{(1)}(k) < 2$.

$$\rho(3) = \frac{x^{(0)}(3)}{x^{(1)}(2)} = \frac{3.337}{6.152} \approx 0.54 \tag{2.53}$$

$$\rho(4) = \frac{x^{(0)}(4)}{x^{(1)}(3)} = \frac{3.390}{9.489} \approx 0.36 \tag{2.54}$$

$$\rho(5) = \frac{x^{(0)}(5)}{x^{(1)}(4)} = \frac{3.679}{12.879} \approx 0.29 \tag{2.55}$$

$\dfrac{\rho(4)}{\rho(3)} = \dfrac{0.36}{0.54} < 1$ $\dfrac{\rho(5)}{\rho(4)} = \dfrac{0.29}{0.36} < 1$ When k>3, $\rho(4) < 0.5$ $\rho(5) < 0.5$

$$\sigma^{(1)}(3)=\frac{x^{(1)}(3)}{x^{(1)}(2)}=\frac{9.489}{6.152}\approx 1.54, \tag{2.56}$$

$$\sigma^{(1)}(4)=\frac{x^{(1)}(4)}{x^{(1)}(3)}=\frac{12.879}{9.489}\approx 1.36 \tag{2.57}$$

$$\sigma^{(1)}(5)=\frac{x^{(1)}(5)}{x^{(1)}(4)}=\frac{16.558}{12.879}\approx 1.29 \tag{2.58}$$

So, the condition of being quasi-smoothness is satisfied

Step3. Generate $z^{(1)}$ data sequence from $z^{(1)}=\frac{1}{2}\left(x^{(1)}(k)+x^{(1)}(k-1)\right)$

$$Z^{(1)}=\left(z^{(1)}(1),z^{(1)}(2),z^{(1)}(3),z^{(1)}(4),z^{(1)}(5)\right)=\left(2.874,4.513,7.820,11.184,14.718\right) \tag{2.59}$$

Step4. Utilizing least squares estimate the parameter sequence $\hat{a}=[a,b]^T$

$$\hat{a}=\left[B^TB\right]^{-1}B^TY \tag{2.60}$$

$$B=\begin{bmatrix}-z^{(1)}(2) & 1\\ -z^{(1)}(3) & 1\\ -z^{(1)}(4) & 1\\ -z^{(1)}(5) & 1\end{bmatrix}=\begin{bmatrix}-4.513 & 1\\ -7.820 & 1\\ -11.184 & 1\\ -14.718 & 1\end{bmatrix},\ Y=\begin{bmatrix}x^{(0)}(2)\\ x^{(0)}(3)\\ x^{(0)}(4)\\ x^{(0)}(5)\end{bmatrix}=\begin{bmatrix}3.278\\ 3.337\\ 3.390\\ 3.679\end{bmatrix}; \tag{2.61}$$

$$B^TB=\begin{bmatrix}-4.513 & 1\\ -7.820 & 1\\ -11.184 & 1\\ -14.718 & 1\end{bmatrix}^T\bullet\begin{bmatrix}-4.513 & 1\\ -7.820 & 1\\ -11.184 & 1\\ -14.718 & 1\end{bmatrix}=\begin{bmatrix}423.244 & -38.236\\ -38.236 & 4\end{bmatrix}. \tag{2.62}$$

$$[B^TB]^{-1}=\begin{bmatrix}423.244 & -38.236\\ -38.236 & 4\end{bmatrix}^{-1}=\begin{bmatrix}0.017317 & 0.165537\\ 0.165537 & 1.83236\end{bmatrix} \tag{2.63}$$

$$Y=\begin{bmatrix}x^{(0)}(2)\\ x^{(0)}(3)\\ x^{(0)}(4)\\ x^{(0)}(5)\end{bmatrix}=\begin{bmatrix}3.278\\ 3.337\\ 3.390\\ 3.679\end{bmatrix}; \tag{2.64}$$

$$\hat{a}=\left[B^T B\right]^{-1} B^T Y=\begin{bmatrix}0.017317 & 0.165537\\ 0.165537 & 1.83236\end{bmatrix}\cdot\begin{bmatrix}-4.513 & 1\\ -7.820 & 1\\ -11.184 & 1\\ -14.718 & 1\end{bmatrix}^T\cdot\begin{bmatrix}3.278\\ 3.337\\ 3.390\\ 3.679\end{bmatrix}$$

$$\begin{bmatrix}0.087385 & 1.085292\\ 0.030118 & 0.537861\\ -0.028136 & -0.019006\\ -0.089335 & -0.604014\end{bmatrix}^{-T}\cdot\begin{bmatrix}3.278\\ 3.337\\ 3.390\\ 3.679\end{bmatrix}=\left[-0.0372044, 3.06536\right]^T. \tag{2.65}$$

Step5: Generate the grey model and the time response sequence

$$\frac{d\hat{x}^{(1)}}{dt}-0.0372\hat{x}^{(1)}=3.06536 \tag{2.66}$$

$$\hat{x}^{(1)}(k+1)=\left[x^{(0)}(1)-\frac{b}{a}\right]e^{-ak}+\frac{b}{a}-85.276151\cdot e^{-ak}+\frac{b}{a}=85.276e^{0.0372k}-82.4 \tag{2.67}$$

Step6: Simulate value of $X^{(1)}$ and estimation value of $X^{(0)}$

$$\hat{X}^{(1)}=(2.874, 6.10604, 9.46059, 12.9423, 16.556) \tag{2.68}$$

$$\hat{x}^{(0)}(k)=\hat{x}^{(1)}(k)-\hat{x}^{(1)}(k-1), \tag{2.69}$$

$$\hat{X}^{(0)}=(2.874, 3.23204, 3.35455, 3.4817, 3.61368). \tag{2.70}$$

Step7. Calculate the Errors and Relative Errors

$$\delta(k)=x^{(0)}(k)-\hat{x}^{(0)}(k)=(0.0459611,\ 0.0175498, -0.0917044, 0.0653211) \tag{2.71}$$

Relative Errors (%):

$$\Delta k=\frac{|\delta(k)|}{x^{(0)}(k)}=(1.40211\%, 0.525914\%, 2.70514\%, 1.77551\%) \tag{2.72}$$

Average relative error:

$$\Delta=\frac{1}{4}\sum_{k=2}^{5}\Delta_K=0.016217=1.60217\% \tag{2.73}$$

Accuracy of GM (1, 1) model

$$accuracy=1-\Delta=1-\frac{1}{4}\sum_{k=2}^{5}\Delta_K=0.983783=98.3783\% \tag{2.74}$$

In here we utilize visual C++ language following the Grey modelling principle to program the solution of the entire questions. (See Figure 2.4.4)

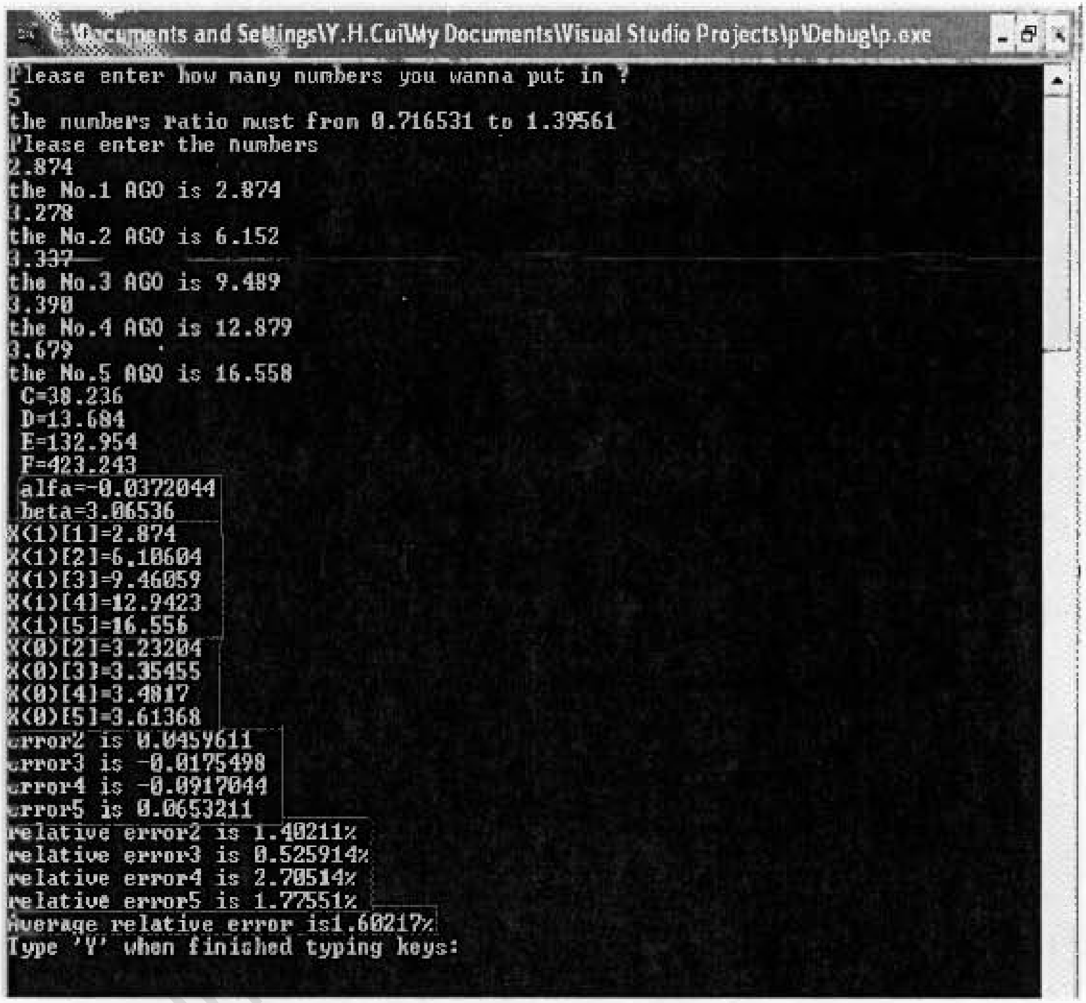

Figure 2.4.4 Visual C++ program GM (1, 1) modelling result

2.4.7 GM (1, 1) model prediction

At the beginning we emphasize, collecting enough information for data analysis is quite inappropriate with today's competitive business environment. So small sample analysis becomes much useful and necessary in today's lots of different academic research area with numerous applications. Actually, the entire GM (n, 1) model has their predictability power could use for data sequence prediction, but the most common model used in grey mathematics application for prediction data sequence is GM (1, 1) model.

Example2.4.3: Let $X^{(0)} = (2.874, 3.278, 3.337, 3.390, 3.679) = (X_1^{(0)}, X_2^{(0)}, X_3^{(0)}, X_4^{(0)}, X_5^{(0)})$, using GM (1, 1) model to predict the value of $(X_6^{(0)}, X_7^{(0)}, X_8^{(0)}, X_9^{(0)}, X_{10}^{(0)})$. The 1-AGO sequence of $X^{(0)}$ is $X^{(1)} = (X_1^{(1)}, X_2^{(1)}, X_3^{(1)}, X_4^{(1)}, X_5^{(1)}) = (2.874, 6.152, 9.489, 12.879, 16.558)$

Assume

$$\frac{dx^{(1)}(t)}{dt} + ax^{(1)} = b \tag{2.74}$$

We evaluate the estimation values for a and b from the least squares method

$$[\hat{a}, \hat{b}]^T = \begin{bmatrix} -0.0372044 \\ 3.06536 \end{bmatrix}. \tag{2.75}$$

The whitening differential equation of GM (1, 1) is

$$\frac{d\hat{x}^{(1)}}{dt} - 0.0372\hat{x}^{(1)} = 3.06536 \tag{2.76}$$

With its time response sequence

$$\begin{cases} \hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1), \\ \hat{x}^{(1)}(k+1) == 85.276e^{0.0372k} - 82.4 \end{cases} \tag{2.77}$$

The predicted result is:

$$\hat{X}^{(0)} = \left(\hat{x}^{(0)}(i)\right)_{i=6}^{9} = (3.75066, 3.89282, 4.04038, 4.19353) \tag{2.78}$$

2.4.8 GM (2, 1) model

GM (2, 1) model is quite different with GM (1, 1) model, which usually apply to describe a non-monotonic data sequence and utilize second order differential equation as its whitening equation.

Definition 2.4.11: Let positive data sequence $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$, after 1-AGO operation, $X^{(1)}$ is

$$X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right), \tag{2.79}$$

Where

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 1, 2, \cdots, n, \tag{2.80}$$

And the 1-IAGO sequence $X^{(-1)}$ of $X^{(0)}$ is

$$X^{(-1)} = \left(x^{(-1)}(1), x^{(-1)}(2), \cdots, x^{(-1)}(n)\right), \tag{2.81}$$

Where

$$x^{(-1)}(k)=x^{(0)}(k)-x^{(0)}(k-1), k=1,2,\cdots,n \tag{2.82}$$

Then the sequence of mean generated of consecutive neighbours of $X^{(1)}$ is

$$Z^{(1)}=\left(z^{(1)}(1),z^{(1)}(2),\cdots,z^{(1)}(n)\right), \tag{2.83}$$

Where

$$z^{(1)}(k)=\frac{1}{2}\left[x^{(1)}(k)+x^{(1)}(k-1)\right], k=1,2,\cdots,n \tag{2.84}$$

Now we have

$$X^{(-1)}+a_1X^{(0)}+a_2Z^{(1)}=b \tag{2.85}$$

Is called a GM (2, 1) grey differential equation.

Definition 2.4.12: we defined the equation

$$\frac{d^2\hat{x}^{(1)}}{dt}+a_1\frac{d\hat{x}^{(1)}}{dt}+a_2\hat{x}^{(1)}=b \tag{2.86}$$

as the whiting equation of GM (2, 1) grey differential equation.

Theorem 2.4.2: Assume that $X^{(0)}, X^{(1)}, Z^{(1)}$, and $\alpha^{(1)}X^{(0)}$ are defined the same way as in Definition 2.4.10, and

$$B=\begin{bmatrix}-x^{(0)}(2) & -z^{(1)}(2) & 1\\ -x^{(0)}(3) & -z^{(1)}(3) & 1\\ \cdots & \cdots & \cdots\\ -x^{(0)}(n) & -z^{(1)}(n) & 1\end{bmatrix} \tag{2.87}$$

$$Y=\begin{bmatrix}\alpha^{(1)}x^{(0)}(2)\\ \alpha^{(1)}x^{(0)}(3)\\ \cdots\\ \alpha^{(1)}x^{(0)}(n)\end{bmatrix}=\begin{bmatrix}x^{(0)}(2)-x^{(0)}(1)\\ x^{(0)}(3)-x^{(0)}(2)\\ \cdots\\ x^{(0)}(n)-x^{(0)}(n-1)\end{bmatrix}. \tag{2.88}$$

,

Then the least squares estimate of the GM (2, 1) parameter sequence

$$\hat{a} = [a_1 \quad a_2 \quad b]^T \tag{2.89}$$

Is given by

$$\hat{a} = [B^T B]^{-1} B^T Y \tag{2.90}$$

Theorem 2.4.3: As for the solution of the GM (2, 1) whitenization equation, the following hold true.

1. If $X^{(1)*}$ is a special solution of

$$\frac{d^2\hat{x}^{(1)}}{dt} + a_1 \frac{d\hat{x}^{(1)}}{dt} + a_2\hat{x}^{(1)} = b \tag{2.91}$$

And $\overline{X}^{(1)}$ a general solution of the homogeneous equation

$$\frac{d^2\hat{x}^{(1)}}{dt} + a_1 \frac{d\hat{x}^{(1)}}{dt} + a_2\hat{x}^{(1)} = 0, \tag{2.92}$$

Then $X^{(1)*} + \overline{X}^{(1)}$ is the general solution of the GM (2, 1) whitenization equation.

2. There are the following three cases for the general solution of the homogeneous equation above.

(a) When the characteristic equation

$$r^2 + a_1 r + a_2 = 0 \tag{2.93}$$

has two distinct real solution r_1 and r_2,

$$\overline{X}^{(1)} = C_1 e^{r_1 t} + C_2 e^{r_2 t} \tag{2.94}$$

(b) When the characteristic equation

$$r^2 + a_1 r + a_2 = 0 \tag{2.95}$$

has a real solution r of multiplicity 2,

$$\overline{X}^{(1)} = e^{rt}(C_1 + C_2 t). \tag{2.96}$$

(c) When the characteristic equation

$$r^2 + a_1 r + a_2 = 0 \tag{2.97}$$

has two complex conjugate solutions $r_1 = \alpha + i\beta, r_2 = \alpha - i\beta$,

$$\overline{X}^{(1)} = e^{\alpha t}\left(C_1 \cos \beta t + C_2 \sin \beta t\right) \tag{2.98}$$

3. There exist three possibilities for a special solution of a whitenization equation:

(a) When zero is not a solution of the characteristic equation,

$$X^{(1)*} = C; \tag{2.99}$$

(b) When zero is a solution of multiplicity 1 of the characteristic equation,

$$X^{(1)*} = Cx; \tag{2.100}$$

(c) When zero is a multiplicate solution of the characteristic equation,

$$X^{(1)*} = Cx^2; \tag{2.101}$$

Example2.4.4: let positive discrete data sequence $X^{(0)} = (2.874, 3.278, 3.337, 3.390, 3.679)$. Then we utilize GM (2, 1) model to calculate the estimate value $\hat{X}^{(0)}$ of $X^{(0)}$.

Step1. Form 1-AGO and 1-IAGO operation to generate $X^{(-1)}$ and $X^{(1)}$ sequence of $X^{(0)}$. Generate $Z^{(1)}$ from equation (2.81).

$$X^{(1)} = \sum_{i=1}^{k} x^{(0)}(i) = (2.874, 6.152, 9.489, 12.879, 16.558) \tag{2.102}$$

$$X^{(-1)}(k) = x^{(0)}(k) - x^{(0)}(k-1) = (0, 0.404, 0.059, 0.053, 0.289) \tag{2.103}$$

And

$$Z^{(1)} = (2.874, 4.513, 7.820, 11.184, 14.7185) \tag{2.104}$$

Step2: using least-square regression method to estimate parameters $\hat{a} = [a_1, a_2, b]^T$

$$B = \begin{bmatrix} -x^{(0)}(2) & -z^{(1)}(2) & 1 \\ -x^{(0)}(3) & -z^{(1)}(3) & 1 \\ -x^{(0)}(4) & -z^{(1)}(4) & 1 \\ -x^{(0)}(5) & -z^{(1)}(5) & 1 \end{bmatrix} = \begin{bmatrix} -3.287 & -4.513 & 1 \\ -3.337 & -7.820 & 1 \\ -3.390 & -11.184 & 1 \\ -3.679 & -14.7185 & 1 \end{bmatrix} \tag{2.105}$$

And

$$Y = \left[x^{(-1)}(i)\right]^T = [0.404, 0.059, 0.053, 0.289]^T \tag{2.106}$$

Then we use least-square method to estimate the parameters

$$\hat{a} = [a_1, a_2, b]^T = \left[B^T B\right]^{-1} B^T Y = [30.48, -1.04, 92.90]^T \tag{2.107}$$

Step3: from the estimated parameters we could obtain the whiting equation

$$\frac{d^2 \hat{x}^{(1)}}{dt} + 30.48 \frac{d\hat{x}^{(1)}}{dt} - 1.04 x^{(1)} = 92.90 \tag{2.108}$$

The general solution of homogeneous equation

$$\frac{d^2 \hat{x}_h^{(1)}}{dt} + 30.48 \frac{d\hat{x}_h^{(1)}}{dt} - 1.04 \hat{x}_h^{(1)} = 0 \tag{2.109}$$

Is

$$\hat{X}_h^{(1)}(t) = C_1 e^{0.0341t} + C_2 e^{-30.514t} \tag{2.110}$$

The solution of characteristic equation

$$-1.04 \hat{x}_p^{(1)}(t) = 92.90 \tag{2.111}$$

Is

$$\hat{X}_p^{(1)}(t) = -\frac{92.9}{1.04} = -89.3269 \tag{2.112}$$

Therefore, we have the solution of whiting differential equation

$$\hat{X}^{(1)}(t) = \hat{X}_h^{(1)} + \hat{X}_p^{(1)} = C_1 e^{0.0341t} + C_2 e^{-30.514t} - 89.3269 \tag{2.113}$$

From

$$\hat{x}^{(1)}(t)\Big|_{t=0} = x^{(1)}(0) = x^{(0)}(1) = 2.874 \tag{2.114}$$

And

$$\left.\frac{d\hat{x}^{(1)}}{dt}\right|_{t=0} = x^{(0)}(t)\Big|_{t=0} = x^{(0)}(0) = 2.643 \tag{2. 115}$$

We could obtain the equation system

$$\begin{cases} 2.874 = C_1 + C_2 - 89.3269 \\ 2.643 = 0.0341 C_1 - 30.514 C_2 \end{cases} \tag{2.116}$$

So, it follows that

$$C_1 = 92.107983, C_2 = 2.931917 \tag{2.117}$$

Then we have

$$\hat{X}^{(1)}(t+1) = 92.107983 e^{0.0341t} + 2.931917 e^{-30.514t} - 89.3269 \tag{2.118}$$

From the equation (2.93) we obtain

$$\hat{X}^{(1)} = (2.8740, 5.9761, 9.2820, 12.7026, 16.2418) \quad (2.119)$$

After IAGO operation

$$x^{(-1)}(k) = \hat{x}^{(0)}(k) - \hat{x}^{(0)}(k-1), k = 1,2,\cdots,n \quad (2.120)$$

We have

$$\hat{X}^{(0)} = (2.8740, 3.1021, 3.3059, 3.4206, 3.5392) \quad (2.121)$$

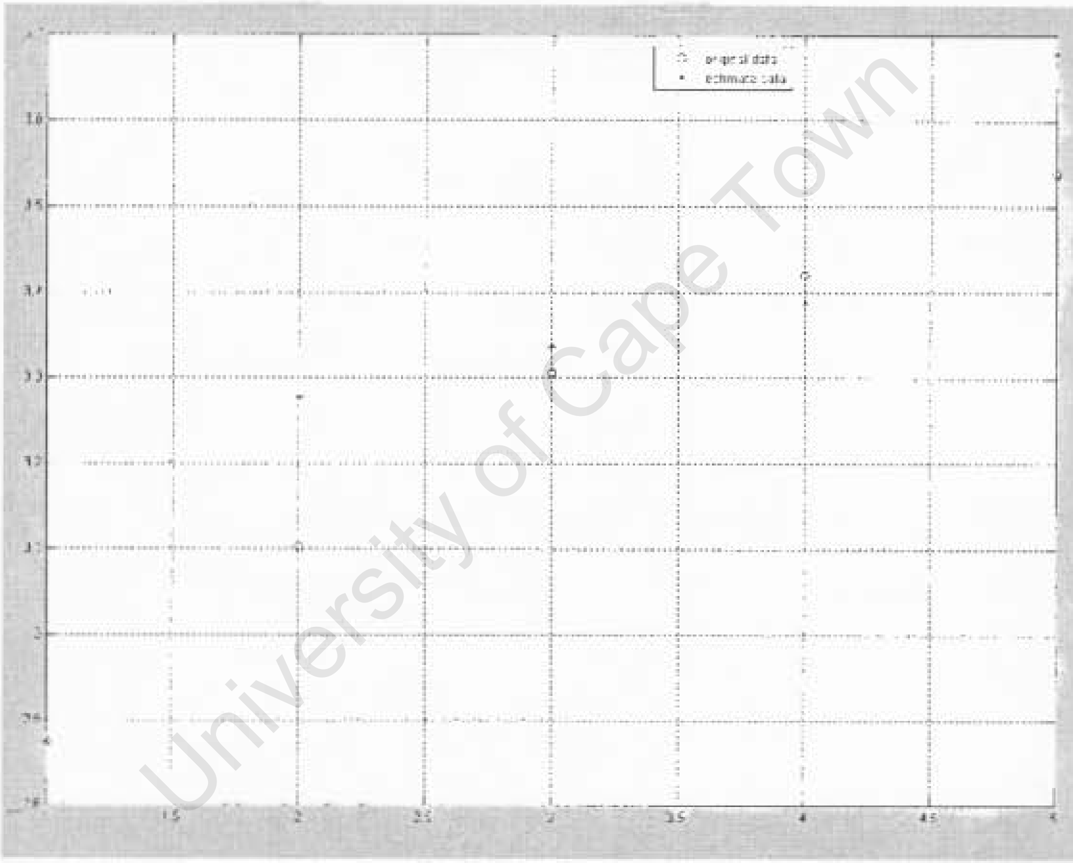

Figure 2.4.5 compares $X^{(0)}$ data sequence with $\hat{X}^{(0)}$ data sequence

The accuracy of GM (2, 1) model of $X^{(0)}$ is

$$accuracy = 1 - \frac{\sum_{i=1}^{k} \left| \frac{x^{(0)}(k) - \hat{x}^{(0)}(k)}{x^{(0)}(k)} \right|}{k-1} = 0.9725 = 97.25\% \quad (2.122)$$

(Modified from Liu and Lin, 2004)

2.5 Summary

After reviewed all fundamental knowledge of grey mathematics in chapter 1, we simply abstract some characteristics from GM (1, 1) model and listed in here:

First, GM (1, 1) model usually dealing with small sample data analysis, as little as 4 sample points needed. Second, GM (1, 1) model have a great power of prediction offered with high accuracy. Third, the modelling procedure and computations is quite easy; all the computation of GM (1, 1) model could be facilitate by a simple Microsoft Excel spreadsheet. Even such lot of advantage of characteristic GM (1, 1) model has, the shortage of information clarity (information incomplete) and lack of mathematical theory foundation still a barrier for people who familiar with statistical analysis to fully accept the whole GM (1, 1) modelling idea. In next chapter we will use statistical method to redefine the definition of GM (1, 1).

I don't pay more attention to GM (2, 1) model in my recently research, so we just introduced some basic knowledge about GM (2, 1) in this thesis.

Chapter 3. Statistical-Grey Consistency

3.1 Statistical-Grey Consistency Theories

Even we realized the importance of regression model in GM (1, 1) modelling, but for common practice, we seemed like ignoring the goodness-of-fit information of the simple regression model itself. The reason for ignoring the goodness-of-fit information in some sense may caused by the sample size extreme small problem. If GM (1, 1) asked for very small sample analysis (such as only 5 data been given), so we can't ensure the small sample actually followed with which distribution. So from the statistical theory, we can't do any hypothesis testing or analysis of variance of it. Now, the only information we could investigate of statistics for very small sample problem is the coefficient of determination (R-square value). If the R-square value close to 1, we can say the estimation regression coefficient is statistical significant. But for other small sample problem (sample size isn't too small), the sample size also smaller than 30, we still need to draw out some information using the F-statistic or t-statistic. In other words, if the regression model in GM (1, 1) is statistically efficient, the t-statistic or F-statistic for indicating the significance level of the regression coefficient estimates must have p-values less than 0.05. And also R-square value close to 1.

According to the definition of coupling principle, the GM (1, 1) model is a coupled differential equation and regression model. If the data assimilated parameter pair (a, b) has a great efficiency, say, R-square value close to 1 and for F-test P value is less than 5%, While the grey model with the significant coefficient also has reasonable model efficiency, then we say the GM (1, 1) model have a quality coupling.

Definition 3.1.1: If the coefficient estimates generated from the regression model is significant, while the GM (1, 1) model with the significant coefficients also has fairly reasonable model efficiency. The phenomena showing in GM (1, 1) modelling we call it statistical-grey consistency.

Example 3.1.1: Let discrete data sequence $X^{(0)} = (6, 14, 15, 15.5, 17, 18, 20, 21, 25, 35, 40, 40, 45)$.

Now let us check the efficiency of regression model and the accuracy of GM (1, 1) model, comparing two of them to show if that GM (1, 1) model is a statistical-grey consistency model.

The following steps results from my program: the statistical-grey consistency checking Matlab toolbox.

Step1:

After 1-AGO treatment of $X^{(0)} = (6,14,15,15.5,17,18,20,21,25,35,40,40,45)$, we have

$X^{(1)} = (6,20,35,50.5,67.5,85.5,105.5,126.5,151.5,186.5,226.5,266.5,311.5)$

Step2:

Now we using the simple regression model of GM (1, 1)

$$y_k = b + ax_k, k = 2,3,4,\cdots,n \tag{3.1}$$

Where

$$\begin{gathered} y_k = x^{(0)}(k), x_k = -z^{(1)}(k), k = 2,3,4,\cdots,n; \\ x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4,\cdots,n; z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right] \end{gathered} \tag{3.2}$$

The estimate parameter pair (α,β), denoted as (a,b), could be calculated by multiple linear regression equation $(a,b)^T = (X^TX)^{-1}X^TY$, where

$$X = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} \tag{3.3}$$

The estimate parameter pair $(a,b) = (-0.1245, 10.1043)$

The simple regression model will become (See Figure 3.1.1)

$$\hat{y}_k = b + ax_k, k = 2,3,4,\cdots,n \tag{3.4}$$

$$\hat{y}_k = \begin{bmatrix} \hat{x}^{(0)}(2) \\ \hat{x}^{(0)}(3) \\ \vdots \\ \hat{x}^{(0)}(13) \end{bmatrix} = 10.1043 - 0.1245 \begin{pmatrix} -13 \\ -27.5 \\ -42.5 \\ -59 \\ -76.5 \\ -95.5 \\ -116 \\ -139 \\ -169 \\ -206 \\ -246.5 \\ -289 \end{pmatrix} = \begin{pmatrix} 11.7224 \\ 13.5272 \\ 15.4254 \\ 17.4481 \\ 19.6263 \\ 21.9913 \\ 24.543 \\ 27.4058 \\ 31.1399 \\ 35.8076 \\ 40.7865 \\ 46.0765 \end{pmatrix} \tag{3.5}$$

The coefficient estimates R-square value for $\hat{y}_k = b + ax_k, k = 2,3,4,\cdots,n$

$$R^2 = 1 - \frac{\sum (y_k - \hat{y}_k)^2}{\sum (y - \bar{y})^2} = 0.9621 = 96.21\% \tag{3.6}$$

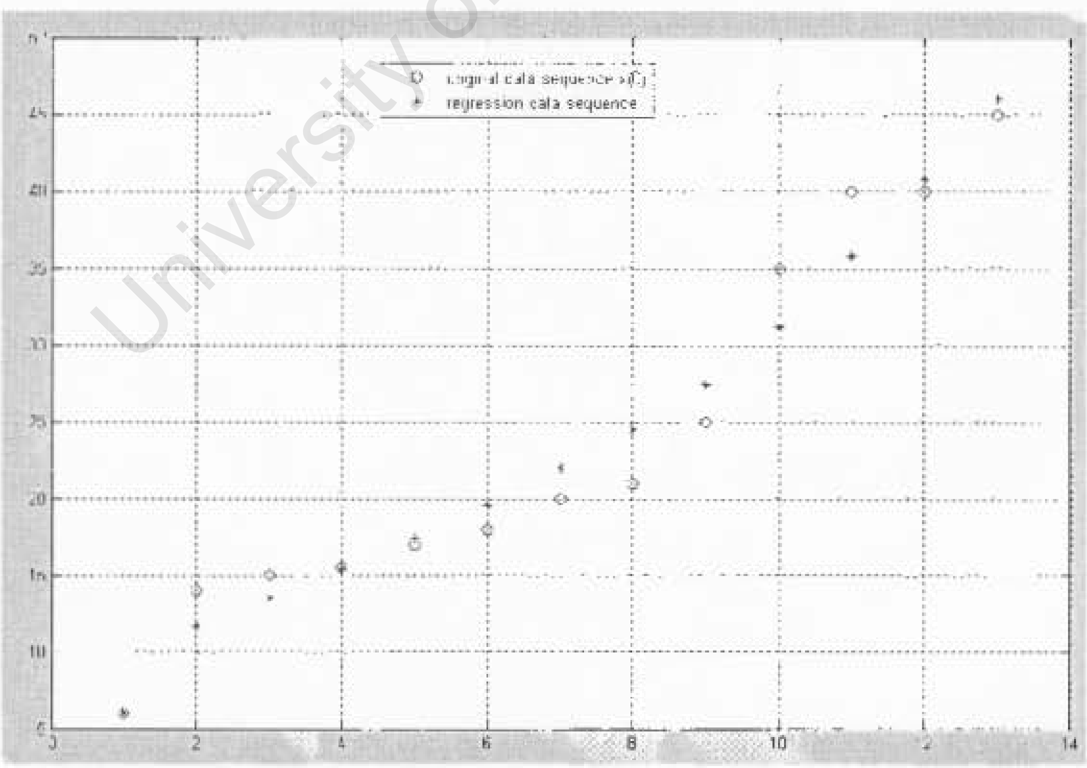

Figure 3.1.1 Whitening differential equation y (t)'s picture

Step 3

Generate the grey model and the time response sequence

$$\frac{d\hat{x}^{(1)}(t)}{dt} + a\hat{x}^{(1)}(t) = b \tag{3.7}$$

$$\frac{d\hat{x}^{(1)}}{dt} - 0.1245\hat{x}^{(1)} = 10.1043 \tag{3.8}$$

$$\begin{aligned} \hat{x}^{(1)}(k+1) &= \left[x^{(0)}(1) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a} \\ &= (6+81.1775)e^{0.1245k} - 81.1775 \\ &= 87.1775e^{0.124k} - 81.1775 \end{aligned} \tag{3.9}$$

Simulate value of $\hat{X}^{(1)}$ and estimation value of $\hat{X}^{(0)}$

$\hat{x}^{(1)}$ = [6.0000 17.5553 30.6423 45.4640 62.2503 81.2616 102.7928 127.1780 154.7954 186.0735 221.4975 261.6170 307.0542]

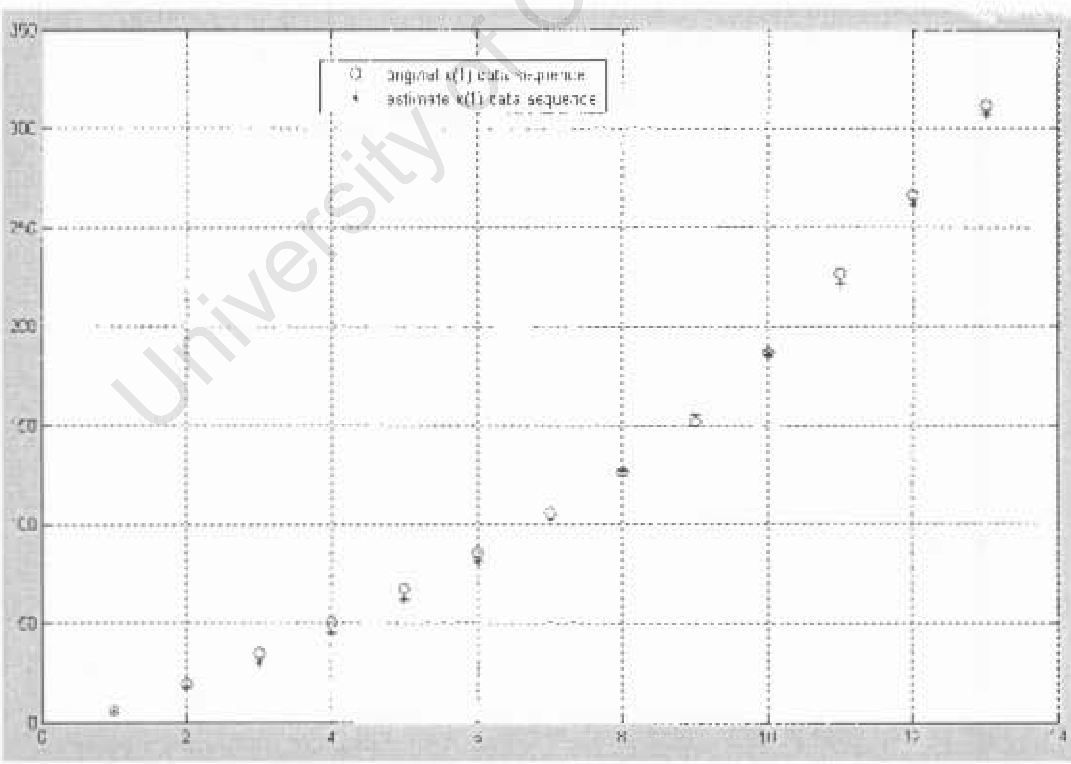

Figure 3.1.2 Comparing $X^{(1)}$ data sequence with $\hat{X}^{(1)}$ data sequence

The coefficient estimates R-square value for $\hat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a}$

$$R^2 = 1 - \frac{\sum (x^{(1)}(k) - \hat{x}^{(1)}(k))^2}{\sum \left(x^{(1)}(k) - \bar{x}^{(1)}(k)\right)^2} = 0.9984 = 99.84\% \tag{3.10}$$

The coefficient $(a,b) = (-0.1245, 10.1043)$ with estimate in regression model $y_k = b - ax_k, k = 2,3,4,\cdots,n$ transferred in differential equation, the fitting having better significant than in regression model.

Step4

Use 1-IAGO treatment $\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$ to get estimation value

$$\hat{x}^{(0)}(k) = [6.0000\ 11.5553\ 13.0870\ 14.8217\ 16.7863\ 19.0113\ 21.5312\ 24.3852\ 27.6174\ 31.2781\ 35.4240\ 40.1194\ 45.4373] \tag{3.11}$$

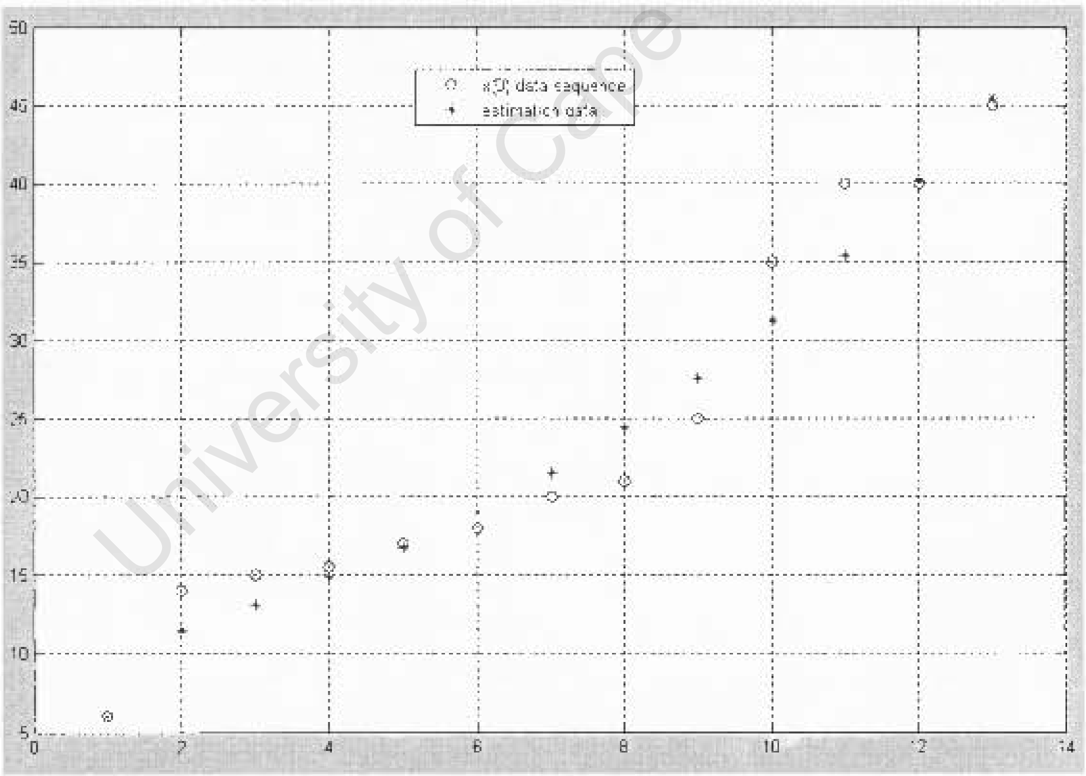

Figure 3.1.3 comparing $X^{(0)}$ data sequence with $\hat{X}^{(0)}$ data sequence

Calculate the Errors and Relative Errors

Errors:

$$\delta(k) = x^{(0)}(k) - \hat{x}^{(0)}(k) = (2.4447\quad 1.9130\quad 0.6783\quad 0.2137\quad 1.0113\quad 1.5312\quad 3.3852\quad 2.6174\quad 3.7219\quad 4.5760\quad 0.1194\quad 0.4373) \tag{3.12}$$

Relative Errors (%):

$$\Delta k = \left|\frac{\varepsilon(k)}{x^{(0)}(k)}\right| = (0.1746 \;\; 0.1275 \;\; 0.0438 \;\; 0.0126 \;\; 0.0562 \;\; 0.0766$$
$$0.1612 \;\; 0.1047 \;\; 0.1063 \;\; 0.1144 \;\; 0.0030 \;\; 0.0097) \tag{3.13}$$

Average relative error:

$$\Delta = \frac{1}{12}\sum_{k=2}^{13}\Delta_K = 0.0825 = 8.25\% \tag{3.14}$$

Accuracy:

$$\text{Accuracy}=0.9175=91.75\% \tag{3.15}$$

Table 3.1.1 Error evaluations for accuracies

No	Real Data $x^{(0)}(k)$	Simulated Values $\hat{x}^{(0)}(k)$	Errors $\varepsilon(k) = x^{(0)}(k) - \hat{x}^{(0)}(k)$	Relative Errors (%) $\Delta k = \left\|\frac{\varepsilon(k)}{x^{(0)}(k)}\right\|$
2	14	11.5553	2.4447	0.1746
3	15	13.0870	1.913	0.1275
4	15.5	14.8217	0.6783	0.0438
5	17	16.7863	0.2137	0.0126
6	18	19.0113	-1.0113	0.0562
7	20	21.5312	-1.5312	0.0766
8	21	24.3852	-3.3852	0.1612
9	25	27.6174	-2.6174	0.1047
10	35	31.2781	3.7219	0.1063
11	40	35.4240	4.576	0.1144
12	40	40.1194	-0.1194	0.003
13	45	45.4373	-0.4373	0.0097

From the last Example 3.1.1's 3 figures, we could see the regression model's coefficient significant is quite usual, coefficient of determination value R-square is 96.21%. When we use the same coefficient $(a,b) = (-0.1245, 10.1043)$ for the whitening differential equation estimate, the R-square value is closed to 1 (99.84%). That means the same parameter pair (a,b) which estimate from the regression model is showing better efficiency when it transfer to the whitening differential equation $\hat{X}^{(1)}$ data sequence estimation. But after the 1-AGO treatment, data transfer back to $\hat{X}^{(0)}$ level, the model Accuracy declined to 91.75%, because the errors is raised when we doing the data transformation.

When the regression model's coefficient significant value R-square is 96.21%, the model efficiency value Relative error is 8.25%, the accuracy is 1-8.25%=91.75%. Last example just showed great statistical-grey consistency.

However, for many cases, even the regression model of GM (1, 1) isn't showing any goodness-of-fit statistics, the GM (1, 1) is still reasonable for generate an efficiency model. From modelling practice experience, when the estimate regression coefficients aren't statistically significant, we may still obtain GM (1, 1) model efficiency being no less than 90%. This may causes a controversy between statistical information sufficiency and grey information sufficiency.

Definition 3.1.2: The coefficient estimates generated from the regression model isn't significant, while the GM (1, 1) model with the insignificant coefficients still has fairly reasonable model efficiency. The phenomena showing in GM (1, 1) modelling we call it statistical-grey inconsistency.

Example 3.1.2 the following sequence of data represents the morbidity rates of rape at Yun meng County of Hubei Province, the People's Republic of China, (data modified from Liu, S., and Lin, Y. (2006))

$$X^{(0)} = \left(x^{(0)}(i)\right)_{i=1}^{13} = (6,20,40,25,40,45,35,21,14,18,15.5,17,15)$$

Now let us check the efficiency of regression model and the accuracy of GM (1, 1) model, comparing two of them to show if that GM (1, 1) model is a statistical-grey consistency model.

The following steps results from the statistical-grey consistency checking Matlab toolbox, which originally made by myself. (See Appendix D)

Step1:

After 1-AGO treatment of $X^{(0)} = \left(x^{(0)}(i)\right)_{i=1}^{13} = (6,20,40,25,40,45,35,21,14,18,15.5,17,15)$, we have $X^{(1)} = (6,26,66,91,131,176,211,232,246,264,279.5,296.5,311.5)$

Step2:

Now we using the simple regression model

$$y_k = b + ax_k, k = 2,3,4,\cdots,n \tag{3.16}$$

Where

$$y_k = x^{(0)}(k), x_k = -z^{(1)}(k), k = 2,3,4,\cdots,n \tag{3.17}$$

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4,\cdots,n\, ; z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right] \tag{3.18}$$

The estimate parameter pair (a,b), could be calculated by multiple linear regression equation $(a,b)^T = \left(X^T X\right)^{-1} X^T Y$, where

$$X = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} \tag{3.19}$$

The estimate parameter pair $(a,b) = (0.0649, 37.2288)$

The simple regression model will become

$$\hat{y}_k = b + ax_k, k = 2,3,4,\cdots,n \tag{3.20}$$

$$\hat{y}_k = \begin{bmatrix} \hat{x}^{(0)}(2) \\ \hat{x}^{(0)}(3) \\ \vdots \\ \hat{x}^{(0)}(13) \end{bmatrix} = 37.2288 - 0.0649 \begin{pmatrix} 16 \\ -46 \\ 78.5 \\ -111 \\ 153.5 \\ -193.5 \\ -221.5 \\ 239 \\ 255 \\ -271.75 \\ 288 \\ -304 \end{pmatrix} = \begin{pmatrix} 36.1911 \\ 34.2453 \\ 32.1374 \\ 30.0295 \\ 27.273 \\ 24.6787 \\ 22.8626 \\ 21.7276 \\ 20.6899 \\ 19.6035 \\ 18.5495 \\ 17.5118 \end{pmatrix} \tag{3.21}$$

The coefficient estimates R-square value for $\hat{y}_k = b + ax_k, k = 2,3,4,\cdots,n$

$$R^2 = 1 - \frac{\sum (y_k - \hat{y}_k)^2}{\sum (y - \bar{y})^2} = 0.4539 = 45.39\% \tag{3.22}$$

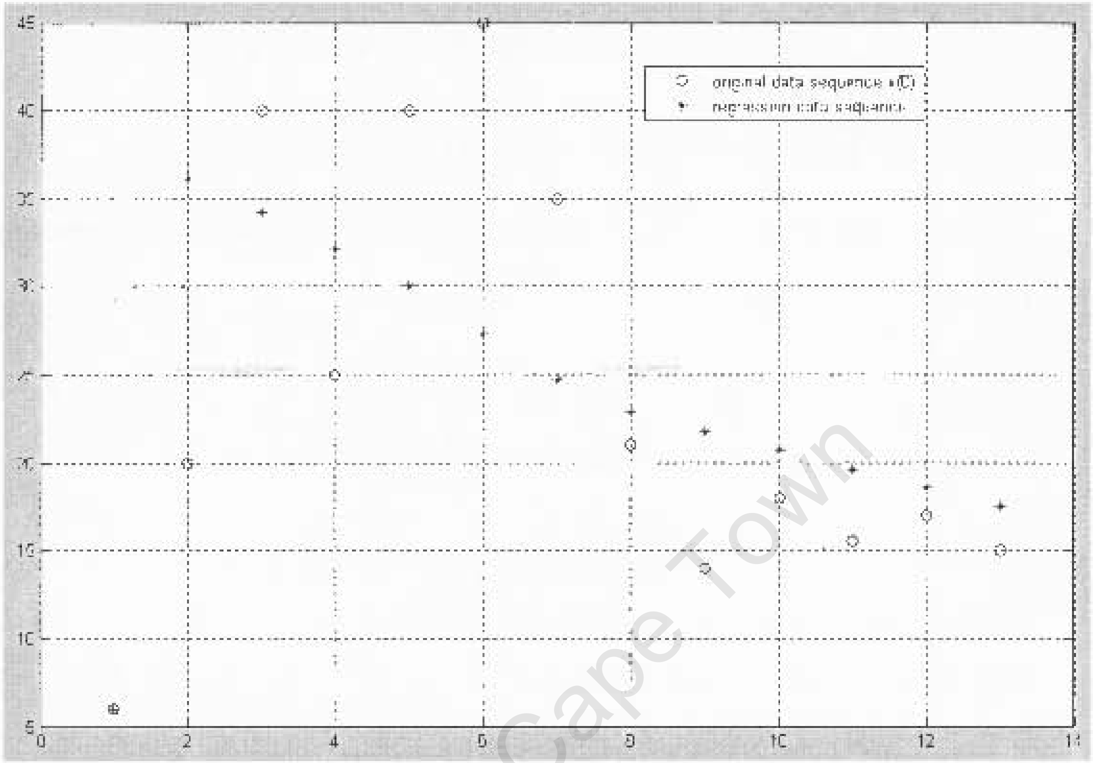

Figure 3.1.4 comparing $y^{(0)} = X^{(0)}$ data sequence with regression value $\hat{y}^{(0)}$ data sequence

Step 3

Generate the grey model and the time response sequence

$$\frac{d\hat{x}^{(1)}(t)}{dt} + a\hat{x}^{(1)}(t) = b \tag{3.23}$$

$$\frac{d\hat{x}^{(1)}}{dt} + 0.0649\hat{x}^{(1)} = 37.2288 \tag{3.24}$$

$$\begin{aligned} \hat{x}^{(1)}(k+1) &= \left[x^{(0)}(1) - \frac{b}{a}\right]e^{-ak} + \frac{b}{a} \\ &= (6 - 573.9991)\mathrm{e}^{-0.0649k} + 573.9991 \\ &= -567.9991\mathrm{e}^{-0.0649k} + 573.9991 \end{aligned} \tag{3.25}$$

Simulate value of $\hat{X}^{(1)}$ and estimation value of $\hat{X}^{(0)}$

$\hat{x}^{(1)}$ = [6.0000 41.6704 75.1007 106.4316 135.7949 163.3142 189.1052 213.2766 235.9300 257.1608 277.0583 295.7062 313.1830]

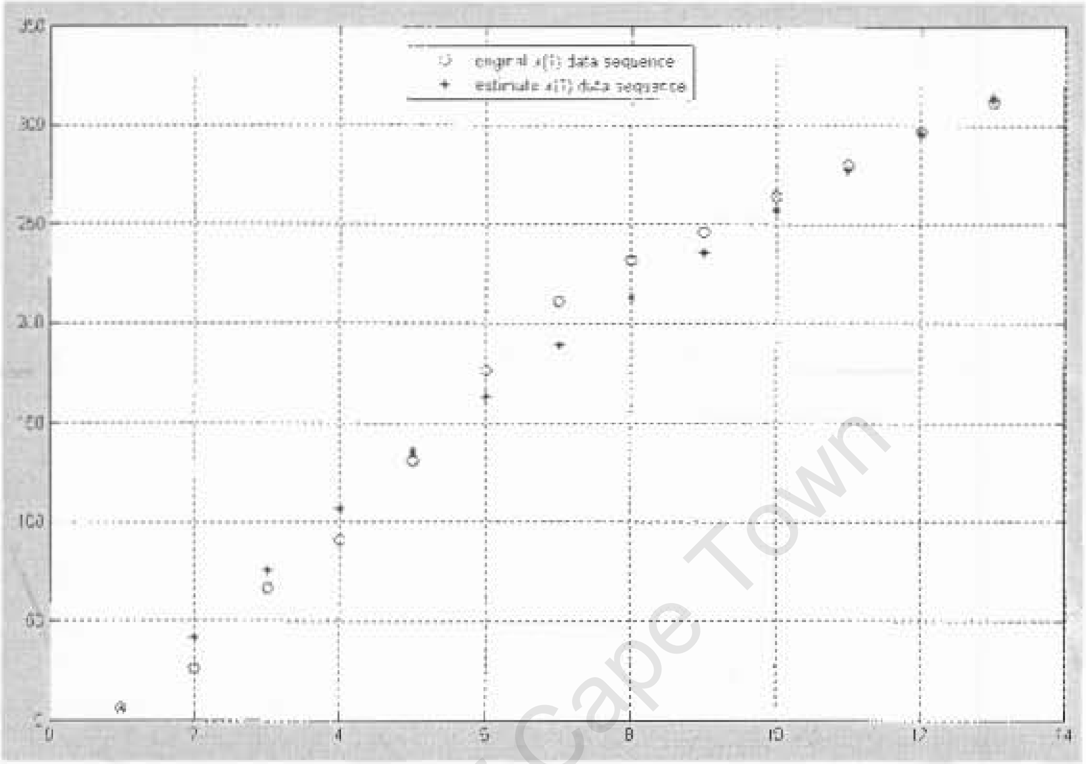

Figure 3.1.5 Comparing $X^{(1)}$ data sequence with $\hat{X}^{(1)}$ data sequence

The coefficient estimates R-square value for $\hat{x}^{(1)}(k+1)=\left[x^{(0)}(1)-\frac{b}{a}\right]e^{-ak}+\frac{b}{a}$

$$R^2=1-\frac{\sum(x^{(1)}(k)-\hat{x}^{(1)}(k))^2}{\sum\left(x^{(1)}(k)-\bar{x}^{(1)}(k)\right)^2}=0.9869=98.69\% \tag{3.26}$$

The coefficient $(a,b)=(0.0649,37.2288)$ with estimate in regression model $y_k=b+ax_k, k=2,3,4,\cdots,n$ transferred in differential equation, the fitting having better significant than in regression model.

Step4

Use 1-IAGO treatment $\hat{x}^{(0)}(k)=\hat{x}^{(1)}(k)-\hat{x}^{(1)}(k-1)$ to get estimation value

$\hat{x}^{(0)}(k)$=[6.0000 35.6704 33.4303 31.3309 29.3633 27.5193 25.7911 24.1714 22.6534 21.2308 19.8975 18.6479 17.4768]

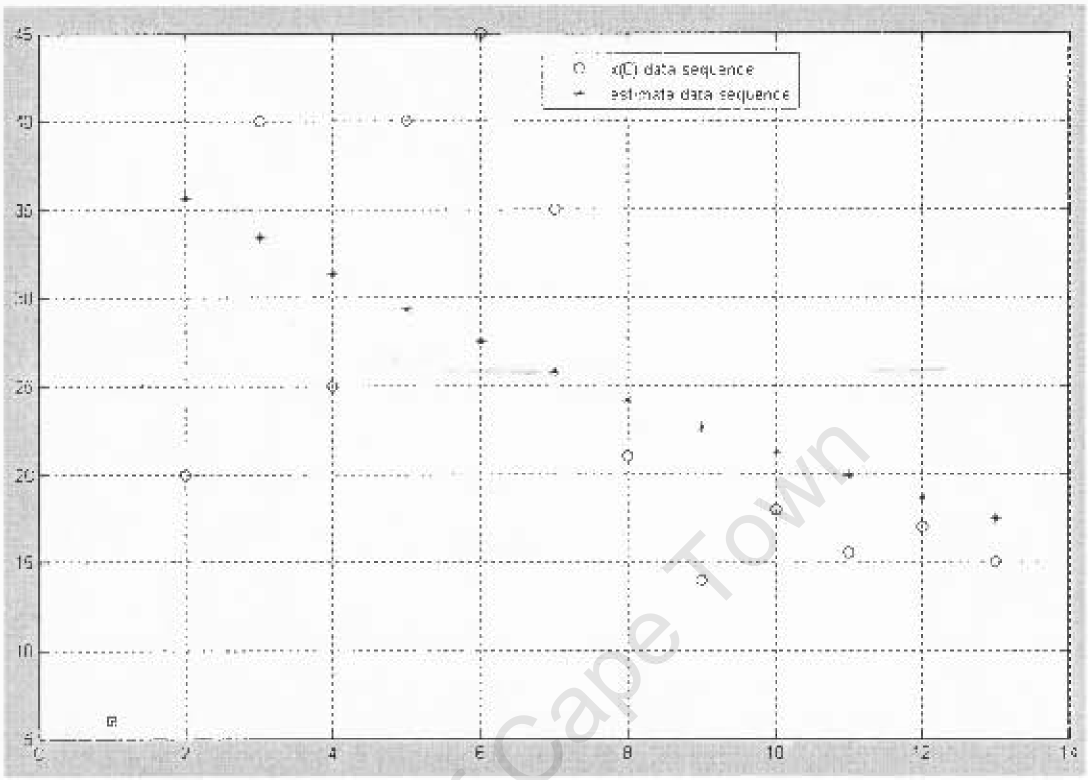

Figure 3.1.6 Comparing $X^{(0)}$ data sequence with $\hat{X}^{(0)}$ data sequence

Calculate the Errors and Relative Errors

Errors:

$$\delta(k)=x^{(0)}(k)-\hat{x}^{(0)}(k)=(-15.6704 \quad 6.5697 \quad -6.3309 \quad 10.6367 \quad 17.4807 \quad 9.2089 \quad -3.1714 \quad -8.6534 \quad -3.2308 \quad -4.3975 \quad -1.6479 \quad -2.4768) \tag{3.27}$$

Relative Errors (%):

$$\Delta k=\left|\frac{\varepsilon(k)}{x^{(0)}(k)}\right|=(0.7835 \quad 0.1642 \quad 0.2532 \quad 0.2659 \quad 0.3885 \quad 0.2631 \quad 0.1510 \quad 0.6181 \quad 0.1795 \quad 0.2837 \quad 0.0969 \quad 0.1651) \tag{3.28}$$

Average relative error:

$$\Delta=\frac{1}{12}\sum_{k=2}^{13}\Delta_k=0.3011=30.11\% \tag{3.29}$$

Accuracy

$$\text{Accuracy}=0.6989=69.89\% \tag{3.30}$$

Table 3.1.2 Error evaluations for accuracies

	Real Data	**Simulated Values**	**Errors**	**Relative Errors(%)**
No.	$x^{(0)}(k)$	$\hat{x}^{(0)}(k)$	$\varepsilon(k)=x^{(0)}(k)-\hat{x}^{(0)}(k)$	$\Delta k=\left\lvert\frac{\varepsilon(k)}{x^{(0)}(k)}\right\rvert$
2	20	35.6704	-15.6704	0.7835
3	40	33.4303	6.5697	0.1642
4	25	31.3309	-6.3309	0.2532
5	40	29.3633	10.6367	0.2659
6	45	27.5193	17.4807	0.3885
7	35	25.7911	9.2089	0.263
8	21	24.1714	-3.1714	0.151
9	14	22.6534	-8.6534	0.6181
10	18	21.2308	-3.2308	0.1795
11	15.5	19.8975	-4.3975	0.2837
12	17	18.6479	-1.6479	0.0969
13	15	17.4768	-2.4768	0.1651

From the last Example 3.1.2's 3 figures, we could see the regression model's coefficient significant is quite worse, coefficient of determination value R-square is only 45.39%. When we use the same coefficient $(a,b)=(0.0649,37.2288)$ for the whitening differential equation estimate, the R-square value is closed to 1 (98.69%). That means the same parameter pair (a,b) which estimate from the regression model is showing much better efficiency when it transfer to the whitening differential equation $\hat{X}^{(1)}$ data sequence estimation. But after the 1-AGO treatment, data transfer back to $\hat{X}^{(0)}$ level, the model Accuracy declined to only 69.89%, which meaning 1-AGO and 1-IAGO transformation could made some errors and result the estimation data lost its efficiency.

The regression model's coefficient significant value R-square is 45.39%, but after transfer parameter $(a,b)=(0.0649,37.2288)$ to the whitening differential equation model, the estimate value of $\hat{X}^{(1)}$ data sequence showing a incredible significant (comparing with $X^{(1)}$ data sequence), the R-square value is closed to 1 (98.69%).But the final GM (1 ,1) model efficiency is not good, the accuracy is 1-30.11%=69.89%.

Last example just showed what we called statistical-grey inconsistency. In next section I will create a new model to decline the errors when data transfer from $\hat{X}^{(1)}$ to $\hat{X}^{(0)}$ level.

3.2 Ratio idea to improve efficiency of Statistical-Grey Consistency model

When we doing GM (1, 1) modelling, normally, if the data sequence is a monotonic one, after IAGO treatment of $\hat{X}^{(1)}$ data sequence, the result generally will showing a good efficiency. But if the original data sequence doesn't following a monotonic tendency, after the IAGO treatment of $\hat{X}^{(1)}$, the estimation value $\hat{X}^{(0)}$ usually isn't such efficient as what we expected.

Example 3.1.2 is a statistical-grey inconsistent model, the regression part could not fit a nice efficient linear model, but after transfer parameter (a,b) to the whitening differential equation model, the estimate value of $\hat{X}^{(1)}$ data sequence just showing a incredible significant (comparing with $X^{(1)}$ data sequence), the R-square value is closed to 1. But the final GM (1, 1) model efficiency isn't expectable nice, the accuracy is only 1-30.11%=69.89%. Now we are using ratio idea to instead of IAGO treatment, to avoid too much errors engendering when data transfer back to $\hat{X}^{(0)}$ level.

Definition 3.2.1: Assume that

$$X^{(0)} = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)),$$

is a non-negative sequence, where $x^{(0)}(k) \geq 0, k = 1,2,\cdots,n, X^{(1)}$ the 1-AGO sequence of $X^{(0)}$ with

$$X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right),$$

where

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i),$$

$k = 1,2,\cdots,n,$ and R is the ratio value generated sequence which given by

$$R = \frac{X^{(1)}}{X^{(0)}} = \frac{x^{(0)}(i)}{\sum_{i=1}^{k} x^{(0)}(i)}$$

After we estimate $\hat{X}^{(1)}$ data sequence, we utilize $\hat{X}^{(1)}/R$ to instead of 1-AGO operation to generate $\hat{X}^{(0)}$ value, which given by

$$\hat{X}^{(0)} = \frac{\hat{X}^{(1)}}{R}$$

We use the Ratio idea to instead of IAGO treatment to improve the efficiency of the same Example 3.1.2 modelling, denoted as Example 3.2.1.

Example 3.2.1: the following sequence of data represents the morbidity in rape at Yun meng County of Hubei Province, the People's Republic of China,

$$X^{(0)} = \left(x^{(0)}(i)\right)_{i=1}^{13} = (6,20,40,25,40,45,35,21,14,18,15.5,17,15)$$

Step 1:

After 1-AGO treatment of $X^{(0)} = \left(x^{(0)}(i)\right)_{i=1}^{13} = (6,20,40,25,40,45,35,21,14,18,15.5,17,15)$, we have $X^{(1)} = (6,26,66,91\ ,131,176,2\ 11,232,246\ ,264,279.5\ ,296.5,311\ .5)$

Then we using the ratio idea to calculate the ration data sequence using $X^{(0)} = \left(x^{(0)}(i)\right)_{i=1}^{13}$ divide by $X^{(1)} = \left(x^{(1)}(i)\right)_{i=1}^{13}$ denote as

$$R = \left(r(i)\right)_{i=1}^{13} = \frac{X^{(1)}}{X^{(0)}} = \frac{\left(x^{(1)}(i)\right)_{i=1}^{13}}{\left(x^{(0)}(i)\right)_{i=1}^{13}} \tag{3.31}$$

$$\begin{aligned} R = (&1.0000 \quad 1.3000 \quad 1.6500 \quad 3.6400 \quad 3.2750 \quad 3.9111 \\ &6.0286 \quad 11.0476 \quad 17.5714 \quad 14.6667 \quad 18.0323 \quad 17.4412 \quad 20.7667) \end{aligned} \tag{3.32}$$

Step 2 and Step 3 is exactly same as example 3.1.2.

Step 4:

Use ratio idea $\hat{x}^{(0)}(k) = \dfrac{\hat{x}^{(1)}(k)}{R(k)}$ instead of 1-IAGO treatment to get estimation value

$$\begin{aligned} \hat{x}^{(0)}(k) = [&6.0000 \quad 32.0542 \quad 45.5156 \quad 29.2395 \quad 41.4641 \quad 41.7565 \\ &31.3682 \quad 19.3052 \quad 13.4269 \quad 17.5337 \quad 15.3646 \quad 16.9545 \quad 15.0810] \end{aligned}$$

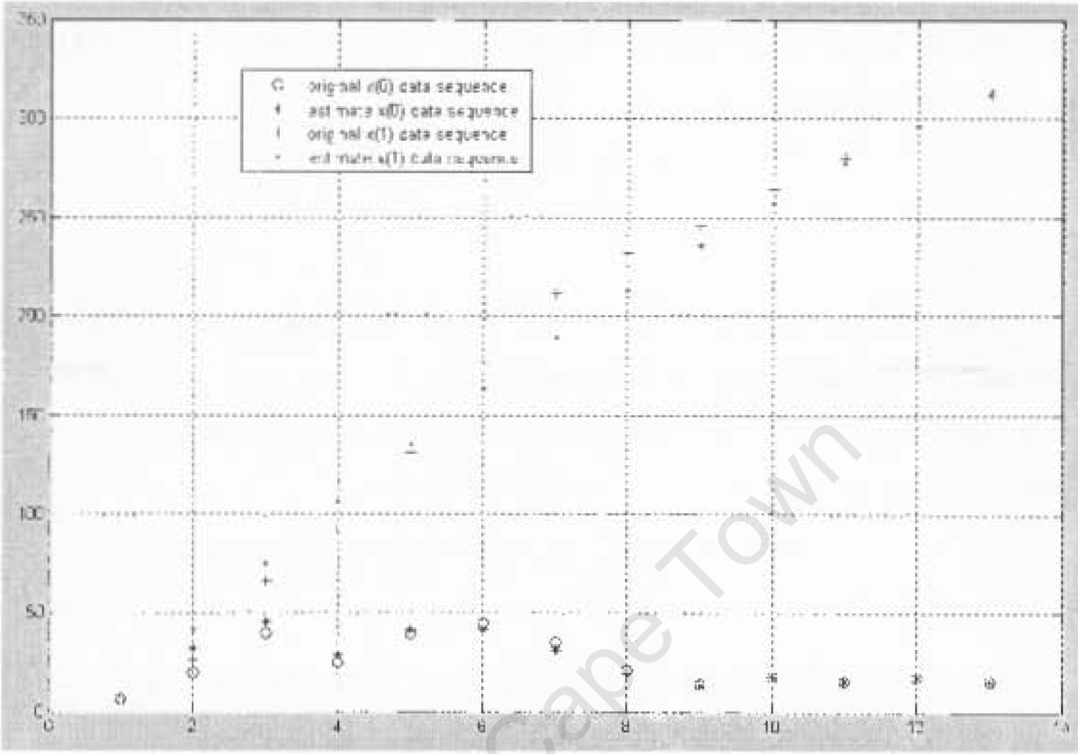

Figure 3.2.1 comparing $X^{(0)}$ data sequence with $\hat{X}^{(0)}$ data sequence; comparing $X^{(1)}$ data sequence with $\hat{X}^{(1)}$ data sequence

Calculate the Errors and Relative Errors

Errors:

$$\delta(k) = x^{(0)}(k) - \hat{x}^{(0)}(k) = (-12.0542\ \ -5.5156\ \ -4.2395\ \ -1.4641\ \ 3.2435\ \ 3.6318\ \ 1.6948\ \ 0.5731\ \ 0.4663\ \ 0.1354\ \ 0.0455\ \ -0.0810) \tag{3.33}$$

Relative Errors (%):

$$\Delta k = \left|\frac{\varepsilon(k)}{x^{(0)}(k)}\right| = (0.6027\ \ 0.1379\ \ 0.1696\ \ 0.0366\ \ 0.0721\ \ 0.1038\ \ 0.0807\ \ 0.0409\ \ 0.0259\ \ 0.0087\ \ 0.0027\ \ 0.0054) \tag{3.34}$$

Average relative error:

$$\Delta = \frac{1}{12}\sum_{k=2}^{13}\Delta_K = 0.1072 = 10.72\% \tag{3.35}$$

Accuracy

$$\text{Accuracy} = 0.8928 = 89.28\% \tag{3.36}$$

Table 3.2.1 Error evaluations for accuracies

	Real Data	**Simulated Values**	**Errors**	**Relative Errors (%)**
No.	$x^{(0)}(k)$	$\hat{x}^{(0)}(k)$	$\varepsilon(k)=x^{(0)}(k)-\hat{x}^{(0)}(k)$	$\Delta k = \left\lvert \frac{\varepsilon(k)}{x^{(0)}(k)} \right\rvert$
2	20	32.0542	-12.0542	0.6027
3	40	45.5156	-5.5156	0.1379
4	25	29.2395	-4.2395	0.1696
5	40	41.4641	-1.4641	0.0366
6	45	41.7565	3.2435	0.0721
7	35	31.3682	3.6318	0.1038
8	21	19.3052	1.6948	0.0807
9	14	13.4269	0.5731	0.0409
10	18	17.5337	0.4665	0.0259
11	15.5	15.3646	0.1354	0.0087
12	17	16.9545	0.0455	0.0027
13	15	15.081	-0.081	0.0054

After utilizing the ratio idea, example 3.1.2's modelling accuracy raised from 69.89% to 89.28%.

The coefficients estimate generated from the regression model isn't significant, R-square value is only 45.39%, far away from 1. While the GM (1, 1) model with the insignificant coefficients still has fairly reasonable model efficiency, Accuracy is 89.28%. The phenomena showing in example 4.3 modelling we call it statistical-grey inconsistency.

3.3 Comparisons between Coupling Model and Variation Model.

Before we realized the essential of GM (1, 1) model is a regression model coupling with a differential equation model, we often regarded GM (1, 1) model as a regression model with differential equation constraint. Why we insist to state the coupling idea is the essential characteristic of GM (1, 1)? Which model is superior if we utilize the GM (1, 1) modelling for a real application testing? In this section, after us examining GM (1, 1) model with the variation calculus approach knowledge and analysis of the constrained regression model, we could simulate a reasonable answer to prove our statement about the coupling model idea is true.

L2-normed GM (1, 1) model could be understand as a kind of approximation method, try to searching paired parameter to minimum the relative errors equation and finally find the efficient estimation of the input value. But the L2-normed GM (1, 1) model is investigated from nonlinear least-square estimation at the "integrated" data level, i.e., at $x^{(1)}$ level. The original observation data sequence is $\left[X^{(0)}\right]^{obs} = \left(\left[X^{(0)}(0)\right]^{obs}, \left[X^{(0)}(1)\right]^{obs}, \cdots, \left[X^{(0)}(n)\right]^{obs}\right)$, after 1-AGO (abbreviate of accumulating generation operator) treatment, we will have the $\left[X^{(1)}\right]^{obs}$ data sequence value.

$$\left[x^{(1)}(k)\right]^{obs} = AGO\left(\left[X^{(0)}\right]^{obs}\right)_K = \sum_{i=1}^{K}\left[x^{(0)}(i)\right]^{obs}, k = 0,1,\cdots,n \tag{3.37}$$

After that, we need utilize least-square approaching method with L2-norm to get the estimate parameter a and b denoted as $\hat{a}$ and $\hat{b}$ respectively. Then we utilize the estimate parameter $\hat{a}$ and $\hat{b}$ to the differential equation $dx^{(1)}/dt + \hat{a}x^{(1)} = \hat{b}$, which using the filtering-predictive equation takes the discrete version of solution (with GM (1, 1) least-square estimated parameter-values).

$$\hat{x}^{(1)}(k) = \left(x^{(1)}(0) - \hat{b}/\hat{a}\right)\exp\left(-\hat{a}(k)\right) + \hat{b}/\hat{a} \tag{3.38}$$

But still now, there is no clarity on fundamental issues of why the parameters estimated by GM (1, 1) modelling can be used in the solution to the differential equation for filtering-prediction purposes and Why are the AGO and MEAN operators introduced, even there are bunches of academic papers of GM (1, 1) modelling have been published.

Because the estimation must take on under the integrated level (i.e. $x^{(1)}$ level), so we give an idea to find appropriate composition of $\left\{z^{(1)}(k), k = 1,2,\cdots,n\right\}$ which could made the integrated observation data sequence $\left\{x^{(1)}(k), k = 0,1,\cdots,n\right\}$ can be formed and lead to an optimal statistical estimation with respect to parameter a, b.

$\left\{z^{(1)}(k), k = 1,2,\cdots,n\right\}$ data sequence is the only one character based on the level $x^{(1)}$, which give by :

$$z^{(1)}(k) = \frac{1}{2}\left(\left[x^{(1)}(k)\right]^{obs} + \left[x^{(1)}(k-1)\right]^{obs}\right) \tag{3.39}$$

Because the weight factor 1/2 is an arbitrary decision, which didn't have any mathematical discussion. Now we discussion the weight factor, denoted as ω, between $\left[x^{(1)}(k)\right]^{obs}$ and

$\left[x^{(1)}(k-1)\right]^{obs}$ should not be predetermined rather than determined by the optimization procedure. In other words, the $z^{(1)}(k)$ data sequence should be defined as:

$$z^{(1)}(k)=\omega\left[x^{(1)}(k)\right]^{obs}+(1-\omega)\left[x^{(1)}(k-1)\right]^{obs}, \omega \in [0,1] \tag{3.40}$$

Now we give the form of the L2-normed estimation equation:

$$\min_{\alpha,\beta,\gamma}\left\{J^{(1)}\right\}=\sum_{i=2}^{n}\left(\left[z^{(1)}(i)\right]^{obs}-x^{(1)}(i)\right)^{2} \tag{3.41}$$

Where

$$\left[x^{(1)}(k)\right]=AGO\left(\left[X^{(0)}\right]\right)_k=\sum_{i=1}^{k}\left[x^{(0)}(i)\right], k=0,1,\cdots,n \tag{3.42}$$

The equation 3.41 could be decomposed as following non-linear equations with the parameter a, b and ω.

$$\begin{cases} \sum_{i=2}^{n}\dfrac{\partial x^{(1)}(i)}{\partial a}\left[z^{(1)}(i)-x^{(1)}(i)\right]=0 \\ \sum_{i=2}^{n}\dfrac{\partial x^{(1)}(i)}{\partial b}\left[z^{(1)}(i)-x^{(1)}(i)\right]=0 \\ \sum_{i=2}^{n}\left(x^{(1)}(i)-x^{(1)}(i-1)\right)\left[z^{(1)}(i)-x^{(1)}(i)\right]=0 \end{cases} \tag{3.43}$$

We simply utilize an example to prove the parameter ω is not necessary to be 0.5.

Example 3.3.1: We have a discrete data sequence denoted as:

$X^{(0)}=\{2.874, 3.278, 3.337, 3.390, 3.679\}$

Step 1

We first using Matlab genetic algorithm toolbox to estimate

In this example, we give an initial value of $x^{(1)}(0)$, $\gamma=2.874$

Then the L2-normed estimation equation will be:

$$\begin{aligned}\min_{\alpha,\beta,\omega}\left\{J^{(1)}\right\}&=\sum_{i=2}^{n}\left(\left[z^{(1)}(i)\right]^{obs}-x^{(1)}(i)\right)^{2}\\ &=\sum_{i=2}^{n}\left(\left[\omega\left[x^{(1)}(i)\right]^{obs}+(1-\omega)\left[x^{(1)}(i-1)\right]^{obs}\right]-\left[\left[x^{(0)}(1)-\frac{b}{a}\right]\exp(-a*i)+\frac{b}{a}\right]\right)^{2}\end{aligned} \tag{3.44}$$

Step 2

We use genetic algorithm to estimate the equation:

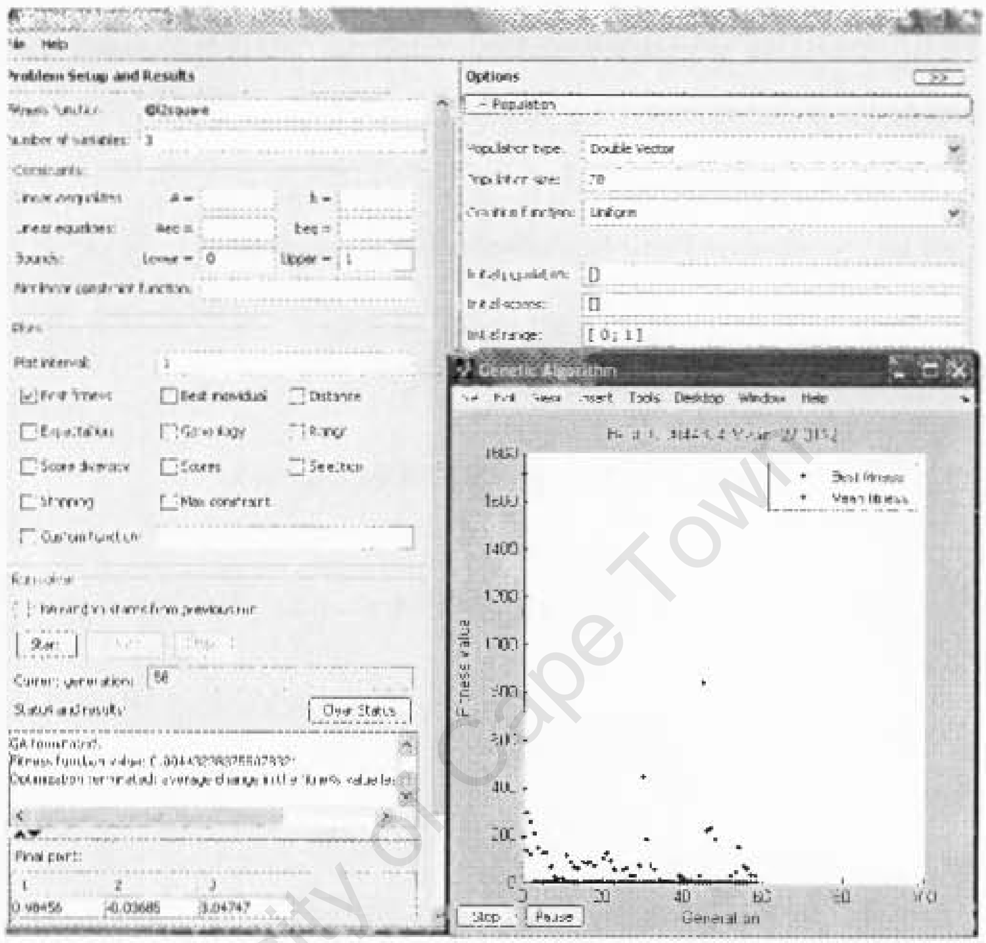

Figure 3.3.1 Using Genetic algorithm to estimate the non-linear equation with parameter α, β and ω.

The equation M-files named l2square.m

```
function e=l2square(x)
e=0;ay(1)=2.874;
y=[2.874 3.278 3.337 3.390 3.679];
for i=2:length(y)
   ay(i)=y(i)+ay(i-1);
end
for i2=2:length(y)
   e=((x(1)*ay(i2)+(1-x(1))*ay(i2-1))-((y(1)-x(3)/x(2))*exp(-x(2)*(i2-1))+x(3)/x(2)))^2+e;
end
```

Step 3

From figure 3.3.1, we know:

$$\min_{a,b,\omega}\left\{J^{(1)}\right\}=\sum_{i=2}^{n}\left(\left[z^{(1)}(i)\right]^{obs}-x^{(1)}(i)\right)^2=0.0044324 \tag{3.45}$$

Where

$$a=-0.03685, \omega=0.98456, b=3.04747 \tag{3.46}$$

Step 4

Now we could determine the model by:

$$\frac{d\hat{x}^{(1)}}{dt}-0.03685\,\hat{x}^{(1)}=3.04747 \tag{3.47}$$

And the time response sequence

$$\hat{x}^{(1)}(k+1)=\left[x^{(0)}(1)-\frac{b}{a}\right]e^{-ak}+\frac{b}{a}=85.5733\,e^{0.03685\,k}-82.699 \tag{3.48}$$

Step5

Solve the model obtained in step 4 for the simulation value of $X^{(1)}$

$$\hat{X}^{(1)}=\left(\hat{x}^{(1)}(i)\right)_{i=1}^{5}=(2.8740, 6.0862, 9.419, 12.8769, 16.4645) \tag{3.49}$$

Restore the $\hat{X}^{(1)}$-value to find the simulation value of $X^{(0)}$. From

$$\hat{x}^{(0)}(k)=a^{(1)}\hat{x}^{(1)}(k)=\hat{x}^{(1)}(k)-\hat{x}^{(1)}(k-1), \tag{3.50}$$

It follows that

$$\hat{X}^{(0)}=\left(\hat{x}^{(0)}(i)\right)_{i=1}^{5}=(2.8740, 3.2122, 3.3328, 3.4579, 3.5876) \tag{3.51}$$

Step6

Evaluate the error. The following Table 3.3.1 gives the relevant error values

Table 3.3.1 Error evaluations for accuracies

No.	Real Data $x^{(0)}(k)$	Simulated Values $\hat{x}^{(0)}(k)$	Errors $\varepsilon(k)=x^{(0)}(k)-\hat{x}^{(0)}(k)$	Relative Errors (%) $\Delta k=\left\|\frac{\varepsilon(k)}{x^{(0)}(k)}\right\|$
2	3.278	3.2122	0.0658	2.01
3	3.337	3.3328	0.0042	0.13
4	3.390	3.4579	-0.0679	2.00
5	3.679	3.5876	0.0914	2.48

And obtain the average relative error:

$$\Delta=\frac{1}{4}\sum_{2}^{5}\Delta_K=1.65\% \tag{3.52}$$

The accuracy is:

$$1\text{-}0.0165=100\%\text{-}1.65\%=98.35\% \tag{3.53}$$

Comparing the L2-normed model with the grey coupling model to figure out which one performed more efficient of GM (1, 1) modelling. (See table 3.3.2)

Table 3.3.2 comparing the L2-normed model with the grey coupling model

GM (1, 1) model	a	b	ω	Average relative errors	Accurac y
L2-normed model	-0.03685	3.04747	0.98456	1.65%	98.35%
Grey coupling model	-0.0372	3.0654	0.5	1.60%	98.40%

According to the result of Table 3.3.2, even we using genetic algorithm to minimize the relative error between $Z^{(1)}$ and $X^{(1)}$ to search best estimate parameter pair (α,β), but the accuracy of the L2-normed model is still little bit lower than coupling model. After checking of example results, we make a conclusion: the GM (1, 1) model can't be just defined as a regression model with differential equation constraint, because if we understand it as a regression model coupling with a differential equation model, the result shows better than first understanding.

3.4 Statistical-Grey Consistency Criterions

In last section, according to the practice result, we understand the GM (1, 1) model can't explain as a regression model with a differential equation constraint, but a regression model coupling with a differential equation model. The coupling steps usually could conclude as follows:

First, we define the modified input data sequence $X^{(0)}$ as vector Y, then generate the $Z^{(1)}$ data values from 1-AGO treatment data sequence $X^{(0)}$, we denoted as $X^{(1)}$.

The least-square regression equation is

$$(a,b)^T = \left(X^T X\right)^{-1} X^T Y \tag{3.54}$$

$$X = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix} Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} \tag{3.55}$$

Where

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4,\cdots,n \, ; z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right] \tag{3.56}$$

Then utilize the linear form regression model to estimate the appropriate parameters (a,b).

Second, transfer the estimate parameters (a,b) to the whitening differential equation $\frac{d\hat{x}^{(1)}(t)}{dt} + a\hat{x}^{(1)}(t) = b$, and using $\hat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a}\right]e^{-ak} + \frac{b}{a}$ as the time response equation to calculate the estimate value $\hat{X}^{(1)}$ data sequence. Then Use 1-IAGO treatment $\hat{x}^{(0)}(k)=\hat{x}^{(1)}(k)-\hat{x}^{(1)}(k-1)$ to get final solution.

If the coefficient estimates generated from the regression model is significant, while the GM (1, 1) model with the significant coefficients also has fairly reasonable model efficiency. The phenomena showing in GM (1, 1) modelling we call it statistical-grey consistency. The quality of coupling mainly judged by two aspects: statistical index and grey index. The two aspects could control the input data sequence with a certain constraint making the final result keeping a high statistical-grey consistency.

Let us check the statistical index first:

From grey modelling practice experience, the sample size of GM (1, 1) modelling usually quite small, it may larger than 4 and less than 20. Because the statistical regression significance asks the observation data highly correlated with values of exploratory variable, so the response data observations linked with the linear transformed response data very tightly. Because grey model's small sample characteristic, when the sample size is extremely small, such as only 4 or 5 data points, we only can obtain one index for regression significant checking, R^2 (R-square) value. When R-square value is much closed to 1, we say the regression model is statistically significant. If the sample size isn't extremely small, could be test for some distribution, then we still can use p-value to test the regression significant. Generally it may expect the p-values less than 0.05.

If we considered the error terms of the regression model, assumed the error terms following a standard normal distribution, then the normal random variable is an identically independent variable with zero mean and constant variance σ^2. After we have the error terms, we may use Maximum likelihood theory to further examine the statistical significant of regression model. According to Maximum likelihood theory, any function's Maximum likelihood function is also a Maximum likelihood function of the all. Then we could do likelihood ratio test to check the p-value larger or less than 0.05 to see if the regression model is statistically significant.

Now let us check the grey index:

There are two method could use for grey index checking, the first one is a traditional idea which we already give some introduce in Chapter2 of GM (1, 1) modelling procedure. Another one is a new idea, which simply made by myself, using spline function and geometric similarity method to justify the grey model efficiency.

The first grey index is the class ratio test with the potential degree of data assimilation.

$$\sigma^{(0)}(k) = \frac{x^{(0)}(k-1)}{x^{(0)}(k)}, k = 2,3,\cdots,n \tag{3.57}$$

With the constraint

$$\sigma^{(0)}(k) \in \left(e^{-2/(n+2)}, e^{2/(n+2)}\right) \tag{3.58}$$

For k=2, 3… n, Then after once accumulating generation operator treatment the original data sequence $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\}$ will become a simply exponential trend data sequence $X^{(1)} = \{x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\}$ where $x^{(1)}(k) = \sum_{i=1}^{n} x^{(0)}(i)$ such that the degree of data assimilation will get into a high level.

The second grey index is Grey relation analysis using spline function for curve fitting.

Suppose given a discrete data sequence $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\}$, we using polynomial spline function fitting the original data point with (n-1) order, and got the fitted response equation according to the GM (1, 1) coupling model, denoted as $\varphi(t)$. Then we could obtain the spline absolute degree of grey relation (ADGR).

$$g(t) = \frac{1 + \left|1 + \int^{t} P_{n-1}(t)dt\right| + \left|\int^{t} \varphi(t)dt\right|}{1 + \left|\int^{t} P_{n-1}(t)dt\right| + \left|\int^{t} \varphi(t)dt\right| + \left|\int^{t} \left(P_{n-1}(t) - \varphi(t)\right)dt\right|} \tag{3.59}$$

We simply comparing the original data curve with fitted response curve to see if the illustration is geometric similar.

After we obtain the grey index and the statistical index, we could figure out whether the two ones are compatible with each other. We could use R-square value as statistical index and the spline-ADGR as grey index. If both of them closed to 1, the coupling GM (1, 1) model should be received a high efficient statistical-grey consistency, otherwise, the statistical-grey inconsistency problem will appears, then we have to alternative equation form of the differential equation for grey differential equation model building.

3.5 Summary

In this chapter, we point out the importance of the goodness-of-fit information of the simple regression model itself. During the modelling procedure, if the data assimilated parameter pair has a great efficiency, while the grey model with the significant coefficient also has reasonable model efficiency. After that, we give the definition of statistical-grey consistency and statistical-grey inconsistency. And we also mentioned about a new ratio idea to improve efficiency of Statistical-Grey Consistency model.

Chapter 4. Role of Spline Models in Statistical-Grey Consistent Models

Definition 4.1: In the mathematical field of numerical analysis, spline interpolation is a form of interpolation where the interpolation is a special type of piecewise polynomial called a spline. Spline interpolation is preferred over polynomial interpolation because the interpolation error can be made small even when using low degree polynomials for the spline. Thus, spline interpolation avoids the problem of Runge's phenomenon which occurs when using high degree polynomials. (http://en.wikipedia.org/wiki/Spline_interpolation)

Definition 4.2: In the mathematical field of numerical analysis, Runge's phenomenon is a problem that occurs when using polynomial interpolation with polynomials of high degree. It was discovered by Carl David Tolme Runge when exploring the behaviour of errors when using polynomial interpolation to approximate certain functions (Runge 1901). (http://en.wikipedia.org/wiki/Runge%27s_phenomenon)

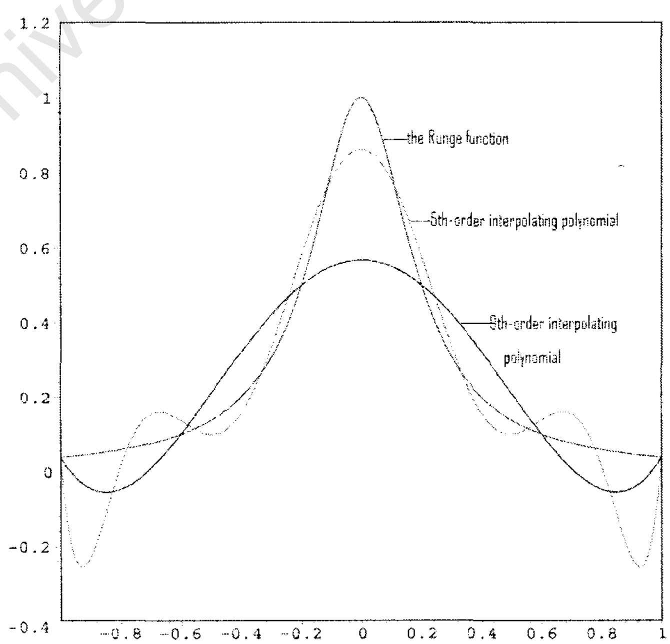

Figure 4.1 Runge's phenomenon

The red curve is the Runge function. The blue curve is a 5th-order interpolating polynomial (using six equally-spaced interpolating points).The green curve is a 9th-order interpolating polynomial (using ten equally-spaced interpolating points).

In grey mathematics, the nature of accumulative generating operation (abbreviated as AGO) and the inverse accumulative generating operation (IAGO) is actually a transformation between the original data sequence $X^{(0)}$ and another kind of data form $X^{(1)}$, which could made data sequence more smoothness and following a tendency of data curve increasing. The problem is the duration of the data transformation always generates some unavoidable errors which may decline the model estimation efficiency, such problem could be solved by anther transformation method (integration and derivative of certain spline function transformation) .The discussion of this chapter is the exchangeability between AGO, IAGO and the integration of certain spline function, the derivative of certain spline function respectively.

In this chapter I present two papers named of role of spline models in statistical-grey consistent models, which utilize smoothing splines as generalizations of interpolation splines where the statistical-grey consistent model are determined to minimize a weighted combination of the average squared approximation error over observed data and the roughness measure.

4.1 Spline function and first-order one-variable grey differential equation model

(YAN HONG CUI, RENKUAN GUO, AND TIM DUNNE)

4.1.1 Introduction

Grey mathematics is a mathematical branch dealing with dynamic law revelation with sparse data availability. Grey differential equations play important roles in grey modelling. However, as Li, Yamaguchi and Nagai (2005) pointed out that there exist some fundamental problems in grey differential equation modelling including initial values, background values and grey derivative specifications. Various remedies are proposed in the grey mathematics literature. We noticed that usage of spline functions is one of them. In this paper, we review a couple of class of spline functions under certain optimal conditions. Then we re-examine the nature and the optimality of the two critical data operations proposed by Deng (2002), accumulative generating operation (abbreviated as AGO) and the inverse accumulative generating operation (abbreviated as IAGO) and then the exchangeability between AGO, IAGO and the integration of certain spline function, the derivative of certain spline function respectively. We further explore the roles of spline functions in the first-order one variable grey differential equation (abbreviated as GM (1, 1)) modelling. Particularly, we propose a distance measure between optimally data-fitted spline functional and constraint functional of GM (1, 1) and in terms of the distance measure for seeking the optimal constraint functional for GM (1, 1) modelling. Finally, we explore the possibility to relax the strictly positive discrete data sequence assumption extend to arbitrary discrete data sequence assumption for grey differential equation modelling in terms of optimally data-fitted spline functional.

4.1.2 A REVIEW ON SPLINE FUNCTION THEORY

Spline functions have been used for interpolations or smoothing data in engineering, science and other fields. Seeking a spline function for a given data set is a problem of variation calculus: selecting the best function to minimize an appropriately specified object functions with respect to given data information. Or in another words, it is a selection of the minimizers of suitable measures of roughness subject to the interpolation constraints. Smoothing splines may be viewed as generalizations of interpolation splines where the functions are determined to minimize a weighted combination of the average squared approximation error over observed data and the roughness measure.

4.1.2.1 Concept of Polynomial splines

The class of polynomial splines plays very flexible roles in the applications of interpolation and smoothing of data. For example, cubic Herrmite spline is used in the developments in Li, G. D., Yamaguchi, D. and Nagai, M. (2005). If the order of polynomial is high, severe oscillations often appear-particularly when the order of polynomial spline $m \geq 3$. Therefore, sometimes, we prefer to work with piecewise spline with lower order, for example, Guo, R. and Love, C. E. (2006).

Therefore, sometimes, we prefer to work with piecewise spline with lower order, for example, Guo, R. and Love, C. E. (2006). Before we end this general commenting paragraph, we give a piecewise definition of polynomial spline function. For a given interval $[a,b]$, we define a partition, denoted as $\Gamma[a,b]=\{a=x_0,x_1,\cdots,x_k=b\}$ where $x_0<x_1<\cdots<x_k$ Partition $\Gamma[a,b]$ divides the interval $[a,b]$ into k subintervals, denoted as $I_i=[x_i,x_{i+1}), i=0,1,\cdots,k-1$, let m be the order of the polynomials, denoted by $\{p_0(x),p_1(x),\cdots,p_\pi(x)\}$ with $p_j(x)=a_j^m+a_j^{m-1}x+\cdots+a_j^0x^m$ or equivalently, $d^{(m+1)}p_i(x)/dx^{m+1}=0$. Then we call the functional class or the space of piecewise polynomial splines

$$PS_m(\Gamma[a,b])=\{f(x):f(x)=p_j(x),x\in[x_j,x_{j+1+m}),j=0,1,\cdots,\pi\} \tag{4.1}$$

From the definition above we gain great flexibility by going over from polynomials to piecewise polynomials. However, we may lose the smoothness feature, which could be patched by adding smoothing constraints between joint points of two polynomials in $PS_m(\Gamma[a,b])$.

4.1.2.2 Nonparametric Regression Splines

Spline function is also treated as nonparametric regression models, say, Wahba (1990), Wahba (1990) and Crainceanu et al(2005),etc. For example, a nonparametric regression model takes the form

$$y(t_i)=\phi(t_i)+\varepsilon_i, i=1,2,\cdots,n \tag{4.2}$$

Where $0\le t_0<t_1<\cdots<t_{n-1}<t_n=1, \varepsilon_i \overset{i.i.d}{\sim} N(0,\sigma^2), i=1,2,\cdots,n.\{t_0,t_1,\cdots,t_n\}$, the partition on unit interval $[0,1]$ is called the knots. A smoothing spline estimate $\hat{g}(\cdot)$ minimizes an object function constituted by two contradictory criteria: residual sum of square-the fidelity criterion and the natural measure of smoothness, $\int_0^1\left[f^{(m)}(t)\right]^2dt, i.e.,$

$$\min_{f,k}\left\{\frac{1}{n}\sum_{i=1}^{n}(y_i-f(t_i))^2+\tau^2\int_0^1\left[f^{(m)}(t)\right]^2dt\right\} \tag{4.3}$$

Function f is a member of the functional class: $W_2^m=\{\phi:\phi\in C^{m-1}[0,1],\phi^{(m)}\in L_2[0,1]\}$. Parameter τ^2 controls the balance between the fidelity and the smoothness. The solution is unique and a natural polynomial spline of order $2m-1$ with knots $\{t_0,t_1,\cdots,t_n\}$

4.1.2.3 A Thin-Plate Regression Spline Model

Recently, low-rank thin-plate splines obtain substantial attention, which takes the form

$$\phi(x,\theta) = \beta_0 + \beta_1 x + \sum_{k=1}^{n} \delta_k |x - q_k|^3 \tag{4.4}$$

Where $\theta = (\beta_0, \beta_1, \delta_1, \delta_2, \cdots, \delta_k)^T$ is the parameter vector (the coefficients of regression model, and sample points, $q_1 < q_2 < \cdots < q_k$ defines the knots of the thin-plate spline and $\{q_k\}$ is the sample quantile of x's, which is $k/(K+1)$. The number of knots must be large enough (typically 5 to 20). The object function is

$$\sum_{i=1}^{n} (y_i - \phi(x_i,\theta))^2 + \frac{1}{\lambda}\theta^T \Lambda \theta \tag{4.5}$$

Where λ is the smoothing parameter and $\Lambda_{(n+2)\times(n+2)}$ is penalty matrix (semi-definite)

$$\Lambda = \begin{bmatrix} 0_{2\times 2} & 0_{2\times K} \\ 0_{K\times 2} & G_{K\times K} \end{bmatrix} \tag{4.6}$$

The sub-matrix $G_K = (g_{ij}), g_{ij} = |q_i - q_j|^3, i, j = 1,2,\cdots,K.$

4.1.2.4 A Bayesian Thin-Plate Spline Model

There are different proposals on Bayesian spline models in the spline function literature. What we present here follows Crainceanu et al (2005). The mixed model is

$$Y = X\beta + Z_K\delta + \varepsilon \tag{4.7}$$

With variance-covariance matrix

$$COV\begin{pmatrix} \delta \\ \varepsilon \end{pmatrix} = \begin{bmatrix} \sigma_\delta^2 G_K^{-1} & 0 \\ 0 & \sigma_\varepsilon^2 I_n \end{bmatrix} \tag{4.8}$$

Where

$$Y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \quad X = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_n \end{bmatrix} \quad Z_K = \begin{bmatrix} |x_1 - q_1| & |x_1 - q_2| & \cdots & |x_1 - q_1| \\ \vdots & \vdots & \ddots & \vdots \\ |x_K - q_1| & |x_K - q_2| & \cdots & |x_K - q_K| \end{bmatrix} \tag{4.9}$$

A re-parameterization with

$$\gamma = G_K^{1/2}\delta, Z = Z_K G_K^{-1/2} \tag{4.10}$$

Leads to

$$Y = X\beta + Z\gamma + \varepsilon \tag{4.11}$$

With variance-covariance matrix

$$COV\begin{pmatrix} \gamma \\ \varepsilon \end{pmatrix} = \begin{bmatrix} \sigma_\gamma^2 I_K & 0 \\ 0 & \sigma_\varepsilon^2 I_n \end{bmatrix} \tag{4.12}$$

Now, we can adopt a Bayesian route by placing priors on parameters. The parameter vector given the precision parameter vector $(\gamma, \varepsilon)^T$ as gamma prior, the prior takes the form $(\beta, \gamma)^T \sim N((\beta, \gamma)^T, \Sigma)$

4.1.3 Distribution filtered vector of GM (1,1) Model

GM (1, 1) Model, as an abbreviation for one variable first order grey differential equation. Except the standard formation proposed by Deng (1985), we will format GM (1, 1) Model in matrix form. Let

$$Y = \begin{bmatrix} x^{(0)}(1) \\ x^{(0)}(2) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} \quad A = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix} \quad A_w = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{2} \end{bmatrix} \tag{4.13}$$

Then GM (1, 1) model can be expressed as

$$Y = 1_{n\times 1}\alpha + [-A_\omega AY]\beta + \varepsilon \tag{4.14}$$

Subject to the differential equation constraint $dx^{(1)}/dt = \alpha - \beta x^{(1)}$, this will lead to

$$A\hat{Y} = \left(x^{(0)} - \frac{a}{b}\right)E + \frac{a}{b} \tag{4.15}$$

Where

$$\hat{Y} = \begin{bmatrix} \hat{x}^{(0)}(1) \\ \hat{x}^{(0)}(2) \\ \vdots \\ \hat{x}^{(0)}(n) \end{bmatrix} \quad E = \begin{bmatrix} 1 \\ \exp(-b(2-1)) \\ \vdots \\ \exp(-b(n-1)) \end{bmatrix} \tag{4.16}$$

Thus

$$\hat{Y} = \left(x^{(0)} - \frac{a}{b}\right)\left(A^T A\right)^{-1} A^T E + \frac{a}{b} 1_{n\times 1} \tag{4.17}$$

With $(a,b)^T$ as estimate for parameter vector $(\alpha,\beta)^T$ based on equation (4.14):

$$\begin{pmatrix} a \\ b \end{pmatrix} = \left[[1,-A_w AY]^T [1,-A_w AY]\right]^{-1} [1,-A_w AY]^T Y \tag{4.18}$$

Thus the distribution for filtered vector $\hat{Y}$ can be derived in terms of equation (4.17) and (4.18).

4.1.4 Mixed spline-GM (1, 1) models

There are various ways to mix spline model and GM (1, 1) model to achieve a better data-assimilating effect.

4.1.4.1 A spline-GM (1, 1) model

It is a well-known fact that polynomial spline or thin-plate spline can fit the original data sequence, denoted as $X^{(0)}$. Let us denote the spline fitted as $\hat{p}(t)$. Then the residual sequence after fitting spline function denoted by $R = \{r(i) : r(i) = x^{(0)}(i) - \hat{p}(i)\}, i = 1,2,\cdots,n.$ Note that the residual data sequence is real-valued sequence which is no longer strictly positive data sequence required for GM (1, 1) model fitting.

Recall that for any real-valued function γ, then $\gamma = \gamma^+ - \gamma^-$ and $|r| = \gamma^+ + \gamma^-$ where:

$$\gamma^+ = \begin{cases} r & if \quad r > 0 \\ 0 & if \quad r \le 0 \end{cases} \tag{4.19}$$

And

$$r^{-} = \begin{cases} 0 & if \quad r \geq 0 \\ -r & if \quad r < 0 \end{cases} \tag{4.20}$$

Notice that the formula:

$$r = 2r^{+} - |r| \tag{4.21}$$

Will "recover" the original value of r once the values of r^{+} and $|r|$ are GM (1, 1) model-fitted respectively. Denoted the GM (1, 1)-fitted residual model as $\hat{r}(t)$, then the spline-GM (1, 1) model is given by

$$\hat{x}^{(0)}(i) = \hat{p}(i) + \hat{r}(i), i = 1,2,\cdots,n \tag{4.22}$$

4.1.4.2 A spline-AGO-GM (1, 1) model

Another way for a better GM (1, 1) modelling is to replace the background values $Z^{(1)} = \left(z^{(0)}(1), z^{(1)}(2), \cdots, z^{(1)}(n)\right)$,

$$z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right] \tag{4,23}$$

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,\cdots,n \tag{4.24}$$

By

$$z^{(1)}(k) = \int_{0}^{k} \hat{p}(u)du, k = 1,2,\cdots,n \tag{4.25}$$

Then the model becomes

$$Y = \left[\underline{1}_{N\times 1}, -Z\right] \cdot \begin{bmatrix} a \\ b \end{bmatrix} + \varepsilon \tag{4.26}$$

Where

$$Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}, Z = \begin{bmatrix} \int_0^2 \hat{p}(u)du \\ \int_0^3 \hat{p}(u)du \\ \vdots \\ \int_0^n \hat{p}(u)du \end{bmatrix} \tag{4.27}$$

subject to the constraint differential equation $dx^{(1)}/dt = a - bx^{(1)}$.

4.1.4.3 A GM (1, 1)-Spline Model

The third way to mix GM (1, 1) model and spline together is to fit GM (1, 1) model first and then fit the spline to the GM (1, 1) model residuals. Denote the GM (1, 1) modelling on the original data sequence as $g(t)$ and the residual spline as $S_g(t)$, then

$$\hat{x}^{(0)}(i) = g(i) + S_g(i), i = 1,2,\cdots,n \tag{4.28}$$

4.1.5 Conclusions

We reviewed two sub-families of the spline functions: polynomial spline and thin-plate spline and also the standard GM (1, 1) model (in matrix form). It is also necessary to point out that the spline fitting may pave a way for a better GM (2, 1) modelling as well as grey relational analysis. For example, if two observational data sequence $X_1^{(0)}$ and $X_2^{(0)}$, the distance measure

$$d\left(\hat{X}_1^{(1)}, \hat{X}_2^{(1)}\right) = \frac{\int_0^n \left|\hat{p}_{X_1}(t) - \hat{p}_{X_2}(t)\right|^2 dt}{\sqrt{\int_0^n \left|\hat{p}_{X_1}(t)\right|^2 dt}\sqrt{\int_0^n \left|\hat{p}_{X_2}(t)\right|^2 dt}} \tag{4.29}$$

Can be considered a candidate to replace current grey coefficient measures in the grey mathematics literature (Deng, 2002)

4.2 Spline function in grey relation analysis and applications to system reliability improvements

(YAN HONG CUI, RENKUAN GUO, AND TIM DUNNE)

4.2.1 Introduction

Relation analysis between (or among) different factors (or variables) is always a focus point in engineering, medical science, financial market research and other business applications. Under probabilistic framework, correlation, concordance, and copulas are developed. However, if we are facing challenge of sparse data availability, the probabilistic based relational measure will not work due to very small sample size. In today's highly competitive business environment, there is often a possibility that not adequate samples are available for system designs and improvements because of the fast-changing in new products and shorter life-cycle.

Deng, J. L. (1982) proposed GRA concept and developed GRA indices. Later, Liu, B. and Liu, Y. K. (2002) and Wen, K. L. (2004) and others have extended GRA methodology with successful applications in many business and scientific environments.

Definition 4.2.1.1: Grey relation analysis □GRA□ is a data processing method to categorize the correlation extent of compared sequences and a certain reference sequence in a system with uncertain information

In this paper, we first review the Liu's absolute degree of grey relation (abbreviated as ADGR) and the relative degree of grey relation (RDGR). We do notice that in the latest book in English version (Liu, S., and Lin, Y. (2006)) he used the term grey incidence, however, we still used the old term grey relation as Li, G. D., Yamaguchi, D. and Nagai, M. (2005), because term grey relation was already used in grey research community more than twenty years since Deng's initiation (1982), and their general properties. Based on a brief review on the spline function theory, we propose a spline-ADGR and spline RDGA because we realize that spline function fitting is no longer an extra burden for grey researcher today due to the advancements in computation power and commercial and freeware software availability, for example, spline function toolbox in Matlab.

Hinted by concordance measure and copula concept in probability theory, we explore grey concordance concept and propose a grey concordance index.

4.2.2 A Review on Liu's GRA Indices

As many authors emphasized that, for example, Li, G. D., Yamaguchi, D. and Nagai, M. (2005),Wen, K. L. (2004)and Liu, S., and Lin, Y. (2006), that grey relation analysis is to investigate the geometric similarity between two data sequences. Since Deng's initial GRA index (1985), there are many extensions to GRA index, for example, Won, Wen, Liu etc.

The absolute degree of grey relation (abbreviated as ADGR) and relative degree of grey relation (abbreviated as RDGR) are typically worth to draw attention because their geometric character.

For given two discrete data sequence (with equal sample size), denoted by $X_1^{(0)} = \{x_1^{(0)}(i)\}_{i=1}^{n}$ and $X_2^{(0)} = \{x_2^{(0)}(i)\}_{i=1}^{n}$ respectively. Liu, S., and Lin, Y. (2006) compare the so-called (absolute) zigzagged lines defined as $X_1^{(0)} = \{x_1^{(0)}(i) - x_1^{(0)}(1)\}_{i=1}^{n}$ and $X_2^{(0)} = \{x_2^{(0)}(i) - x_2^{(0)}(1)\}_{i=1}^{n}$ respectively. These zigzagged curves start with common value zero Quantity:

$$s_i^a = \int_1^n \left(x_i^{(0)}(t) - x_i^{(0)}(1)\right)dt, i = 1,2 \tag{4.30}$$

Then the ADGR is defined as:

$$\varepsilon_{ij}^a = \frac{1 + |s_i^a| + |s_j^a|}{1 + |s_i^a| + |s_j^a| + |s_i^a - s_j^a|} \tag{4.31}$$

And the relative zigzagged lines defined as $X_1^{(0)} = \{x_1^{(0)}(i) / x_1^{(0)}(1)\}_{i=1}^{n}$ and $X_2^{(0)} = \{x_2^{(0)}(i) / x_2^{(0)}(1)\}_{i=1}^{n}$ respectively, which are scaled data sequences. Notice that the relative zigzagged curves start with common value one. Let us define the relative cumulative area under the relative zigzagged curves

$$s_i^r = \int_1^n \left(x_i^{(0)}(t) / x_i^{(0)}(1)\right)dt, i = 1,2 \tag{4.32}$$

which is actually corresponding to total 1-AGO of the relative values of the original data sequence. The RDGR is defined as

$$\varepsilon_{ij}^{r} = \frac{1+\left|s_i^r\right|+\left|s_j^r\right|}{1+\left|s_i^r\right|+\left|s_j^r\right|+\left|s_i^r - s_j^r\right|} \tag{4.33}$$

As a general axiomatic framework, the grey analysis index, denoted as $\gamma(x_i, x_j)$, must satisfy the following four axioms .Wen, K. L. (2004)

Normality $\gamma(x_i, x_j) \in [0,1]$;

Duality $\gamma(x_i, x_j) = \gamma(x_j, x_i), i \neq j$;

Wholeness $\gamma(x_i(k), x_j(k)) \neq \gamma(x_j(k), x_i(k)), i \neq j$; and

Closeness the term $\left|x_j(k), x_i(k)\right|$ play the key role in defining any grey relation grade (i.e., index).

As Liu, S., and Lin, Y. (2006) demonstrated, the two GRA (ADGR, RDGR) indices satisfy the four axioms. In principle, we notice that ADGR and RDGR are both geometrically motivated. However, for computational conveniences, the integration calculations of s_i^a and s_i^r are both carried in terms of trapezoid rule which leads to 1-AGO (from 2 to (n-1)) formation with correction term $x_i^{(0)}(n)/2$.

Li, G. D., Yamaguchi, D. and Nagai, M. (2005), noticed that there could be some logical link between spline function and AGO and IAGO, as pointed by Wahba (1991) too. Therefore, we will briefly review spline function theory in the next section to prepare our extension or modification to Liu's ADGR and RDGR.

4.2.3 A Review on Spline Function Theory

Seeking a spline function for a given data set is a problem of variation calculus: seeking the best function to minimize an appropriately specified object functions with respect to given data information. Or in another words, it is a selection of the minimizers of suitable measures of roughness subject to the interpolation constraints.

4.2.3.1 Concept of Polynomial splines

We prefer to work with piecewise spline with lower order, for example, Guo and Love (1996). Before we end this general commenting paragraph, we give a piecewise definition of polynomial spline function. For a given interval $[a,b]$, we define a partition, denoted as $\Gamma[a,b]=\{a=x_0,x_1,\cdots,x_k=b\}$ where $x_0<x_1<\cdots<x_k$ Partition $\Gamma[a,b]$ divides the interval $[a,b]$ into k subintervals, denoted as $l_i=[x_i,x_{i+1}),i=0,1,\cdots,k-1$, let m be the order of the polynomials, denoted by $\{p_0(x),p_1(x),\cdots,p_\pi(x)\}$ with $p_j(x)=a_j^m+a_j^{m-1}x+\cdots+a_j^0x^m$ or equivalently, $d^{(m+1)}p_i(x)/dx^{m+1}=0$. Then we call the functional class or the space of piecewise polynomial splines

$$PS_m(\Gamma[a,b])=\{f(x):f(x)=p_j(x),x\in[x_j,x_{j+1+m}),j=0,1,\cdots,\pi\} \tag{4.34}$$

4.2.4 Spline-Modification to ADGR and RDGR

Wahba (1991) pointed out that "A practical application of this result we will return is the estimation of $f(t)$ given data:

$$y_i=f(t_i)+\varepsilon_i,i=1,2,\cdots,n \tag{4.35}$$

One can take the smoothing spline for the data and use its derivative as an estimate of f'.

Therefore, we will fit spline functions to the shifted data sequences $X_1^{(0)}=\{x_1^{(0)}(i)-x_1^{(0)}(1)\}_{i=1}^n$ and $X_2^{(0)}=\{x_2^{(0)}(i)-x_2^{(0)}(1)\}_{i=1}^n$ respectively, denoted them as $\varphi_1^S(t)$ and $\varphi_2^S(t)$ respectively, In other words, for data

$$x_1^{(0)}(i)-x_1^{(0)}(1)=\varphi_1(t_i)+\varepsilon_i,i=1,2,\cdots,n \tag{4.36}$$

which leads to

$$s_1^a=\int_1^n\left(x_1^{(0)}(i)-x_1^{(0)}(1)\right)dt=\int_1^n\varphi_1(t)dt \tag{4.37}$$

Similarly, for data

$$x_2^{(0)}(i)-x_2^{(0)}(1)=\varphi_2(t_i)+\varepsilon_i,i=1,2,\cdots,n \tag{4.38}$$

which leads to

$$s_2^a = \int_1^n \left(x_2^{(0)}(i) - x_2^{(0)}(1)\right)dt = \int_1^n \varphi_2(t)dt \tag{4.39}$$

Thus, the spline-ADGR (abbreviated as SADGR) will take the form

$$\varepsilon_{ij}^a = \frac{1 + \left|\int^{t'} \varphi_i(t)dt\right| + \left|\int^{t'} \varphi_j dt\right|}{1 + \left|\int^{t'} \varphi_i(t)dt\right| + \left|\int^{t'} \varphi_j(t)dt\right| + \left|\int^{t'} (\varphi_i(t) - \varphi_j(t))dt\right|} \tag{4.40}$$

Note that

$$\left|\int^{t'} (\varphi_i(t) - \varphi_j(t))dt\right| \le \int^{t'} \left|\varphi_i(t) - \varphi_j(t)\right|dt \tag{4.41}$$

Therefore, we can enlarge the denominator of equation 4.40 and accordingly adjust other two integrations, i.e., $\int^{t'} |\varphi_i(t)|dt$ and $\int^{t'} |\varphi_j(t)|dt$, respectively, a modified version of SADGR is

$$\varepsilon_{ij}^a = \frac{1 + \int^{t'} |\varphi_i(t)|dt + \int^{t'} |\varphi_j(t)|dt}{1 + \int^{t'} |\varphi_i(t)|dt + \int^{t'} |\varphi_j(t)|dt + \int^{t'} |\varphi_i(t) - \varphi_j(t)|dt} \tag{4.42}$$

For the relative (i.e., scaled) data sequence

$$x_i^{(0)}(k)/x_i^{(0)}(1) = \psi_1(t_k) + \varepsilon_k, k = 1,2,\cdots,n \tag{4.43}$$

which leads to

$$s_i^r = \int_1^n \left(\frac{x_1^{(0)}(t)}{x_1^{(0)}(1)}\right)dt = \int_1^n \psi_i(t)dt \tag{4.44}$$

Thus a modified version of spline-RDGR (abbreviated as SRDGR) takes the form

$$\varepsilon_{ij}^r = \frac{1 + \int^{t'} |\psi_i(t)|dt + \int^{t'} |\psi_j(t)|dt}{1 + \int^{t'} |\psi_i(t)|dt + \int^{t'} |\psi_j(t)|dt + \int^{t'} |\psi_i(t) - \psi_j(t)|dt} \tag{4.45}$$

4.2.5 Further Extensions

4.2.5.1 Local GRA Grade

A close examination on equation on both equation (4.40) and equation (4.42), we should realize that both GRA grades are overall measures over interval [1, n]. However, due to the convenience of replacing 1-AGO like summation by the integral of a continuous function, it is possible to define a local GRA index:

$$\varepsilon_{ij}^{a}(u,u)=\frac{1+\int^{u}|\varphi_i(t)|dt+\int^{u}|\varphi_j(t)|dt}{1+\int^{u}|\varphi_i(t)|dt+\int^{u}|\varphi_j(t)|dt+\left|\int^{u}\varphi_i(t)dt-\int^{u}\varphi_j(t)dt\right|} \tag{4.46}$$

Where $u \in [0,1]$. However, the second parameter u can be replaced by v, both within [1, n] but not necessarily equal each other. This may create the GRA analysis for unequal length data sequence. Note that $\varepsilon_{ij}^{a}(u,u)=\varepsilon_{ji}^{a}(u,u), i \neq j$.

4.2.5.2 Spline-Grey Relation Index

We notice that both ADGR and RDGR are transformed data sequences. As our basic aim is to explore the geometric similarity between the geometric curves underlying the discrete data sequences, it is not necessary to engage transformations, particularly, when spline curves are fitted. It is true that the spline curves to be fitted are not necessarily the true geometric curves for the data sequences, however, these splines are at least functions in average sense conditional the sample data (i.e., localized expected curves). Therefore, we define a spline-grey relation index with respect to original data sequences, denoted by $X_i^{(0)}=\{x_i^{(0)}(k)\}_{k=1}^{n}$ and $X_j^{(0)}=\{x_j^{(0)}(k)\}_{k=1}^{n}$, and let the splines be $\phi_i(t)$ and $\phi_j(t)$ respectively. The spline-grey relation index takes the form

$$\varepsilon_{ij}^{s}=\frac{1+\int^{n}|\phi_i(t)|dt+\int^{n}|\phi_j(t)|dt}{1+\int^{n}|\phi_i(t)|dt+\int^{n}|\phi_j(t)|dt+\int^{n}|\phi_i(t)-\phi_j(t)|dt} \tag{4.47}$$

4.2.5.3 Grey self-similarity grade

A close examination can alert that under probabilistic framework, autocorrelation and self-similar measure were developed to measure the internal relation between different pieces of the sample information. Grey relation analysis does not provide such a kind measure yet because of the sparse data availability. However, in terms of spline function, a grey self-similarity index may be defined. Recall that in probability theory, given a stationary stochastic process $\{X_n, n=1,2,\cdots\}$

in wide sense, for each integer $m = 1,2,\cdots$, let $\{X_k^{(m)}, k = 1,2,\cdots\}$ be a subsequence of the original process. The process is H-similar if

$$X_k = \sum_{i=(k-1)m+1}^{km} X_i \tag{4.48}$$

Holds for all $m > 0$

For the grey self-similar analysis, we propose an (δ, Δ)-self-similar grey index where $\delta > 0$ as allowance and $\Delta > 1$ as shift parameter.

$$\varepsilon_\delta(u,\Delta) = \frac{1 + (\Delta - 1)^{-1}\left[\left|\int^{\Delta}\phi(t)dt\right| + \left|\int_t^{t+\Delta}\phi(t)dt\right|\right]}{1 + (\Delta - 1)^{-1}\left[\left|\int^{\Delta}\phi(t)dt\right| + \left|\int_t^{t+\Delta}\phi(t)dt\right| + \left|\int^{\Delta}\phi(t)dt - \int_t^{t+\Delta}\phi(t)dt\right|\right]} \tag{4.49}$$

If $\varepsilon_\delta(u,\Delta) < \delta$ for a given positive number $\delta > 0$, then we call the data sequence $X^{(0)} = \{x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\}$ as $\delta - (u,\Delta)$-self-similar.

4.2.5.4 Grey distance metric

In terms of the fitted spline function, we can define a grey distance metric. For example, if two observational data sequence $X_1^{(0)}$ and $X_2^{(0)}$ the grey distance metric takes the form:

$$d\left(X_1^{(0)}, X_2^{(0)}\right) = \frac{\int_1^n \left|\phi_{X_0}(t) - \phi_{X_1}(t)\right| dt}{\int_1^n \left|\phi_{X_0}(t)\right| dt + \int_1^n \left|\phi_{X_1}(t)\right| dt} \tag{4.50}$$

which can be also considered a candidate to replace current grey coefficient measures in the grey mathematics literature (Deng, 2002). Note that $d\left(X_1^{(0)}, X_2^{(0)}\right) = 0$ if and only if $\phi_{X_0}(t) = \phi_{X_1}(t)$ and $d\left(X_1^{(0)}, X_2^{(0)}\right) = 1$, if and only if $\phi_{X_0}(t) = -\phi_{X_1}(t)$. In general $d\left(X_1^{(0)}, X_2^{(0)}\right) \in [0,1]$.

4.2.5.5 Illustrative Examples

For comparison purposes, we use the data sequences in Liu and Lin (2006). The original data sequences X_0 and X_1 sequence are listed in the Column 2 and 3 respectively. Column 4 and 5 list the shifted data (where shifted values are $x_0^{(0)}(1)$ and $x_1^{(0)}(1)$ respectively). Column 6 and 7 list the scaled data (where the scaled values are $x_0^{(0)}(1)$ and $x_1^{(0)}(1)$ respectively). Finally, the Column 8 and 9 record the scaled-shifted data-minus 1.0 from scaled data).

Table 4.2.1 Original and Transformed Data Sequences.

No	X_0	X_1	Shifted X_0	Shifted X_1	Scaled X_0	Scaled X_1	Scaled& Shifted X_0	Scaled& Shifted X_1
1	10	46	0	0	1.0000	1.0000	0.0000	0.0000
2	9	58	-1	12	0.9000	1.2609	-0.1000	0.2609
3	15	70	5	24	1.5000	1.5217	0.5000	0.5217
4	14	77	4	31	1.4000	1.6739	0.4000	0.6739
5	14	84	4	38	1.4000	1.8261	0.4000	0.8261
6	15	91	5	45	1.5000	1.9783	0.5000	0.9783
7	16	98	6	52	1.6000	2.1304	0.6000	1.1304

Table 4.2.2 Coefficients of the fitted polynomial splines

Coef	Original data		Shifted data		Scaled data		Scaled-shifted data	
t^6	0.051389	0.027778	0.051389	0.027778	0.051389	0.0006093	0.051389	0.00060386
t^5	-1.3292	-0.70833	-1.3292	-0.70833	-0.13292	-0.015399	-0.13292	-0.015399
t^4	13.66	7.1528	13.66	7.1528	1.366	0.1555	1.366	0.1555
t^3	-70.521	-36.042	-70.521	-36.042	-7.0521	-0.78351	-7.0521	-0.78351
t^2	189.29	92.819	189.29	92.819	18.929	2.0178	18.929	2.0178
t	-242.15	-101.25	-242.15	-101.25	-24.215	-2.2011	-24.215	-2.2011
Const.	121	84	111	38	12.1	1.8261	11.1	0.82609

The grey relation indices in Liu and Lin (2006) as well as extended versions are calculated based on the data in Table 4.2.1 are listed in Table 4.2.3.

Table 4.2.3 grey relation index comparisons

Name	Liu & Lin	Equation 4.40+4.42	Equation 4.45	Equation 4.45	Equation 4.47	Equation 4.50
Absolute	0.5581	0.55949	*	*	0.58624	0.70712
Relative	0.78	0.55100	0.8997	0.77763	*	*

4.2.6 GRA in System Reliability Improvements

In this section, we explore the spline role in defining grey relational analysis indices. The work is of critical importance to system reliability improvements because today's product design and production process are both short-lifecycled and thus the data information available for design and improvements will be sparse. Therefore, the grey relation indices will help to reveal necessary information on the relation of different subsystems (in terms of SADGR or SRDGR and their modified versions) and even the relation between different development stages (using grey self-similar index) under a little data information for system engineering team. In the future, we will explore the basic characters of different spline-GRA indices proposed and their suitable application range. We have to point out that the grey relation indices in the literature often utilized certain transformation schemes for certain reasons, however, the grey relation indices in terms of transformed data may not truly reveal the geometric closeness of the original data sequences. We believe that the spline function usage may not require any transformations. Furthermore, we need to investigate the invariance issue of the grey relation indices under monotonic transformation.

4.3 Summary

In this chapter, we reviewed a couple of class of spline functions under certain optimal conditions. Then we re-examined the nature and the optimality of the two critical data operations proposed by Deng (2002), accumulative generating operation (abbreviated as AGO) and the inverse accumulative generating operation (abbreviated as IAGO) and then the exchangeability between AGO, IAGO and the integration of certain spline function, the derivative of certain spline function respectively. We further explored the roles of spline functions in the first-order one variable grey differential equation (abbreviated as GM (1, 1)) modelling. Particularly, we proposed a distance measure between optimally data-fitted spline functional, constraint functional of GM(1,1) and in terms of the distance measure for seeking the optimal constraint functional for GM(1,1) modelling. Finally, we explore the possibility to relax the strictly positive discrete data sequence assumption extend to arbitrary discrete data sequence assumption for grey differential equation modelling in terms of optimally data-fitted spline functional.

Chapter 5. Coupling Principle in Grey Differential Equation Modelling

5.1 Coupling Idea of Grey Differential Equation Model.

Because of the characteristic of small sample analysis, we simply know the GDE model is just a data-oriented modelling process (Do not following any distribution). The most important part of GDE modelling is justifying appropriated model utilized available data information. The essential of GDE model fitting is quite similar to fitting many possible regression equations to data sets. Then, after analysis of GDE modelling procedure we could abstract GDE modelling into 3-step approach: step1, the AGO or IAGO (abbreviations of accumulative generating operation and inverse of accumulative generating operation respectively) of the original data sequence, step2: fitting a regression model for seeking parameter estimation in GDE model, step3: use the solution to whitening differential equation for filtering or prediction.

If we examine the coupling feature in GM(1,1) in detail, we will find that a GM(1,1) model actually starts with a motivated differential equation, then the coupled regression model is specified in the discretized or discretized form of the motivated differential equation, in turn, in terms of coupling regression model. The parameters specifying the motivated differential equation are estimated under least-squares estimation. Furthermore, the solution to the motivated differential equation (or the discretized solution) equipped with data-assimilated parameter estimates, is used for system analysis or prediction. We should emphasize here that the way a GM (1, 1) model uses system sampling information to solve the motivated differential equation is different from that in common algorithms for solving a differential equation numerically. In a GM (1, 1) modelling, we will obtain a closed-form functional solution (i.e., the primitive function) to the motivated differential equation with optimal data-assimilated parameters. The availability of the closed-form primitive function $x(t)$ will provide the great conveniences in the further investigation on the system under study. We acknowledge that the idea of obtaining a closed-form solution to the motivated differential equation was suggested by the founder of Grey Mathematics, Deng (1985). (Guo, R. (2007d).)

Table 5.1.1. Coupling Rule in univariate GM (1, 1) Model

Term	*Motivated DE*	*Coupled REG*
Discretization rule between MDE and CREG		
Intrinsic feature	*Continuous*	*Discrete*
Independent Variable	t	k
1st-order Derivative	$dx(t)/dt$	$\Delta x(k)=x(k)-x(k-1)$
pst-order Derivative	$d^{(p)}x(t)/dt^{p}$	$\Delta^{p}x(k)=\Delta^{p-1}x(k)-\Delta^{p-1}x(k-1)$
Primitive function	$x(t)$	$x(k)$
Model Formation	$\frac{dx(t)}{dt}=\alpha+\beta x(t)$	$\Delta x(k)=\alpha+\beta\cdot x(k)+\varepsilon_k$
Parameter Coupling		
Parameter	(α,β)	(a,b)
Dynamics (Solution)	$x(t)=\left[x(0)-\frac{\alpha}{\beta}\right]e^{\beta t}+\frac{\alpha}{\beta}$	$\hat{x}(k+1)=\left[x(1)-\frac{a}{b}\right]e^{bk}+\frac{a}{b}$
Filtering (Prediction)	$dx(t)/dt=[\alpha-\beta dx(0)/dt]e^{\beta t}$	$\Delta\hat{x}(k+1)=\hat{x}(k+1)-\hat{x}(k)$

A fundamental idea list in Table 5.1.1 is that the differenced observations are treated as the approximated derivatives of the dynamic law $x(t)$, but the modelling is still required to return back to the derivative level because that is the observational continuum. (R. Guo, D. Guo, T. Dunne and C. Thiart (2007)

From the procedure of 3-step approach GDE modelling, we can easily realize the GM (1, 1) model is just a data AGO treatments and a simple regression model coupled with the whitening differential equation model.

At beginning, we explore two GM (1, 1) modelling examples, which will lead us to find the relation between GM (1, 1) model and the regression model built in GM (1, 1) modelling procedure.

Example 5.1.1: Let $X^{(0)}=(5.081,4.611,5.1177,9.3775,11.057,11.3524)$ following the GM (1, 1) modelling procedure:

First, take 1-time accumulating generation operation of $X^{(0)}$, to generate the data sequence $X^{(1)}=(5.081,9.692,14.8097,24.1872,35.2442,46.5966)$

Second, we utilize a simple multiple linear regression to generate value $\hat{X}^{(1)} = (5.081, 10.2129, 16.609, 24.5806, 34.5157, 46.8981)$. The correlation coefficient value of $X^{(1)}$ is $r^2 = \dfrac{\sum (y_{est} - \bar{y})^2}{\sum (y - \bar{y})^2} = \dfrac{\sum \left(\hat{X}^{(1)} - \overline{X}^{(1)}\right)}{\sum \left(\hat{X}^{(1)} - \overline{X}^{(1)}\right)} \approx 0.9686$, then take the IAGO

$\hat{X}^{(0)} = (5.081, 5.13195, 6.39606, 7.97156, 9.93513, 12.3824)$; The average relative error is 14.0978%. The accuracy is 85.9022%

Example 5.1.2: Let $Y^{(0)} = (2.874, 3.278, 3.337, 3.390, 3.679)$, following the GM (1, 1) modelling procedure:

First, take 1-time accumulating generation operation of $Y^{(0)}$, to generate the data sequence $Y^{(1)} = (2.874, 6.152, 9.489, 12.879, 16.558)$

Second, we utilize a simple multiple linear regression to generate value $\hat{Y}^{(1)} = (2.874, 6.10604, 9.46059, 12.9423, 16.556)$. The correlation coefficient value of $Y^{(1)}$ is $r^2 = \dfrac{\sum (y_{est} - \bar{y})^2}{\sum (y - \bar{y})^2} = \dfrac{\sum \left(\hat{Y}^{(1)} - \overline{Y}^{(1)}\right)}{\sum \left(Y^{(1)} - \overline{Y}^{(1)}\right)} \approx 0.9999 = 1$, then we take the IAGO

$\hat{Y}^{(0)} = (2.874, 3.23204, 3.35455, 3.4817, 3.61368)$; The average relative error is 1.602175%. The accuracy is 98.397825%

Comparing with two examples, the first ones AGO data sequence $X^{(1)}$ after a simple regression, the coefficient of the determination value R-square equal to 0.9686 and the accuracy is 85.9022%. The second example, R-square equal to 0.9999 and accuracy is 98.397825%. After analysis with this two examples ,we will easy to realized that during the GDE modelling procedure, if the regression part showing a better goodness-of-fit, say, R-square value close to 1, then the GDE model should hold a superior modelling efficiency. Such phenomena which exist in the GDE model tell us, the regression model truly linked with the whitening differential equation, fitting a regression model for seeking parameter estimation in GDE model is very important, which directly effect to the GM (1, 1) modelling efficiency.

5.2 The role of difference equation in GM (1, 1)

Normally we just following the whitening differential equation to generate a solution form for calculate the final estimation value. But in here we try to simulate a solution form for GM (1, 1) modelling just using difference equation.

Given a discrete positive data sequence $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), x^{(0)}(3), \cdots, x^{(0)}(n)\right)$, the difference equation is

$$x^{(0)}(k) + \beta z^{(1)}(k) = \alpha, k = 2,3,4,\cdots,n \tag{5.1}$$

Where

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4,\cdots,n\,;\; z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right]. \tag{5.2}$$

Now we rewrite the equation (5.9) as

$$\left(1-\frac{\beta}{2}\right)x^{(1)}(k) - \left(1-\frac{\beta}{2}\right)x^{(1)}(k-1) = \alpha, k = 2,3,4\cdots,n \tag{5.3}$$

So the equation (5.11) could be written as:

$$x^{(1)}(k) - \frac{(1-\frac{\beta}{2})}{\left(1+\frac{\beta}{2}\right)}x^{(1)}(k-1) = \frac{\alpha}{\left(1-\frac{\beta}{2}\right)} \tag{5.4}$$

Because $x^{(1)}(k) - x^{(1)}(k-1) = x^{(0)}(k)$, then $x^{(1)}(k) = x^{(1)}(k-1) + x^{(0)}(k)$

So the equation (5.12) could be written as:

$$x^{(0)}(k) + \frac{\beta}{\left(1+\frac{\beta}{2}\right)}x^{(1)}(k-1) = \frac{\alpha}{\left(1+\frac{\beta}{2}\right)} \tag{5.5}$$

Then

$$x^{(0)}(k) = \frac{\alpha}{\left(1+\frac{\beta}{2}\right)} - \frac{\beta}{\left(1-\frac{\beta}{2}\right)}x^{(1)}(k-1) \tag{5.6}$$

Substituting the response value $\hat{x}^{(1)}(k-1) = \left(x^{(0)}(1) - \frac{\alpha}{\beta}\right)e^{-\beta(k-2)} + \frac{\alpha}{\beta}$ of $x^{(1)}(k-1)$ into the equation above gives that

$$x^{(0)}(k) = \frac{\alpha}{\left(1+\frac{\beta}{2}\right)} - \frac{\beta}{\left(1-\frac{\beta}{2}\right)}\left(\left(x^{(0)}(1) - \frac{\alpha}{\beta}\right)e^{-\beta(k-2)} + \frac{\alpha}{\beta}\right) \tag{5.7}$$

Then the solution form of $x^{(0)}$ is

$$x^{(0)}(k)=\frac{\beta}{1+\frac{\beta}{2}}\left(\frac{\alpha}{\beta}-x^{(0)}(1)\right)e^{-\beta(k-2)},k=2,3,4,\cdots,n \tag{5.8}$$

In here we notice the parameter (α,β) should be estimated in terms of multiple linear regression function: $(a,b)^T=\left(X^TX\right)^{-1}X^TY$.

Example 5.2.1: Let $X^{(0)}=(2.874,3.278,3.337,3.390,3.679)$, we try to estimate the parameter (α,β) following by modified equation (5.8): $\min f=\sum_{k=2}^{n}\left|\frac{\beta}{1+\frac{\beta}{2}}\left(\frac{\alpha}{\beta}-x^{(0)}(1)\right)e^{-\beta(k-2)}\right|,k=2,3,4,\cdots,n$,

then we could generate estimation value from equation: $\hat{x}^{(0)}(k)=\frac{\beta}{1+\frac{\beta}{2}}\left(\frac{\alpha}{\beta}-x^{(0)}(1)\right)e^{-\beta(k-2)},k=2,3,4,\cdots,n$

We use matlab program to calculate all the results from above question conditions.

```
function f=f(r)
f=abs((r(2)/(1+r(2)/2))*((r(1)/r(2))-2.874)*exp(-r(2)*(2-2))-
2.874)+abs((r(2)/(1+r(2)/2))*((r(1)/r(2))-2.874)*exp(-r(2)*(3-2))-
3.278)+abs((r(2)/(1+r(2)/2))*((r(1)/r(2))-2.874)*exp(-r(2)*(4-2))-
3.337)+abs((r(2)/(1+r(2)/2))*((r(1)/r(2))-2.874)*exp(-r(2)*(5-2))-
3.390)+abs((r(2)/(1+r(2)/2))*((r(1)/r(2))-2.874)*exp(-r(2)*(6-2))-3.679);
```

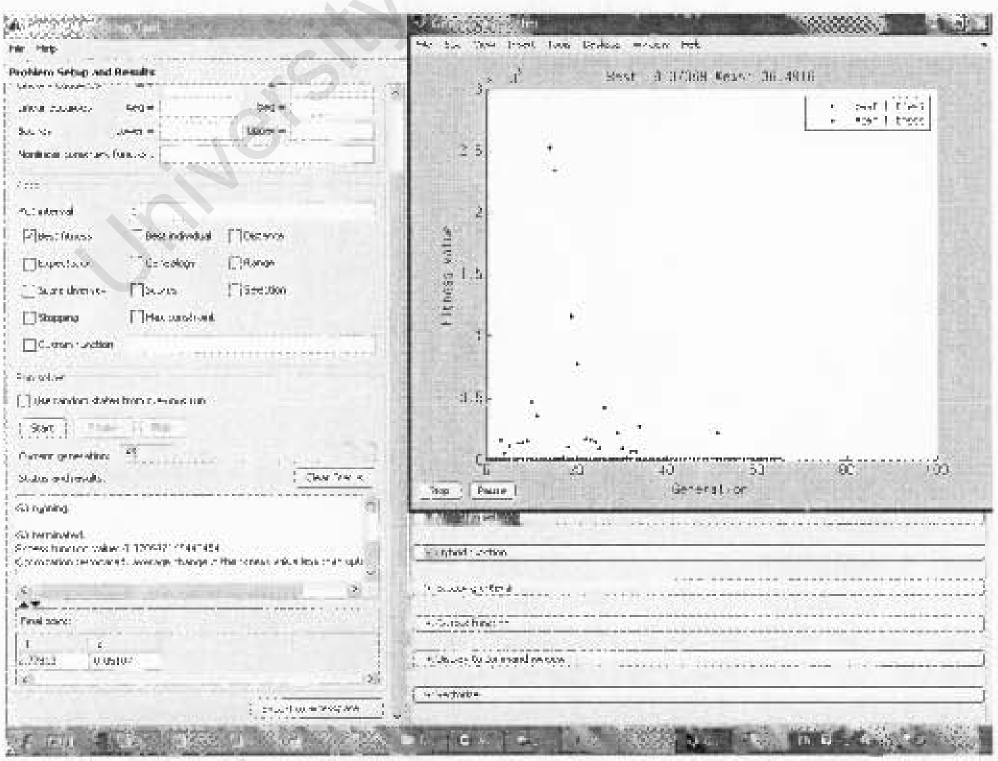

Figure 5.2.1 Genetic algorithm for generating the parameter (α,β)

$\min f = 0.37086$ $\alpha = 2.77813$ $\beta = -0.05107$, then $\hat{x}^{(0)}(k) = (3.0015, 3.1588, 3.3243, 3.4985, 3.6818)$ $X^{(0)} = (2.874, 3.278, 3.337, 3.390, 3.679)$

Calculate the Errors and Relative Errors

$$\delta(k) = x^{(0)}(k) - \hat{x}^{(0)}(k) = (0.1275, -0.1192, -0.0127, 0.1085, 0.028) \tag{5.9}$$

Relative Errors (%):

$$\Delta k = \frac{|\delta(k)|}{x^{(0)}(k)} = (4.4363\%, 3.3636\%, 0.3805\%, 3.2005\% 0.7611\%) \tag{5.10}$$

Average relative error:

$$\Delta = \frac{1}{5}\sum_{k=1}^{5} \Delta_K = 0.024284 = 2.4284\% \tag{5.11}$$

Accuracy of GM (1, 1) model

$$accuracy = 1 - \Delta = 1 - \frac{1}{5}\sum_{k=1}^{5} \Delta_K = 0.975716 = 97.5716\% \tag{5.12}$$

The same data sequence we are using for example 2.5.1 and example 2.4.2, but the result showing example 2.5.1's accuracy 97.5716% is lower than example 2.4.2's accuracy 98.3783%.

From this example we know, even we could only utilize the difference equation to calculate the final estimation value , because the role of whitening differential equation exist and which could be help differential equation generate more accuracy result, so today the usage of the grey difference equation only for introducing grey differential equation concept and related models.

5.3 The role of whitening differential equation in GM (1, 1)

The role of whitening differential equation is quite extraordinary in GM (1, 1) modelling which utilizes a first-order differential equation form to represent all estimation values with specified equal gapped input value t or t index. The general form of the whitening differential equation is

$$y' + p(x)y = q(x) \tag{5.13}$$

The solution form of the whitening differential equation is

$$y = \exp\left(-\int p(x)dx\right)\left[\int q(x)\exp\left(\int p(x)dx\right)dx + c\right] \tag{5.14}$$

Then the solution will be

$$\begin{cases} y(t) = \left(y(0) - \dfrac{\alpha}{\beta}\exp(-\beta t) + \dfrac{\alpha}{\beta}\right) \\ y(0) = y_0 \end{cases} \tag{5.15}$$

The structure of the whitening differential equation solution in GM (1, 1) model $\begin{cases} y(t) = \left(y(0) - \dfrac{\alpha}{\beta}\exp(-\beta t) + \dfrac{\alpha}{\beta}\right) \\ y(0) = y_0 \end{cases}$ is a simple exponential equation, only can represent as an increase tendency curve (continue case) or a discrete increase data sequence with equal-gapped index t (discrete case). Even after the IAGO (IAGO is the abbreviation of Inverse accumulative generation operator) treatment, the estimation data sequence still can't showing any undulation at all. Actually, in the real world, the data sequence collected from widespread divisions of science research appeared various structures, even some of them showing anomaly form, so it might be difficult for using a simple formed exponential differential equation to represent.

Example 5.3.1: Let $X^{(0)} = (0.94764, 1.83403, 0.99801, 0.21138, 0.31537, 1.26251)$, from the figure 5.3.1 we could see, $X^{(0)}$ isn't a monotone data sequence. We utilize GM (1, 1) whitening differential equation to calculate the estimation value.

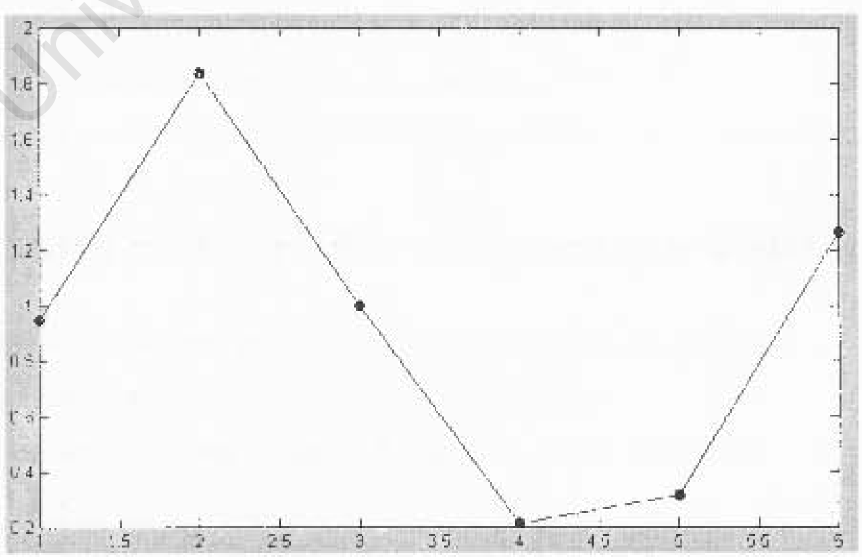

Figure 5.3.1 $X^{(0)}$ original data sequence

We utilize the software discussed in Chapter 2 to calculate estimation value. After 1-AGO treatment the data sequence $X^{(1)} = (0.94764, 2.78167, 3.77968, 3.99106, 4.30643, 5.56894)$.

Utilize the regression model to simulate the whitening differential equation is

$$\begin{cases} y(t) = \left(y(0) - \dfrac{\alpha}{\beta}\exp(-\beta t) + \dfrac{\alpha}{\beta} \right) \\ y(0) = y_0 \end{cases} - \begin{cases} y(t) = \left(y(0) - 0.153446\exp(-2.08168t) + 0.153446 \right) \\ y(0) = 0.94764 \end{cases} \tag{5.16}$$

(Whitening differential equation y (t)'s shape form see figure 5.3.2)

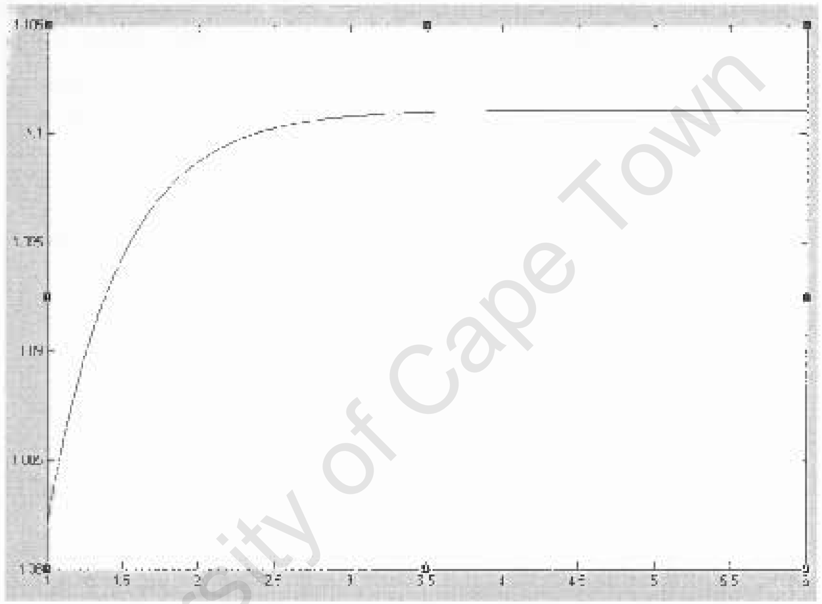

Figure 5.3.2 Whitening differential equation y (t)'s shape form

The estimation value of $X^{(1)}$ is $\hat{X}^{(1)} = (0.94764, 2.47047, 3.57691, 4.38081, 4.9649, 5.38928)$.

The estimation value of $X^{(0)}$ is $\hat{X}^{(0)} = (0.94764, 1.52283, 1.10644, 0.8039, 0.584087, 0.424377)$.

The average relative error is 91.9472%

From the above example we realized, because of the exponential differential equation form is quite simple, lot of no regular data sequence can't be estimated using GM (1, 1) model. Even some of them using transformation idea to regular data sequence to an increase monotone data structure, but after they transfer the data back, most of them still can't have an efficient result.

The inefficient estimation results phenomena causing by whitening differential equation's singularity structure are still reflect a grey-statistical inconsistency. These phenomena remind us, using various different of whitening differential equation form to represent the estimation value could change the grey-statistical inconsistency from essential part. In next section we will discuss this question, and created some new whitening differential equation form to calculate the estimation value.

5.4 The role of regression model in GM (1, 1)

Even we can easily derived the solution form from the difference equation or whitening differential equation in GM (1, 1) modelling , we still need to seek appropriate parameter pair (α, β) from a simply regression model.

Before discussion with regression model in GM (1, 1), we first need to review some definitions about linear regression.

Definition 5.4.1: linear models represent the relationship between the conditional expected value of one variable y given the values of some other variables X. (either continuous or categorical)

The simple linear regression model represent as a typical state form:

$$y = \beta + \alpha x - \varepsilon \tag{5.17}$$

Y is an n-by-1 vector of observations of the response variable.

X is the n-by-p design matrix determined by the predictors.

α, ß is a p-by-1 vector of unknown parameters to be estimated.

ε is an n-by-1 vector of independent. identically distributed random disturbances.

For linear models method, it contains various forms of models to represent different data structure of diverse purpose of regression, such as: Multiple Linear Regression; Quadratic Response Surface Models; Stepwise Regression; Generalized Linear Models; Robust and Nonparametric Methods, etc.

The most important part for a general linear regression model is using least-square or some other methods to estimate the parameters. By recognized the $y = \beta - \alpha x + \varepsilon$ regression model , we could express the regression model as a system of linear equations by using data matrix X, aiming vector Y and parameter vector ε. The linear model can be written as:

$$\begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_3 & 1 \end{bmatrix} * \begin{bmatrix} \alpha \\ \beta \end{bmatrix} + \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix} \tag{5.18}$$

When we using the pure matrix notation it may becomes:

$$Y = X\delta + \varepsilon \tag{5.19}$$

The least-square estimator for δ is

$$\hat{\delta} = (X'X)^{-1} X'Y \tag{5.20}$$

We could easily put parameter δ back into the model formula to get the predicted y values at the data points.

$$\hat{Y} = X\delta = HY \tag{5.21}$$

$$H = X(X'X)^{-1} X' \tag{5.22}$$

Then the residuals or errors are denoted as the difference between the observed Y and predicted $\hat{Y}$ values.

$$\varepsilon = Y - \hat{Y} = (I - H)Y \tag{5.23}$$

The residuals are useful for detecting failures in the model assumptions, since they correspond to the errors, ε, in the model equation. By assumption, these errors each have independent distributions with mean equal to zero and a constant variance."

The residuals, however, are correlated and have variances that depend on the locations of the data points. It is a common practice to scale the residuals so they all have the same variance.

After the review of linear models, we are able to follow the linear regression definition to re-write the regression model to a strictly regression model form.

In GM (1, 1) modelling, the regression model is always coupling with a whitening differential equation interactively. We could simply derive the solution form from the whitening differential equation. And then utilize the parameter pair (a,b) we estimated before to complete the GM (1, 1) modelling.

Before we emphasize the coupling idea of GM (1, 1) modelling, we simply consider the

GM (1, 1) model as a regression model with a differential equation constrained. Now we realized that GM (1, 1) modelling essential is a differential equation coupled with a regression model, it is no longer a constrained regression question anymore. Differential equation gives a solution form and regression model estimate the parameter of the general solution. Then we found the quality of GM (1, 1) model (measured by model efficiency) should depend on the quality of coupling. Furthermore, the significance of the coupling is depending upon the dynamics underlying the data and the form of whitening differential equation. A quality coupling model should have high statistical-grey consistency, only a high statistical-grey consistent GM (1, 1) model possesses intrinsic coupling of regression model and differential equation model.

5.5 The role of AGO and IAGO in GM (1, 1)

Because of the data sequence we collected from different widespread divisions of actual world is quite various and complicated, so we need utilize AGO (abbreviation of accumulative generation operator) and IAGO (abbreviation of Inverse accumulative generation operator) to manipulate the complicated data structure. The solution forms of whitening differential equation $\hat{x}^{(1)}(k+1)=\left[x^{(0)}(1)-\frac{a}{b}\right]e^{-ak}+\frac{a}{b}$ where $\hat{x}^{(0)}(k)=\hat{x}^{(1)}(k)-\hat{x}^{(1)}(k-1)$ is a kind of an exponential equation. Then most of raw data sequence after accumulating generation once may transform the pattern from no regularity condition $X^{(0)}$ level to a tendency of growth $X^{(1)}$ level, and after that, we still can using inverse accumulating generation operator transform the estimates from $X^{(1)}$ level back to $X^{(0)}$ level. In fact, the AGO treatment could be understand as an approximation of integral and IAGO could be understand as an approximation of derivative. Therefore, the interactive coupling in GDE modelling is facilitated by AGO and IAGO. For more clarity, let us restate the coupling of the regression model and the (whitening) differential equation model in the standard GM (1, 1) model in matrix language (i.e., linear model language). Let

$$Y_{n\times 1}=\begin{bmatrix} x^{(0)}(1) \\ x^{(0)}(2) \\ \vdots \\ x^{(0)}(n) \end{bmatrix} A_{n\times n}=\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix} A^{(1/2)}_{(n-1)\times n}=\begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & \cdots & 0 \\ 0 & 0 & \frac{1}{2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{2} \end{bmatrix} \tag{5.24}$$

The exploratory vector $X^{(1)}=\left[x^{(1)}(1),\quad x^{(1)}(2),\quad \cdots,\quad x^{(1)}(n)\right]^{T}_{n\times 1}$ obtained the value by linear transformation

$$X^{(1)}=A_{n\times n}\times Y_{n\times 1}=\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 1 & 1 & 0 & \cdots & 0 \\ 1 & 1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & 1 & \cdots & 1 \end{bmatrix}_{n\times n} \cdot \begin{bmatrix} x^{(0)}(1) \\ x^{(0)}(2) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}_{n\times 1} \tag{5.25}$$

$$=\left[x^{(0)}(1), x^{(0)}(1)+x^{(0)}(2), x^{(0)}(1)+x^{(0)}(2)+x^{(0)}(3),\cdots, x^{(0)}(1)+x^{(0)}(2)+\cdots+x^{(0)}(n)\right]^{T}_{n\times 1}$$
$$=\left[x^{(1)}(1), x^{(1)}(2),\cdots, x^{(1)}(n)\right]^{T}_{n\times 1}$$

Then we can get the "mean" vector $Z^{(1)} = \left[z^{(1)}(1), z^{(1)}(2), \cdots, z^{(1)}(n-1)\right]_{n-1}^{T}$, the so called background value vector.

$$Z^{(1)} = -A_{1/2}AY = \begin{bmatrix} \frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & \cdots & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & \frac{1}{2} & \frac{1}{2} \end{bmatrix}_{(n-1)\times n} \begin{bmatrix} x^{(1)}(1) \\ x^{(1)}(2) \\ \vdots \\ x^{(1)}(n) \end{bmatrix}_{n\times 1} \tag{5.26}$$

$$= \begin{bmatrix} \frac{1}{2}\left(x^{(1)}(1) + x^{(1)}(2)\right) \\ \frac{1}{2}\left(x^{(1)}(2) + x^{(1)}(3)\right) \\ \vdots \\ \frac{1}{2}\left(x^{(1)}(n-1) + x^{(1)}(n)\right) \end{bmatrix}_{(n-1)\times 1} = \left[z^{(1)}(1), z^{(1)}(2), \cdots, z^{(1)}(n-1)\right]_{n-1}^{T} \tag{5.27}$$

Then the 1-AGO on observation vector is to create the linear model:

$$\hat{Y} = I_{(n-1)\times 1}\alpha + \left(-A^{1/2}_{(n-1)\times n}A_{n\times n}Y_{n\times 1}\right)\beta + \varepsilon \tag{5.28}$$

From the equation (3.37) we obtained, can cleanly figure out the regression model in equation (3.37) is actually a mixed linear model, with a pair of parameter (α, β) need to be estimated. The matrix form $A^{1/2}_{(n-1)\times n}$ and $A_{n\times n}$ to obtained the liner mixed model in $\hat{Y} = I_{(n-1)\times 1}\alpha + \left(-A^{1/2}_{(n-1)\times n}A_{n\times n}Y_{n\times 1}\right)\beta + \varepsilon$ is hinted by form of the whitening differential equation $dx^{(1)}/dt = \alpha - \beta x^{(1)}$. $z = \left(-A^{1/2}_{(n-1)\times n}A_{n\times n}Y_{n\times 1}\right)$ is a "design" matrix of random variables which associated with the random effect parameter β. The 1-AGO treatment we use $X^{(1)} = A_{n\times n} \times Y_{n\times 1}$ a matrix form to calculate, after multiply $-A^{1/2}_{(n-1)\times n}$ the 1-AGO data sequence $X^{(1)}$ become a Z data sequence, where $Z^{(1)} = \left(\frac{1}{2}\left(x^{(1)} + x^{(2)}\right), \frac{1}{2}\left(x^{(2)} + x^{(3)}\right), \cdots, \frac{1}{2}\left(x^{(n-1)} + x^{(n)}\right)\right)$. The design matrix Z is a linear transformation of the observation vector $Y_{n\times 1}$. The term coupling of regression model and whitening differential equation in GM (1, 1) model building is just in this sense.

The Inverse accumulating generation operator (abbreviated as IAGO) treatment also needs to translate into a linear transformation using matrix form.

$$A_{(n-1)\times n}^{(-1)} = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ 0 & -1 & 1 & \cdots & 0 \\ 0 & 0 & -1 & \ddots & 0 \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} \tag{5.29}$$

The 1-IAGO treatment would be

$$x^{(-1)}(k+1) = x^{(0)}(k+1) - x^{(0)}(k) \tag{5.30}$$

We translate the equation (5.30) into a matrix language, the second order grey derivate vector will be

$$X_{(n-1)\times 1}^{(-1)} = A_{(n-1)\times n}^{(-1)} Y_{n\times 1} \tag{5.31}$$

After we utilize mathematical language to translate the AGO and IAGO treatment into a matrix operation, we essentially know the AGO and IAGO operation is just a data transformation. The transformation result making the transformed data more fit for the whitening differential equation to estimate a better prediction value and make the regression model easier to generate an accuracy pair of parameter. After that we need describe a new idea for using integral and derivative instead of AGO and IAGO operation to modify a new GM (1, 1) model.

5.6 A new modified model of GM (1, 1) makes use of integral and derivative method.

From lot of modelling practice experience, the AGO and IAGO treatment is only some kind of mathematical idea using summation for data transformation which quite useful for many mathematics problem solving. But for $z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right]$

where $x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4,\cdots,n$ equation, which use for generate parameter of whitening differential equation (exponential equation) seemed like inapprehension. So in here, we think about to via integral idea to instead of AGO and $z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right]$ equation to generate an appropriate parameter pair for exponential differential equation.

Example 5.6.1: Let $X^{(0)}=(2.874,3.278,3.337,3.390,3.679)=(X_1^{(0)},X_2^{(0)},X_3^{(0)},X_4^{(0)},X_5^{(0)})$ (the same data sequence we used before), via integral and derivative method to instead of AGO and IAGO treatment to generate parameter for whitening differential equation of GM (1, 1) model.

Step 1

Because of the operation of integral only can deal with continuous equation, so at the beginning, we need to utilize spline function to joint all the discrete points and make the function curve very smooth.

The spline function we manipulate to fit the data sequence is polynomial spline function (fitting result illustrate in Figure 5.6.1), the equation and parameters of equation list below:

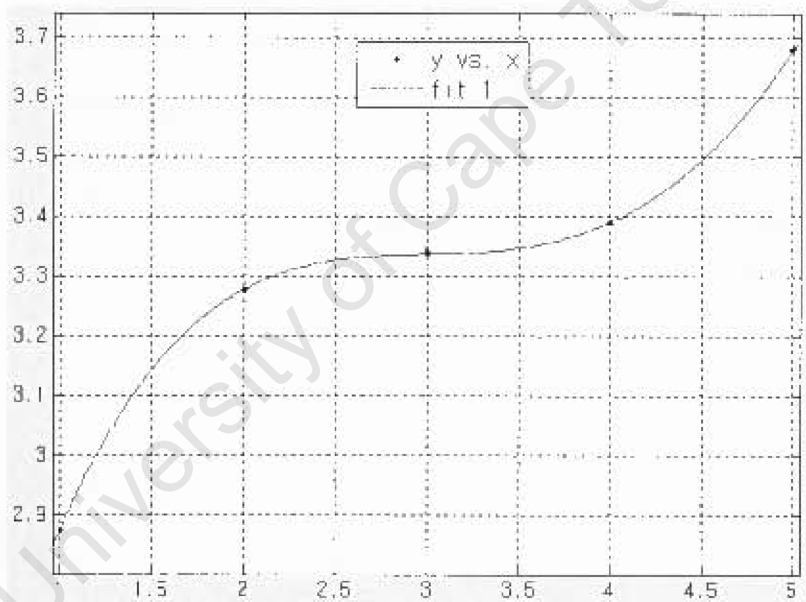

Figure 5.6.1 polynomial curve fitting

Linear model Poly4:

$$F(x) = p1*x^4 + p2*x^3 + p3*x^2 + p4*x + p5 \tag{5.32}$$

Coefficients:

p1 =-0.004042

p2 =0.09692

p3 =-0.653

p4 =1.745

p5 =1.689

Goodness of fit:

SSE: 2.524e-029

R-square: 1

*Step 2

We denote the polynomial spline function as f(x), we need to calculate the integral value

$$F_k = \int_{k}^{k+1} f(x)dt$$

$$\begin{aligned} F_1 &= \int_1^2 f(x)dt = 3.12143 \quad F_2 = \int_2^3 f(x)dt = 6.44208 \\ F_3 &= \int_3^4 f(x)dt = 9.79451 \quad F_4 = \int_4^5 f(x)dt = 13.3018 \end{aligned} \tag{5.33}$$

Step3 Utilizing least squares estimate the parameter sequence $\hat{a} = [a,b]^T$

$$\hat{a} = \left[B^T B\right]^{-1} B^T Y \tag{5.34}$$

$$B = \begin{bmatrix} -F_1 & 1 \\ -F_2 & 1 \\ -F_3 & 1 \\ -F_4 & 1 \end{bmatrix} = \begin{bmatrix} -3.12143 & 1 \\ -6.44208 & 1 \\ -9.79451 & 1 \\ -13.3018 & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ x^{(0)}(4) \\ x^{(0)}(5) \end{bmatrix} = \begin{bmatrix} 3.278 \\ 3.337 \\ 3.390 \\ 3.679 \end{bmatrix}; \tag{5.35}$$

$$\hat{a} = \left[B^T B\right]^{-1} B^T Y = [a,b]^T = \begin{pmatrix} -0.0373 \\ 3.1167 \end{pmatrix} \tag{5.36}$$

*Step4: Generate the grey model and the estimation value using derivate method instead of IAGO

$$\frac{d\hat{x}^{(1)}}{dt} - 0.0373\hat{x}^{(1)} = 3.1167 \tag{5.37}$$

Where

$$x^{(0)}(t) = \frac{d\hat{x}^{(1)}}{dt} \tag{5.38}$$

$$\hat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a}\right]e^{-ak} + \frac{b}{a} = 86.5139 \bullet e^{-ak} + \frac{b}{a} = 86.5139e^{0.0373k} - 83.6399 \tag{5.39}$$

$$\hat{x}^{(0)}(k+1) = \frac{\hat{x}^{(1)}(k+1)}{dk} = 3.2238 * \exp(0.0373 * k) \tag{5.40}$$

The estimation value is $\hat{X}^{(0)} = [2.874, 3.2238, 3.3462, 3.4733, 3.6052]$

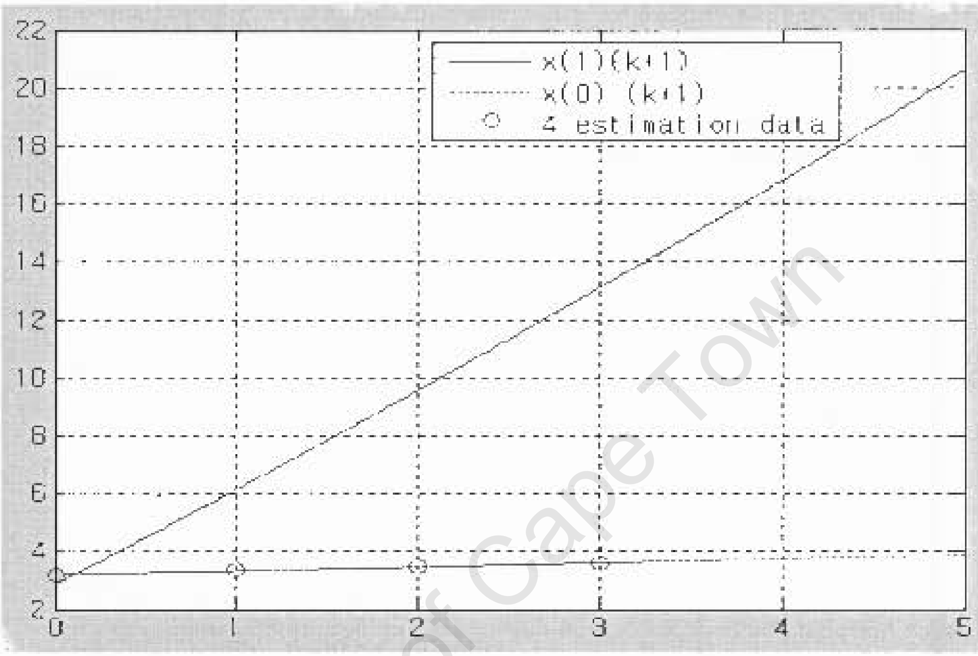

Figure 5.6.2 Whitening equation, derivative equation and estimation value image

Step5:

Relative Errors (%):

$$\Delta k = \frac{|\delta(k)|}{x^{(0)}(k)} = (1.65345\%, 0.27570\%, 2.45722\%, 2.00600\%) \tag{5.41}$$

Average relative error:

$$\Lambda = \frac{1}{4}\sum_{k=2}^{5} \Lambda_k = 0.015980877454979 = 1.59809\% \tag{5.42}$$

From the estimation result and Average relative error we can see, the method of utilize integral, derivative operation instead of AGO, IAGO treatment is quite successful from the result. The accuracy looks even little better than using the traditional summation way. $(1.59809\% < 1.60217\%)$.

The estimation result please see figure 5.6.3

The details of spline functions using for GM (1, 1) model modified ideas we already discussed in the chapter 4.

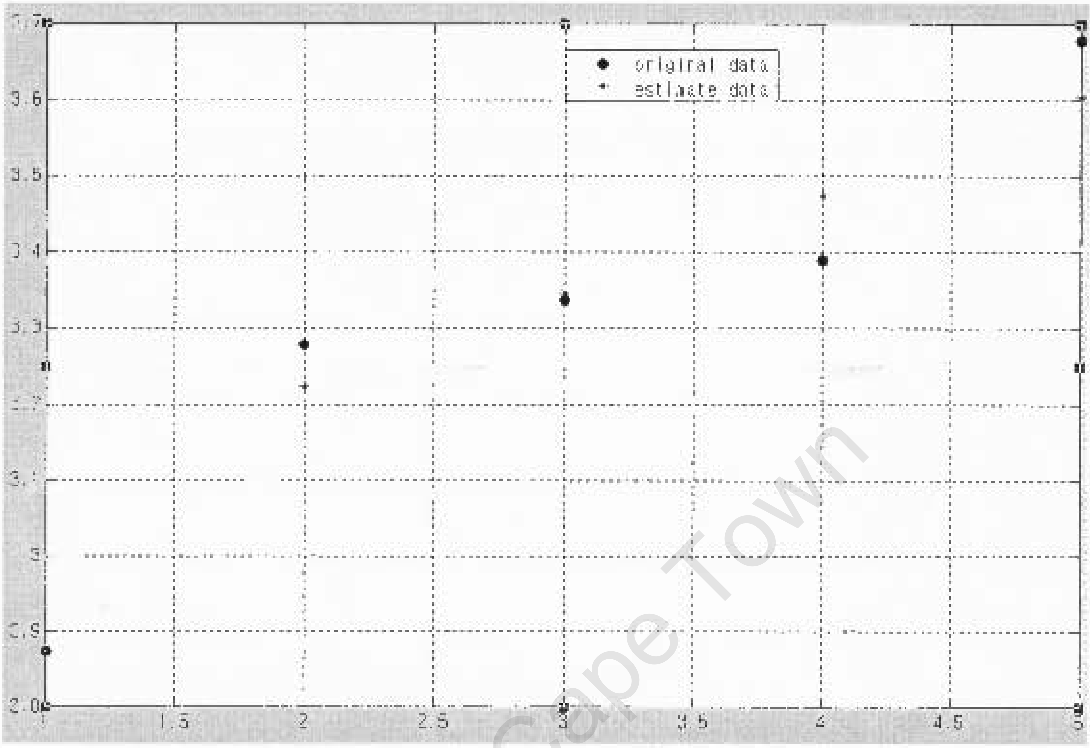

Figure 5.6.3 compares with original data and estimate data

5.7 Coupling Principle in Grey Differential Equation Modelling

After we systematically explore the nature of the GM (1, 1) model and finally identified that the definition of GM (1, 1) model couples a simple regression model and a whitening differential equation model together organically. The form of the whitening differential equation (abbreviated as WDE) (ex. $\frac{dx^{(1)}(t)}{dt}+\alpha x^{(1)}(t)=\beta$) determines the form of the coupling regression (abbreviated as CREG) (ex. $y_K=\beta+\alpha x_K, K=2,3,4,\cdots,n$ regression parameter pair (α,β), denoted as (a, b) .The data assimilated parameter pair (a, b) in CREG determines the system parameter pair (α.β).Because the standard GM (1, 1) model is equal-gapped strictly positive data sequence, according to the law of grey mathematics we can easily assume the discrete attribute and utilize WDE (a function form) to calculate the estimation data sequence value. The coupling translation rule is listed in previously (see Table 5.1.1)

Now, after we examined the three basic component models: simple regression model, differential equation model and difference model in GM (1, 1), and list the coupling translation rule in table 5.1.1, we try to conclude all the analysis we made previously and give a definition about Coupling Principle at the end.

Definition 5.7.1: nature of the GM (1, 1) model dynamics is a coupled differential equation model and regression model. The differential equation plays the role of defining the basic formation of regression model and the response function, while the appropriately formed regression model in terms of AGO or IAGO operations plays the role of estimating the two parameters assimilated by data. We state such coupling nature in GM (1, 1) model building as the Coupling Principle.

5.8 Summary

For answering the question of the grey modelling waste the statistical information during the coefficient estimates procedure, we give a new coupling idea of grey differential equation modelling. After analysis each role of the 3-steps modelling procedure, we point out the essential of grey differential equation modelling is just a regression model coupled with a whitening differential equation. At the end of this chapter, we give everyone the definition of Coupling Principle.

Chapter 6. A Family of Extended GDE Models for Statistical-Grey Consistency Modelling

In chapter 5 we have shown that the GM (1, 1) model is just a simple regression model coupled with a whitening differential equation. The regression model is only using for estimate appropriate parameter pair (α, β), and parameter pair (α, β) need to transferred to whitening differential equation to simulate the final solution. So, the form of regression model and whitening differential equation both depend on the form of whitening differential equation. Because GM (1, 1) usually deals with small sample data sequence, normally larger than 4 and less than 20, but GM (1, 1) only can solve certain formed data sequence which obey the grey index and statistical index law to keep the statistical-grey consistency. When we try to modelling other formed data sequence, normally we should do some transformation. In this chapter we first list some unformed data pattern which not following the grey index and statistical index law and could not solved just without any data transformation by GM (1, 1) model in previously research, after that we will represent our recently research outcome to build different formed grey differential equation models to deal with different kind of formed data sequence.

Up to our research now, we totally built seven elementary types of models in GDE family, but in this chapter we discuss list 5 types of them, and each type of them we give an application for testing model's efficiency. The Table 6.1 will illustrate the details about the seven elementary types of models in GDE family.

Table 6.1: Seven Elementary Types of Models in GDE Family

New model	*Order* $p=1$	*Order* $p=2$
Type I	$\begin{cases} \frac{dx}{dt}=\alpha_0+\alpha_1 x \\ \Delta x(k)=\alpha_0+\alpha_1\hat{x}(k)+\varepsilon_k \end{cases}$	$\begin{cases} \frac{d^2x}{dt^2}=\alpha_0+\alpha_1 x+\alpha_2\frac{dx}{dt} \\ \Delta^2 x(k)=\alpha_0+\alpha_1\hat{x}(k)+\alpha_2\Delta x(k)+\varepsilon_k \end{cases}$
Type II	$\begin{cases} \frac{dx}{dt}=\alpha_0 e^{\delta t}+\alpha_1 x \\ \Delta x(k)=\alpha_0 e^{\delta k}+\alpha_1\hat{x}(k)+\varepsilon_k \end{cases}$	$\begin{cases} \frac{d^2x}{dt^2}=\alpha_0 e^{\delta t}+\alpha_1 x+\alpha_2\frac{dx}{dt} \\ \Delta^2 x(k)=\alpha_0 e^{\delta k}+\alpha_1\hat{x}(k)+\alpha_2\Delta x(k)+\varepsilon_k \end{cases}$
Type III	$\begin{cases} \frac{dx}{dt}=\alpha_0\sin(\omega t+\varpi)+\alpha_1 x \\ \Delta x(k)=\alpha_0\sin(\omega k+\varpi)+\alpha_1\hat{x}(k)+\varepsilon_k \end{cases}$	$\begin{cases} \frac{d^2x}{dt^2}=\alpha_0\sin(\omega t+\varpi)+\alpha_1 x+\alpha_2\frac{dx}{dt} \\ \Delta^2 x(k)=\alpha_0\sin(\omega k+\varpi)+\alpha_1\hat{x}(k)+\alpha_2\Delta x(k)+\varepsilon_k \end{cases}$
Type IV	$\begin{cases} \frac{dx}{dt}=\alpha_0 e^{\delta t}\sin(\omega t+\varpi)+\alpha_1 x \\ \Delta x(k)=\alpha_0 e^{\delta k}\sin(\omega k+\varpi)+\alpha_1\hat{x}(k)+\varepsilon_k \end{cases}$	$\begin{cases} \frac{d^2x}{dt^2}=\alpha_0 e^{\delta t}\sin(\omega t+\varpi)+\alpha_1 x+\alpha_2\frac{dx}{dt} \\ \Delta^2 x(k)=\alpha_0 e^{\delta k}\sin(\omega k+\varpi)+\alpha_1\hat{x}(k)+\alpha_2\Delta x(k)+\varepsilon_k \end{cases}$
*Type V**	$\begin{cases} \frac{dx}{dt}=\alpha_0 p_q(t)+\alpha_1 x \\ \Delta x(k)=\alpha_0 p_q(k)+\alpha_1\hat{x}(k)+\varepsilon_k \end{cases}$	$\begin{cases} \frac{d^2x}{dt^2}=\alpha_0 p_q(t)+\alpha_1 x+\alpha_2\frac{dx}{dt} \\ \Delta^2 x(k)=\alpha_0 p_q(k)+\alpha_1\hat{x}(k)+\alpha_2\Delta x(k)+\varepsilon_k \end{cases}$
*Type VI**	$\begin{cases} \frac{dx}{dt}=\alpha_0 e^{\delta t}p_q(t)+\alpha_1 x \\ \Delta x(k)=\alpha_0 e^{\delta k}p_q(k)+\alpha_1\hat{x}(k)+\varepsilon_k \end{cases}$	$\begin{cases} \frac{d^2x}{dt^2}=\alpha_0 e^{\delta t}p_q(t)+\alpha_1 x+\alpha_2\frac{dx}{dt} \\ \Delta^2 x(k)=\alpha_0 e^{\delta k}p_q(k)+\alpha_1\hat{x}(k)+\alpha_2\Delta x(k)+\varepsilon_k \end{cases}$
*Type VII**	$\begin{cases} \frac{dx}{dt}=\alpha_0 p_q(t)\sin(\omega t+\varpi)+\alpha_1 x \\ \Delta x(k)=\alpha_0 p_q(k)\sin(\omega k+\varpi)+\alpha_1\hat{x}(k)+\varepsilon_k \end{cases}$	$\begin{cases} \frac{d^2x}{dt^2}=\alpha_0 p_q(t)\sin(\omega t+\varpi)+\alpha_1 x+\alpha_2\frac{dx}{dt} \\ \Delta^2 x(k)=\alpha_0 p_q(k)\sin(\omega k+\varpi)+\alpha_1\hat{x}(k)+\alpha_2\Delta x(k)+\varepsilon_k \end{cases}$

(From R. Guo, D. Guo, T. Dunne and C. Thiart (2007))

6.1 The Weakness of GM (1, 1) in Data Pattern Catching

Now we utilize an example to show the weakness of GM (1, 1) dealing with unformed data pattern and also to illustrate how scholars exploit data transformation idea to solve that problem in previously. Besides, even data transformation idea could solve some trouble of unformed data estimation, but the procedure of the data transforming may still cause of some new troubles, like missing data information problem, like the procedure of data transforming causing new estimate errors problem, etc. Summarized overall discussions, we need to extend the families of Grey differential equation models (abbreviate as GDE model), from changing the form of GDE model, to adapted itself to various formed data patterns.

Example 6.1.1: let the data sequence, $X^{(0)} = \{4.250, 1.287, 0.810, 1.524, 2.612, 3.679, 4.608, 5.425, 6.203, 7.012, 7.908\}$ and then we use remnant method, one data transformation idea, to calculate the estimate value of $X^{(0)}$ sequence, try to improve the inefficient accuracy of GM (1, 1) model with considered of the error part sequence.

We give the discrete data sequence

$$X^{(0)} = \left(x^{(0)}(i)\right)_{i=1}^{11} = \{4.250, 1.287, 0.810, 1.524, 2.612, 3.679, 4.608, 5.425, 6.203, 7.012, 7.908\} \tag{6.1}$$

After 1-AGO treatment we obtain

$$X^{(1)} = \left(x^{(1)}(i)\right)_{i=1}^{11} = \{4.250, 5.537, 6.347, 7.871, 10.483, 14.162, 18.770, 24.195, 30.398, 37.410, 45.318\} \tag{6.2}$$

Apply a consecutive neighbour generation to $X^{(1)}$, let

$$Z = \frac{1}{2}\left(x^{(1)}(i) + x^{(1)}(i-1)\right) = \left(z(i)\right)_{i=2}^{11} = \{4.2500, 4.8935, 5.9420, 7.1090, 9.1770, 12.3225, 16.4660, 21.4825, 27.2965, 33.9040, 41.3640\} \tag{6.3}$$

Perform a least square estimate for the parametric sequence $\hat{a} = [a, b]^T$ .we can obtain that

$$\hat{a} = [a, b]^T = [B^T B]^{-1} B^T Y \tag{6.4}$$

Where

$$B = \begin{bmatrix} -z(2) & 1 \\ -z(3) & 1 \\ \vdots & \vdots \\ -z(11) & 1 \end{bmatrix}, \quad Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(11) \end{bmatrix} \tag{6.5}$$

Then

$$\hat{a} = [a, b]^T = [B^T B]^{-1} B^T Y = [-0.1929, 0.6363] \tag{6.6}$$

The whitening differential equation will be

$$\frac{d\hat{x}^{(1)}}{dt} - 0.1929\hat{x}^{(1)} = 0.6363 \tag{6.7}$$

The time response sequence

$$\hat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a}\right]e^{-ak} + \frac{b}{a} \tag{6.8}$$
$$= 7.5496 \cdot e^{0.1929k} - 3.2996$$

The simulate value of $X^{(1)}$ sequence

$$\hat{X}^{(1)} = \left(\hat{x}^{(1)}(i)\right)_{i=1}^{11} = \{4.2500, 5.8558, 7.8032, 10.1647, 13.0286, 16.5016, 20.7134, 25.8210, 32.0149, 39.5264, 48.6355\} \tag{6.9}$$

The simulate value of $X^{(0)}$ sequence

$$\hat{X}^{(0)} = \left(\hat{x}^{(0)}(i)\right)_{i=1}^{11} = \{4.25, 1.6058, 1.9474, 2.3615, 2.8639, 3.473, 4.2118, 5.1076, 6.1939, 7.5115, 9.1091\} \tag{6.10}$$

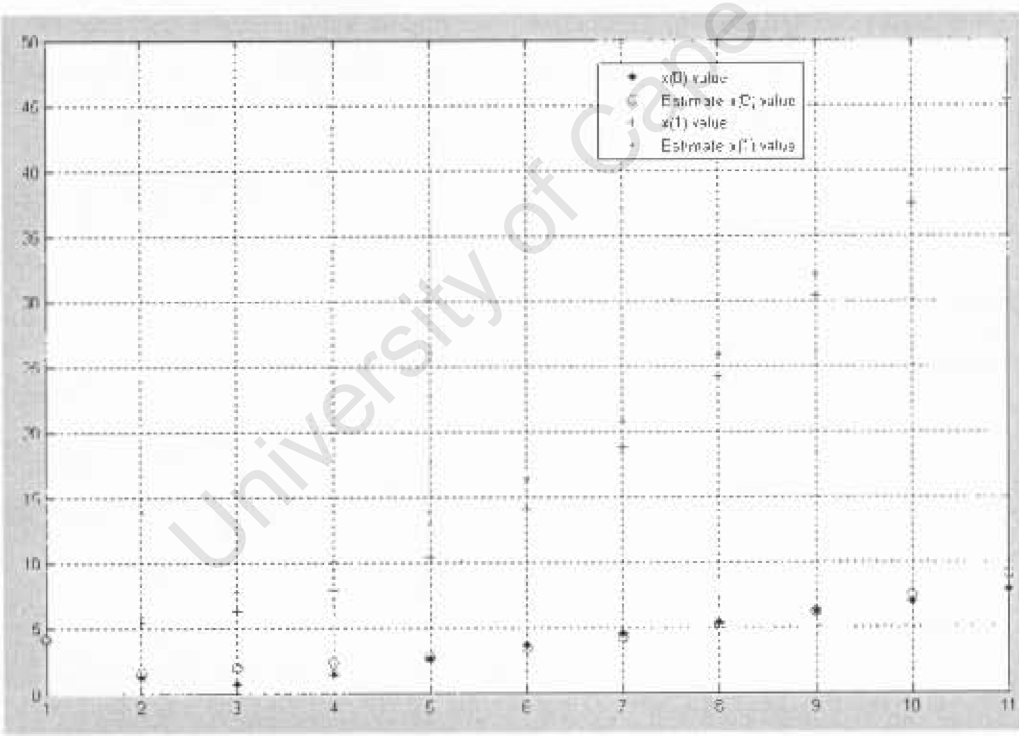

Figure 6.1.1 Compare $X^{(0)}$ value vs. $\hat{X}^{(0)}$ value. $X^{(1)}$ value vs. $\hat{X}^{(1)}$ value

Evaluate the relevant error values: (See Table 6.1.1)

Table 6.1.1 relevant error values

	Real Data	**Simulated Data**	**Errors**	**Relative Errors (%)N**
No.	$x^{(0)}(i)$	$\hat{x}^{(0)}(i)$	$\varepsilon(k) = x^{(0)}(i) - \hat{x}^{(0)}(i)$	$\Delta_K = \frac{\lvert\varepsilon(k)\rvert}{x^{(0)}(k)}$
2	1.287	1.6058	-0.3188	0.2477
3	0.81	1.9474	-1.1374	1.4042
4	1.524	2.3615	-0.8375	0.5495
5	2.612	2.8639	-0.2519	0.0964
6	3.679	3.4730	0.2060	0.0560
7	4.608	4.2118	0.3962	0.0860
8	5.425	5.1076	0.3174	0.0585
9	6.203	6.1939	0.0091	0.0015
10	7.012	7.5115	-0.4995	0.0712
11	7.908	9.1091	-1.2011	0.1519

The average relative error:

$$\Delta = \frac{1}{10}\sum_{k=2}^{11}\Delta_k = 27.23\% \tag{6.11}$$

From the average relative error of example 6.1.1 we could understand when GM (1, 1) model facing some unformed data pattern, the efficient of the model will be very poor. And now we try to operate some data transformation idea to improve the model efficiency.

Definition 6.1.1: When the accuracy of a GM (1, 1) model is not meeting a predetermined requirement, one can establish a GM (1, 1) model using the error sequence to remedy the original model in order to improve the accuracy, we call it as remnant GM (1, 1) model. (liu and lin, 2004)

We use remnant methods to calculate the estimated value of $X^{(0)}$ sequence; try to improve the accuracy of GM (1, 1) model with consideration of the error part sequence. But first we need to make use of ratio idea which we already discuss in section 3.2 to decline the estimate errors during the AGO and IAGO operation taken.

First we need to have:

$$R = \frac{X^{(1)}}{X^{(0)}} = (r(i))_{i=1}^{11} = \{1.0000,4.3023,7.8358,5.1647,4.0134,3.8494,4.0734,4.4599,4.9005,5.3351,5.7307\} \tag{6.12}$$

Then the simulation value of $X^{(0)}$ sequence:

$$\hat{X}^{(0)} = \frac{\hat{X}^{(1)}}{R} = \left(\hat{x}^{(0)}(i)\right)_{i=1}^{11} = \{4.2500,1.3611,0.9958,1.9681,3.2463,4.2868,5.0851,5.7896,6.5330,7.4087,8.4869\} \tag{6.13}$$

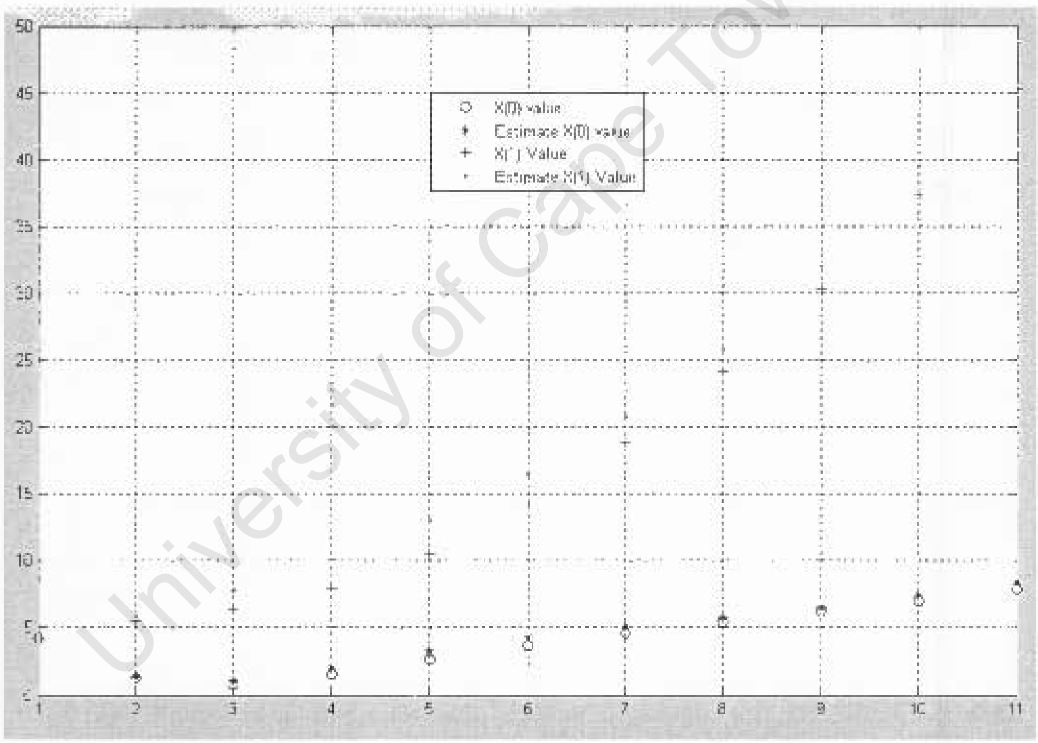

Figure 6.1.2 Compare $X^{(0)}$ value vs. $\hat{X}^{(0)}$ value, $X^{(1)}$ value vs. $\hat{X}^{(1)}$ value

Evaluate the relevant error values:

Table 6.1.2 relevant error values

	Real Data	Simulated Data	Errors	Relative Errors (%)N
No.	$x^{(0)}(i)$	$\hat{x}^{(0)}(i)$	$\varepsilon(k) = x^{(0)}(i) - \hat{x}^{(0)}(i)$	$\Delta_k = \frac{\lvert\varepsilon(k)\rvert}{x^{(0)}(k)}$
2	1.287	1.3611	-0.0741	0.0576
3	0.81	0.9958	-0.1858	0.2294
4	1.524	1.9681	-0.4441	0.2914
5	2.612	3.2463	-0.6343	0.2428
6	3.679	4.2868	-0.6078	0.1652
7	4.608	5.0851	0.4771	0.1035
8	5.425	5.7896	-0.3646	0.0672
9	6.203	6.5330	-0.3300	0.0532
10	7.012	7.4087	-0.3967	0.0566
11	7.908	8.4869	-0.5789	0.0732

The average relative error:

$$\Delta = \frac{1}{10}\sum_{i=2}^{11}\Delta_k = 13.4\% \tag{6.14}$$

The relative error is large than 5%, so, it is necessary to apply a remnant model to do some remedies.

$$\begin{aligned}\varepsilon^{(0)} = \left(\varepsilon^{(0)}(i)\right)_{i=2}^{11} = (&-0.0741,-0.1858,-0.4441,-0.6343,-0.6078,\\ &-0.4771,-0.3646,-0.3300,-0.3967,-0.5789)\end{aligned} \tag{6.15}$$

Taking absolute value

$$\begin{aligned}\left|\varepsilon^{(0)}\right| = \left|\left(\varepsilon^{(0)}(i)\right)_{i=2}^{11}\right| = (&0.0741,0.1858,0.4441,0.6343,0.6078,\\ &0.4771,0.3646,0.3300,0.3967,0.5789)\end{aligned} \tag{6.16}$$

Follow the GM (1, 1) modelling procedure to estimate $\left|\varepsilon^{(0)}\right|$ sequence value.

$$\begin{aligned}\left|\hat{\varepsilon}^{(0)}\right| = \left|\left(\hat{\varepsilon}^{(0)}(i)\right)_{i=2}^{11}\right| = (&0.0741,0.3466,0.5706,0.6318,05531,0.4367,\\ &0.3498,0.3325,0.4083,0.589)\end{aligned} \tag{6.17}$$

Then the estimate value of $\varepsilon^{(0)}$

$$\begin{aligned}\hat{\varepsilon}^{(0)} = \left(\hat{\varepsilon}^{(0)}(i)\right)_{i=2}^{11} = (&-0.0741,-0.3466,-0.5706,-0.6318,-05531,-0.4367,\\ &-0.3498,-0.3325,-0.4083,-0.589)\end{aligned} \tag{6.18}$$

The remnant value of $X^{(0)}$ data sequence only could from $i = 3$, so

$$Remnant(X^{(0)}) = (rx^{(0)}(i))_{i=3}^{11} = (x^{(0)}(i))_{i=3}^{11} + (\hat{\varepsilon}^{(0)}(i))_{i=3}^{11} = (0.6492, 1.3975, 2.6145, 3.7337, 4.6484, 5.4398, 6.2005, 7.0004, 7.8979) \tag{6.19}$$

Table 6.1.3 Remnant model relevant error values

	Real Data	Simulated Data	Errors	Relative Errors (%)N
No.	$x^{(0)}(i)$	$rx^{(0)}(i)$	$r\varepsilon(k) = x^{(0)}(i) - rx^{(0)}(i)$	$r\Delta_K = \frac{\lvert r\varepsilon(k)\rvert}{x^{(0)}(k)}$
3	0.81	0.6492	0.1608	0.2478
4	1.524	1.3975	0.1265	0.0905
5	2.612	2.6145	-0.0025	0.0010
6	3.679	3.7337	-0.0547	0.0147
7	4.608	4.6484	-0.0404	0.0087
8	5.425	5.4398	-0.0148	0.0027
9	6.203	6.2005	0.0025	0.0004
10	7.012	7.0004	0.0116	0.0017
11	7.908	7.8979	0.0101	0.0013

Then the average relative error

$$r\Delta = \frac{1}{9}\sum_{i=3}^{11} r\Delta_k = 4.10\% \tag{6.20}$$

Much less than 13.4% and 27.23% average relative error used estimate.

Then the accuracy of this model is

$1 - r\Delta = 100\% - 4.10\% = 95.90\%$

6.2 Families of Linear Differential Equation with Various Right-Side Functional Terms

Previously, from the discussion of the statistical-grey consistency problem, we known the nature of the grey differential equation modelling is just a regression model coupling with a differential equation model. But in actual world, we have various formed small sample data sequence need to investigate, which will be hard for us to test followed with any distribution, so we need to build different formed grey differential equation models to deal with diverse formed data sequences. Because the form of the data sequence is quite various, so the selection of a typical whitening differential equation adapt to a certain kind of formed data sequence is very difficult. Therefore, it is necessary for us to explore the forms of whitening differential equation, the corresponding regression, the close-form solution, the possible dynamic pattern and the potential suitable data pattern for modelling applications. Based on such considerations, we will discuss Linear whitening differential equation families with constant coefficients in left-side but functional term in right side. We are expecting the discussions will provide wider choices of grey differential models.

6.2.1 The Differential Equation with Constant Term on Right Side

We first list GM (1, 1) model and GM (2, 1) model as the form of grey differential equation model with constant term on right side. The modelling procedure and application examples we already performed in previously, so in here, we just list the concernment steps with no example and explanation gives.

GM (1, 1) model we could perform as a basic form of the first order linear differential equation with constant term on right side, α, β are parameters need for estimate in coupling regression model

$$\frac{d\hat{x}}{dt}+\beta\hat{x}=\alpha \tag{6.21}$$

To solve the differential equation in (6.21) we simply divide it as

$$\frac{d\hat{x}}{dt}+\beta\hat{x}=\alpha \Rightarrow \begin{cases} d\hat{x}_h / dt+\beta\hat{x}_h=0 \\ \\ 0+\beta\hat{x}_p=\alpha_p \end{cases} \tag{6.22}$$

Then the solution will be

$$\hat{x}=\hat{x}_h+\hat{x}_p=c_0e^{-\beta t}+\frac{\alpha}{\beta}$$

$$\text{Where } \hat{x}_h=c_0e^{-\beta t}, \hat{x}_p=\alpha/\beta \tag{6.23}$$

Where c_0 is a constant value

The coupling regression model represent as

$$x^{(0)}(k)=\alpha+\beta\left(-z^{(1)}(k)\right) \tag{6.24}$$

Where $x^{(0)}(k)=\left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$ is the initial discrete data sequence

$$x^{(1)}(k)=\sum_{i=1}^{k}x^{(0)}(i), k=2,3,4\cdots n \tag{6.25}$$

$$z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right] \tag{6.26}$$

GM (2, 1) model could represent as a second order linear differential equation with constant term in right side

$$\frac{d^2\hat{x}}{dt^2} + \beta_1 \frac{d\hat{x}}{dt} + \beta_2 \hat{x} = \alpha \tag{6.27}$$

To solve the differential equation in (6.27) we simply divide it as

$$\frac{d^2\hat{x}}{dt^2} + \beta_1 \frac{d\hat{x}}{dt} + \beta_2 x = \alpha \Rightarrow \begin{cases} \dfrac{d^2\hat{x}_h}{dt^2} + \beta_1 \dfrac{d\hat{x}_h}{dt} + \beta_2 \hat{x}_h = 0 \\ \beta_2 \hat{x}_p = \alpha \end{cases} \tag{6.28}$$

Then the solution will be

$$\hat{x} = \hat{x}_h + \hat{x}_p$$

$$\text{Where } \hat{x}_h = c_1 e^{\left(\left(-\beta_1 + \sqrt{\beta_1^2 - 4\beta_2}\right)/2\right)t} + c_2 e^{\left(\left(-\beta_1 - \sqrt{\beta_1^2 - 4\beta_2}\right)/2\right)t} \quad \hat{x}_p = \alpha / \beta_2 \tag{6.29}$$

Then

$$\hat{x} = c_1 e^{\left(\left(-\beta_1 + \sqrt{\beta_1^2 - 4\beta_2}\right)/2\right)t} + c_2 e^{\left(\left(-\beta_1 - \sqrt{\beta_1^2 - 4\beta_2}\right)/2\right)t} + \frac{\alpha}{\beta_2} \tag{6.30}$$

Where $\beta_1^2 \neq 4\beta_2$.

The coupling regression is

$$x^{(-1)}(k) = \alpha + \beta_1\left(-x^{(0)}(k)\right) + \beta_2\left(-z^{(1)}(k)\right), k = 2,3\cdots,n \tag{6.31}$$

Where

$$x^{(-1)}(k) = \left(x^{(0)}(2) - x^{(0)}(1), x^{(0)}(3) - x^{(0)}(2), \cdots, x^{(0)}(n) - x^{(0)}(n-1)\right) \tag{6.32}$$

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4\cdots n \tag{6.33}$$

$$z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right] \tag{6.34}$$

For the second order grey differential equation modelling, the minimum sample size n should be no less than 6.

6.2.2 The Differential Equation with Exponential Term on Right Side

Now we change the right side term with other kind of formed equations, first we alter it into a simple exponential form and with the example to show how it works.

The basic form of the first order linear differential equation with exponential term on right side

$$\frac{d\hat{x}}{dt} + \beta\hat{x} = \alpha e^{\delta t} \tag{6.35}$$

Where δ is a given constant

The homogeneous equation $d\hat{x}_h/dt + \alpha\hat{x}_h = 0$ with the solution $\hat{x}_h = c_0 e^{-\beta t}$, where c_0 is a constant

The particular solution $\hat{x}_p = A_0 e^{\delta t}$ satisfies the equation $A_0\delta e^{\delta t} + \beta A_0 e^{\delta t} = \alpha e^{\delta t}$, and then we have

$$A_0 = \alpha/(\delta + \beta) \tag{6.36}$$

$$\hat{x}_p = \alpha e^{\delta t}/(\delta + \beta) \tag{6.37}$$

Then the general solution is

$$\hat{x} = \hat{x}_h + \hat{x}_p = c_0 e^{-\beta t} + \frac{\alpha}{(\delta + \beta)} e^{\delta t} \tag{6.38}$$

The coupling regression model is

$$x^{(0)}(k) = \alpha e^{\delta k} + \beta\left(-z^{(1)}(k)\right) \tag{6.39}$$

Where $x^{(0)}(k) = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$ is the initial discrete data sequence

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4\cdots n \tag{6.40}$$

$$z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right] \tag{6.41}$$

The second order linear differential equation with exponential term on right side could represent as

$$\frac{d^2\hat{x}}{dt^2} + \beta_1\frac{d\hat{x}}{dt} + \beta_2\hat{x} = \alpha e^{\delta t} \tag{6.42}$$

Then the solution will be

$$\hat{x} = \hat{x}_h + \hat{x}_p$$

Where $\hat{x}_h = c_1 e^{\left(\left(-\beta_1 + \sqrt{\beta_1^2 - 4\beta_2}\right)/2\right)t} + c_2 e^{\left(\left(-\beta_1 - \sqrt{\beta_1^2 - 4\beta_2}\right)/2\right)t}$ $\hat{x}_p = \alpha e^{\delta t} / \left(\delta^2 + \beta_1 \delta + \beta_2\right)$ (6.43)

Then

$$\hat{x} = c_1 e^{\left(\left(-\beta_1 + \sqrt{\beta_1^2 - 4\beta_2}\right)/2\right)t} + c_2 e^{\left(\left(-\beta_1 - \sqrt{\beta_1^2 - 4\beta_2}\right)/2\right)t} + \frac{\alpha e^{\delta t}}{\left(\delta^2 + \beta_1 \delta + \beta_2\right)} \tag{6.44}$$

The coupling regression

$$x^{(-1)}(k) = \alpha e^{\delta t} + \beta_1\left(-x^{(0)}(k)\right) + \beta_2\left(-z^{(1)}(k)\right) \tag{6.45}$$

Where $x^{(0)}(k) = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$ is the initial discrete data sequence

$$x^{(-1)}(k-1) = x^{(0)}(k) - x^{(0)}(k-1), k = 2,3,\cdots,n \tag{6.46}$$

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4\cdots n \tag{6.47}$$

$$z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right] \tag{6.48}$$

For the second order grey differential equation modelling, the minimum sample size n should be no less than 6.

6.2.3 The Differential Equation with Sine and Cosine Term on Right Side

The basic form of the first order linear differential equation with sine term on right side is

$$\frac{d\hat{x}}{dt} + \beta \hat{x} = \alpha \sin \omega t \tag{6.49}$$

In order to satisfy the homogeneous equation

$$\frac{d\hat{x}_h}{dt} + \beta \hat{x}_h = 0 \tag{6.50}$$

We obtain the general solution:

$$\hat{x}_h = c_0 e^{-\beta t} \tag{6.51}$$

The particular solution having a form

$$\hat{x}_p = A_0 \sin \omega t + B_0 \cos \omega t \tag{6.52}$$

Need be satisfied the equation

$$\frac{d\hat{x}_p}{dt} + \beta\hat{x}_p = \mathrm{A}_0\omega\cos\omega t - B_0\omega\sin\omega t + A_0\beta\sin\omega t + B_0\beta\cos\omega t = \alpha\sin\omega t \tag{6.53}$$

Which leads to

$$\begin{cases} A_0\omega + B_0\beta = 0 \\ A_0\beta - B_0\omega = \alpha \end{cases} \tag{6.54}$$

Which will result

$$\begin{cases} A_0 = \dfrac{\alpha\beta}{\omega^2 + \beta^2} \\ B_0 = -\dfrac{\alpha\omega}{(\omega^2 + \beta^2)} \end{cases} \tag{6.55}$$

Then the particular solution will be

$$\hat{x}_p = \frac{\alpha\beta}{\omega^2 + \beta^2}\sin\omega t - \frac{\alpha\omega}{(\omega^2 + \beta^2)}\cos\omega t \tag{6.56}$$

The general solution is

$$\begin{gathered} \hat{x} = \hat{x}_h + \hat{x}_p == c_0 e^{-\beta t} + \frac{\alpha\beta}{\omega^2 + \beta^2}\sin\omega t - \frac{\alpha\omega}{\omega^2 + \beta^2}\cos\omega t \\ \text{where } \hat{x}_h = c_0 e^{-\beta t}\ \hat{x}_p = \frac{\alpha\beta}{\omega^2 + \beta^2}\sin\omega t - \frac{\alpha\omega}{\omega^2 + \beta^2}\cos\omega t \end{gathered} \tag{6.57}$$

The coupling regression is

$$x^{(0)}(k) = \alpha\sin(\omega k) + \beta(-z^{(1)}(k)), k = 2,3,\cdots,n \tag{6.58}$$

Where $x^{(0)}(k) = (x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n))$ is the initial discrete data sequence

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4\cdots n \tag{6.59}$$

$$z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right] \tag{6.60}$$

The second order linear differential equation with sine term on right side is

$$\frac{d^2\hat{x}}{dt} + \beta_1\frac{d\hat{x}}{dt} + \beta_2\hat{x} = \alpha\sin\omega t \tag{6.61}$$

In order to satisfy the homogeneous equation

$$\frac{d\hat{x}_h}{dt} + \beta \hat{x}_h = 0 \tag{6.62}$$

We obtain the general solution:

$$\hat{x}_h = c_0 e^{-\beta t} \tag{6.63}$$

The particular solution having a form

$$\hat{x}_p = A_0 \sin \omega t + B_0 \cos \omega t \tag{6.64}$$

Satisfied the equation

$$\begin{aligned}
&\frac{d^2\hat{x}_p}{dt^2} + \beta_1 \frac{d\hat{x}_p}{dt} + \beta_2 \hat{x}_p \\
&= -A_0\omega^2 \sin \omega t - B_0\omega^2 \cos \omega t + A_0\omega\beta_1 \cos \omega t \\
&\quad - B_0\omega\beta_1 \sin \omega t + A_0\beta_2 \sin \omega t + B_0\beta_2 \cos \omega t \\
&\quad - B_0\omega\beta_1 \sin \omega t + A_0\beta_2 \sin \omega t + B_0\beta_2 \cos \omega t \\
&= \alpha \sin \omega t
\end{aligned} \tag{6.65}$$

Which leads to

$$\begin{cases} \alpha = A_0\left(\beta_2 - \omega^2\right) - B_0\omega\beta_2 \\ 0 = A_0\omega\beta_1 + B_0\left(\beta_2 - \omega^2\right) \end{cases} \tag{6.66}$$

Then the coefficients of the particular solution x_p are

$$\begin{cases} A_0 = \dfrac{\alpha\left(\beta_2 - \omega^2\right)}{\left(\beta_2 - \omega^2\right)^2 + \beta_1\beta_2\omega^2} \\ B_0 = -\dfrac{\alpha\beta_1\omega}{\left(\beta_2 - \omega^2\right)^2 + \beta_1\beta_2\omega^2} \end{cases} \tag{6.67}$$

Then the particular is

$$\hat{x}_p = \frac{\alpha\left(\beta_2 - \omega^2\right)}{\left(\beta_2 - \omega^2\right)^2 + \beta_1\beta_2\omega^2} \sin \omega t - \frac{\alpha\beta_1\omega}{\left(\beta_2 - \omega^2\right)^2 + \beta_1\beta_2\omega^2} \cos \omega t \tag{6.68}$$

The general solution will be

$$\hat{x} = \hat{x}_h + \hat{x}_p == c_0 e^{-\beta t} + \frac{\alpha\left(\beta_2 - \omega^2\right)}{\left(\beta_2 - \omega^2\right)^2 + \beta_1\beta_2\omega^2} \sin \omega t - \frac{\alpha\beta_1\omega}{\left(\beta_2 - \omega^2\right)^2 + \beta_1\beta_2\omega^2} \cos \omega t$$

Where $\hat{x}_h = c_0 e^{-\beta t}$ $\hat{x}_p = \dfrac{\alpha\left(\beta_2 - \omega^2\right)}{\left(\beta_2 - \omega^2\right)^2 + \beta_1\beta_2\omega^2} \sin \omega t - \dfrac{\alpha\beta_1\omega}{\left(\beta_2 - \omega^2\right)^2 + \beta_1\beta_2\omega^2} \cos \omega t$ (6.69)

The coupling regression is

$$x^{(-1)}(k) = \alpha \sin(\omega k) + \beta_1\left(-x^{(0)}(k)\right) + \beta_2\left(-z^{(1)}(k)\right), k = 2,3,\cdots,n \tag{6.70}$$

Where $x^{(0)}(k) = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$ is the initial discrete data sequence

$$x^{(-1)}(k-1) = x^{(0)}(k) - x^{(0)}(k-1), k = 2,3,\cdots,n \tag{6.71}$$

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4\cdots n \tag{6.72}$$

$$z^{(1)}(k) = \frac{1}{2}\left[x^{(1)}(k) + x^{(1)}(k-1)\right] \tag{6.73}$$

6.2.4 The Differential Equation with Polynomial Term on Right Side

The basic form of the first order linear differential equation with a polynomial term on right side is

$$\frac{d\hat{x}}{dt} + \beta\hat{x} = \alpha_0 + \alpha_1 t + \alpha_2 t^2 \tag{6.74}$$

In order to satisfy the homogeneous equation

$$\frac{d\hat{x}_h}{dt} + \beta\hat{x}_h = 0 \tag{6.75}$$

We obtain the general solution:

$$\hat{x}_h = c_0 e^{-\beta t} \tag{6.76}$$

The particular solution having a form

$$\hat{x}_p = A_0 + A_1 t + A_2 t^2 \tag{6.77}$$

Satisfying the equation

$$\frac{d\hat{x}_p}{dt} + \beta\hat{x}_p = \left(A_1 + 2A_{2t}\right) + \beta\left(A_0 + A_1 t + A_2 t^2\right) \tag{6.78}$$

Which leads to

$$\begin{cases} a_0 = A_1 + \beta A_0 \\ a_1 = 2A_2 + \beta A_1 \\ a_2 = \beta A_2 \end{cases} \tag{6.79}$$

Then

$$\begin{cases} A_0 = \left(a_0\beta^2 - a_1\beta + 2a_2\right)/\beta^3 \\ A_1 = \left(a_1\beta - 2a_2\right)/\beta^2 \\ A_2 = a_2/\beta \end{cases} \tag{6.80}$$

The particular solution is

$$\hat{x}_p = \frac{a_0\beta^2 - a_1\beta + 2a_2}{\beta^3} + \frac{a_1\beta - 2a_2}{\beta^2}t + \frac{a_2}{\beta}t^2 \tag{6.81}$$

The general solution is

$\hat{x} = \hat{x}_h + \hat{x}_p$

Where $\hat{x}_h = c_0 e^{-\beta t}$ $\hat{x}_p = \dfrac{a_0\beta^2 - a_1\beta + 2a_2}{\beta^3} + \dfrac{a_1\beta - 2a_2}{\beta^2}t + \dfrac{a_2}{\beta}t^2$ (6.82)

$$\hat{x} = c_0 e^{-\beta t} + \frac{a_0\beta^2 - a_1\beta + 2a_2}{\beta^3} + \frac{a_1\beta - 2a_2}{\beta^2}t + \frac{a_2}{\beta}t^2 \tag{6.83}$$

The coupling regression is

$$x^{(0)}(k) = a_0 + a_1 k + a_2 k^2 + \beta\left(-z^{(1)}(k)\right) \tag{6.84}$$

Where $x^{(0)}(k) = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$ is the initial discrete data sequence

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4 \cdots n \tag{6.85}$$

$$z^{(1)}(k) = (1-\omega) * x(k-1) + \omega * x(k) \tag{6.86}$$

For the first order grey differential equation with polynomial term on right side modelling, the minimum sample size n should be no less than 7.

The second order linear differential equation with polynomial term on right side is

$$\frac{d^2\hat{x}}{dt^2} + \beta_1 \frac{d\hat{x}}{dt} + \beta_2 \hat{x} = a_0 + a_1 t + a_2 t^2 \tag{6.87}$$

In order to satisfy the homogeneous equation

$$\frac{d\hat{x}_h}{dt} + \beta \hat{x}_h = 0 \tag{6.88}$$

We obtain the general solution:

$$\hat{x}_h = c_0 e^{-\beta t} \tag{6.89}$$

The particular solution having a form

$$\hat{x}_p = A_0 + A_1 t + A_2 t^2 \tag{6.90}$$

Satisfying the equation

$$\begin{aligned} &\frac{d^2\hat{x}_p}{dt^2}+\beta_1\frac{d\hat{x}_p}{dt}+\beta_2\hat{x}_p \\ &=2A_2+\beta_1A_1+2\beta_1A_2t+\beta_2A_0+\beta_2A_1t+\beta_2A_2t^2 \\ &=a_0+a_1t+a_2t^2 \end{aligned} \tag{6.91}$$

Which leads to

$$\begin{cases} a_0=2A_2+\beta_1A_1+\beta_2A_0 \\ a_1=2\beta_1A_2+\beta_2A_1 \\ a_2=\beta_2A_2 \end{cases} \tag{6.92}$$

Then the coefficients of the particular solution x_p are

$$\begin{cases} A_0=\dfrac{a_0\beta_2^2-a_1\beta_1\beta_2^2+2a_2\beta_2^1\beta_2}{\beta_2^3} \\ A_1=\dfrac{a_1\beta_2-2a_2\beta_1}{\beta_2^2} \\ A_2=\dfrac{a_2}{\beta_2} \end{cases} \tag{6.93}$$

The particular solution is

$$\hat{x}_p=\frac{a_0\beta_2^2-a_1\beta_1\beta_2^2+2a_2\beta_2^1\beta_2}{\beta_2^3}+\frac{a_1\beta_2-2a_2\beta_1}{\beta_2^2}t+\frac{a_2}{\beta_2}t^2 \tag{6.94}$$

The general solution is

$$\hat{x}=\hat{x}_h+\hat{x}_p=c_0e^{-\beta t}+\frac{a_0\beta_2^2-a_1\beta_1\beta_2^2+2a_2\beta_2^1\beta_2}{\beta_2^3}+\frac{a_1\beta_2-2a_2\beta_1}{\beta_2^2}t+\frac{a_2}{\beta_2}t^2$$

Where $\hat{x}_h=c_0e^{-\beta t}$ $\hat{x}_p=\dfrac{a_0\beta_2^2-a_1\beta_1\beta_2^2+2a_2\beta_2^1\beta_2}{\beta_2^3}+\dfrac{a_1\beta_2-2a_2\beta_1}{\beta_2^2}t+\dfrac{a_2}{\beta_2}t^2$ (6.95)

The coupling regression is

$$x^{(-1)}(k)=a_0+a_1k+a_2k^2+\beta_1\left(-x^{(0)}(k)\right)+\beta_1\left(-z^{(1)}(k)\right) \tag{6.96}$$

For the second order grey differential equation with polynomial term on right side modelling, the minimum sample size n should be no less than 8.

6.2.5 The Differential Equation with Product of Exponential and Sine

The basic form of the first order linear differential equation with product of exponential and sine term on right side is

$$\frac{d\hat{x}}{dt} + \beta\hat{x} = \alpha e^{\delta t} \sin \omega t \tag{6.97}$$

In order to satisfy the homogeneous equation

$$\frac{d\hat{x}_h}{dt} + \beta\hat{x}_h = 0 \tag{6.98}$$

We obtain the general solution

$$\hat{x}_h = c_0 e^{-\beta t} \tag{6.99}$$

The particular solution having a form

$$\hat{x}_p = e^{\delta t}\left(A_0 \sin \omega t + B_0 \cos \omega t\right) \tag{6.100}$$

Need be satisfied the equation

$$\begin{aligned}
&\frac{d\hat{x}_p}{dt} + \beta\hat{x}_p \\
&= A_0\delta e^{\delta t} \sin \omega t + B_0\delta e^{\delta t} \cos \omega t \\
&+ A_0\omega e^{\delta t} \cos \omega t - B_0\omega e^{\delta t} \sin \omega t \\
&+ A_0\beta e^{\delta t} \sin \omega t + B_0\beta \cos \omega t \\
&= \alpha e^{\delta t} \sin \omega t
\end{aligned} \tag{6.101}$$

Which leads to

$$\begin{aligned}
&\begin{cases} A_0(\delta+\beta) - B_0\omega = \alpha \\ A_0\omega + B_0(\delta+\beta) = 0 \end{cases} \\
&\Rightarrow \\
&\begin{cases} A_0 = \dfrac{(\delta+\beta)\alpha}{\omega^2 + (\beta+\delta)^2} \\ B_0 = -\dfrac{\alpha\omega}{\omega^2 + (\beta+\delta)^2} \end{cases}
\end{aligned} \tag{6.102}$$

Then the particular solution will be

$$\hat{x}_p = A_0 e^{\delta t} \sin(\omega t + \overline{\omega}) + B_0 e^{\delta t} \cos(\omega t + \overline{\omega}) \tag{6.103}$$

The general solution is

$$\begin{aligned}
&\hat{x} = \hat{x}_h + \hat{x}_p = c_0 e^{-\beta t} + A_0 \sin(\omega t + \overline{\omega}) + B_0 \cos(\omega t + \omega) \\
&\text{Where } \hat{x}_h = c_0 e^{-\beta t} \; \hat{x}_p = A_0 \sin(\omega t + \overline{\omega}) + B_0 \cos(\omega t + \overline{\omega})
\end{aligned} \tag{6.104}$$

The coupling regression is

$$x^{(0)}(t_k) = \alpha e^{\delta t_k} \sin(\omega t_k + \overline{\omega}) + \beta\left(-z^{(1)}(t_k)\right), k = 2,3,\cdots,n \tag{6.105}$$

Where $x^{(0)}(k) = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$ is the initial discrete data sequence

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4\cdots n \tag{6.106}$$

$$z^{(1)}(k) = (1-\theta) * x(k-1) + \theta * x(k) \tag{6.107}$$

The parameter pair (α, β) is obtained by least-square estimation $(a,b)^{T} = \left(X^{T}X\right)^{-1}X^{T}Y$, where

$$X = \begin{bmatrix} e^{\delta t_1} \sin(\omega t_1 + \overline{\omega}) & -z^{(1)}(t_1) \\ e^{\delta t_2} \sin(\omega t_2 + \overline{\omega}) & -z^{(1)}(t_2) \\ e^{\delta t_3} \sin(\omega t_3 + \overline{\omega}) & -z^{(1)}(t_3) \\ e^{\delta t_4} \sin(\omega t_4 + \overline{\omega}) & -z^{(1)}(t_4) \end{bmatrix}, Y = \begin{bmatrix} z^{(0)}(t_1) \\ z^{(0)}(t_2) \\ \vdots \\ z^{(0)}(t_n) \end{bmatrix} \tag{6.108}$$

Since δ and ω are given

Example 6.2.4: the data we used in table 6.2.1 is an exploration of the possible dynamics of the cement roller functioning times. It will help us for the purpose to explore whether the new statistical-grey model could well reveal the underlying mechanism behind the data set.

Table 6.2.1 Cement Roller data

Functioning	Failure	Covariate	Covariate	Covariate	Repair
54	pm	12	10	800	93
133	failure	13	16	1200	142
147	pm	15	12	1000	300
72	failure	12	15	1100	237
105	failure	13	16	1200	0
115	pm	11	13	900	525
141	pm	16	13	1000	493
59	failure	8	16	1100	427
107	pm	9	11	800	48
59	pm	8	10	900	1115
36	failure	11	13	1000	356

Now we use new statistical-grey consistent differential equation model with product of exponential and sine term to estimate function time in table 5.3

Step1. According to The coupling regression function

$$x^{(0)}(t_k) = \alpha e^{\delta t_k}\sin(\omega t_k + \overline{\omega}) + \beta\left(-z^{(1)}(t_k)\right), k = 2,3,\cdots,n \tag{6.109}$$

To fit the functioning time data $X^{(0)} = (54,133,147,72,105,115,141,59,107,59,36)$ and estimate the parameters α, δ, ω, $\overline{\omega}$, β, θ.

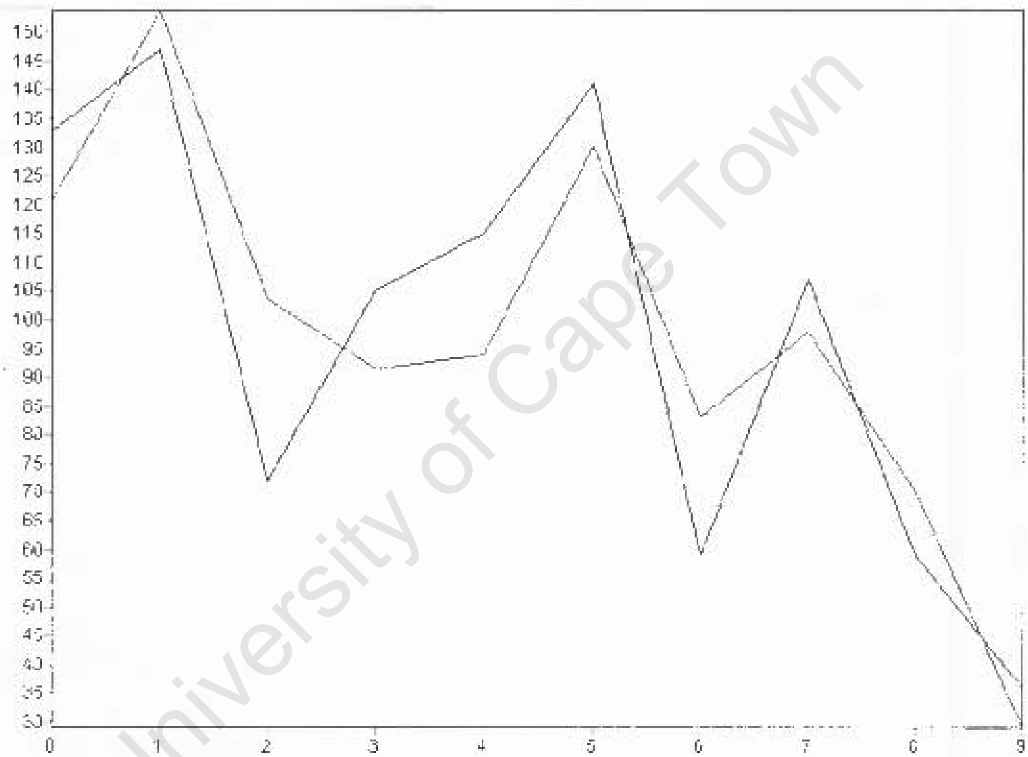

Figure 6.2.10 the coupling regression

Where blue line stands as functioning time data, red line stands as estimate data sequence.,

The regression result information we list as follows:

Optimization algorithm: Genetic algorithm

(RMSE): 16.6536014749416

(SSE): 2773.42442086178

(R): 0.891007090394582

(R^2): 0.793893635133418

(DC): 0.793834228772428

(Chi-Square): 16.1201415585113

(F-Statistic): 30.8149099867858

Table 6.2.2 best estimate parameters

Parameters	Best estimate
α	-1083.24678250239
δ	0.0120878405592725
ω	6.16633109749509
$\overline{\omega}$	3.18043250854117
β	-1.10361913789015
θ	9.13450734807867E-8

Table 6.2.3 Observation data vs. fitted data

No	Observation data T	Fitted data $\hat{T}$	Difference D
1	133	120.9269783	12.0730
2	147	153.7186621	-6.7187
3	72	103.5698702	-31.5699
4	105	91.4757710	13.5242
5	115	93.8787938	21.1212
6	141	130.2588807	10.7411
7	59	83.2180889	-24.2181
8	107	97.9649364	9.0351
9	59	70.2096111	-11.2096
10	36	29.0946332	6.9054

Step2. From statistics information (R):0.891007090394582, (R^2): 0.793893635133418, we know the regression isn't performing very well, so we need to further explore the difference sequence $\{D_i\}$.Now we need run another round statistical-grey differential equation fitting : $\frac{dD}{dt} = \alpha\sin(\omega t) + \beta D$, try to obtain a highest R^2 (approaching 1).

Fitting: $\frac{d|D|}{dt} + \beta_1 |D| = \alpha e^{\gamma t} \sin \omega t$ using coupling regression

$$|D|(t_k) = \alpha e^{\gamma t_k} \sin(\omega t_k + \overline{\omega}) + \beta\left(-z^{(1)}(t_k)\right), k = 2,3,\cdots,n$$

The difference data sequence has negative numbers, so we directly take the absolute value. $\{|D_i|\}$ = (12.0730,6.7187,31.5699,13.524221.1212,10.7411,24.2181,9.0351,11.2096,6.9054)

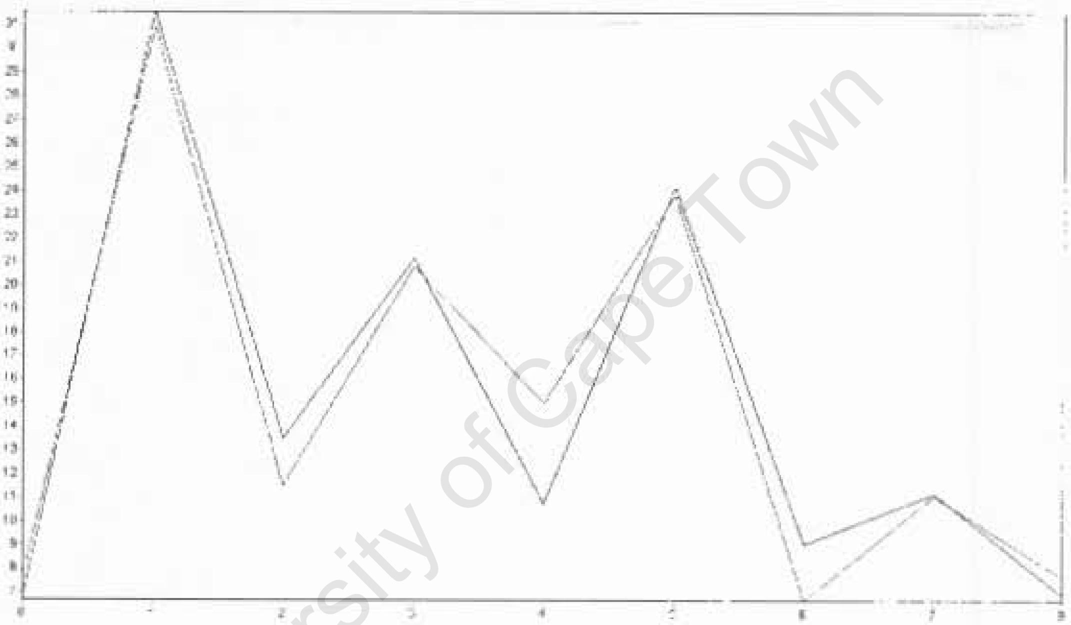

Figure 6.2.11 statistical-grey differential equation fitting: $\frac{d|D|}{dt} = \alpha \sin(\omega t) - \beta|D_i|$

Optimization algorithm: (Levenberg-Marquardt) + (General global optimization)

(RMSE): 1.82687738657493

(SSE): 30.0373288702097

(R): 0.974706019625614

(R^2): 0.950051824694408

(DC): 0.95005180799867

(Chi-Square): 1.341417993262

(F-Statistic): 133.14525970514

Table 6.2.4 best estimate parameters

Parameters	Best estimate
α	-251.586239698936
δ	0.00318597752310604
ω	6.42178672156538
$\overline{\omega}$	-34.5853403254242
β	1.67942647905306
θ	0.999999999977642

Table 6.2.5 Observation data vs. fitted data

No	Observation data D_i	Fitted data $\hat{D}_i$
1	6.1787	7.6268620
2	31.5699	30.9303348
3	13.5242	11.5161471
4	21.1212	20.7732794
5	10.7411	15.0107950
6	24.2181	23.8335301
7	9.0351	6.6399246
8	11.2096	11.0887255
9	6.9054	7.6268824

Table 6.2.6 True data D_i, fitted data $\hat{D}_i$ and residual $e = D_i - \hat{D}_i$

No	True data D_i	Fitted data $\hat{D}_i$	Residual $e = D_i - \hat{D}_i$
1	-6.1787	-7.6268620	0.9082
2	-31.5699	-30.9303348	-0.6396
3	13.5242	11.5161471	2.0081
4	21.1212	20.7732794	0.3479
5	10.7411	15.0107950	-4.2697
6	-24.2181	-23.8335301	-0.3846
7	9.0351	6.6399246	2.3952
8	-11.2096	-11.0887255	-0.1209
9	6.9054	7.6268824	-0.7215

Step3. Then we will have the residual part $e = D_i - \hat{D}_i$. The interpretation of the partition is $T_i = \hat{T}_i + \hat{D}_i + \hat{e}_i$, where

$\hat{T}_i$ are the average functioning time during i^{th} operation;

$\hat{D}_i$ are the average repair effects due to previous repairs;

$\hat{e}_i$ are the residuals after the two stage Statistical-grey consistent grey differential equation modelling, which follows normal random fuzzy $N(\zeta, \sigma^2)$ approximately, where fuzzy mean ζ is triangular $(-\sigma, 0, \sigma)$, therefore, we just need to estimate parameter $(0, \sigma)$.

From Table 6.2.9 we have e= [0.9082; -0.6396; 2.0081; 0.3479; -4.2697; -0.3846; 2.3952; -0.1209;-0.7215;]

Then we fit e residual sequence:

```
>> [muhat,sigmahat] = normfit(e)

muhat =

   -0.0530


sigmahat =

    1.9369
```

Parameter: Fuzzy Mean $\zeta = -0.0530$; Standard Deviation $\sigma = 1.9369$

Perform another round Statistical-grey consistent model fitting with respect to data sequence $t = \hat{T} + \hat{e}$, using the first round parameter estimates as initial value and thus you should obtain the refitted first-stage Statistical-grey consistent model with high R^2 regression.

$t = \hat{T} + \hat{e} =$ [122.8639; 155.6555; 105.5068; 93.4127; 95.8157; 132.1958; 85.1550; 99.9018; 72.1465; 31.0315;]

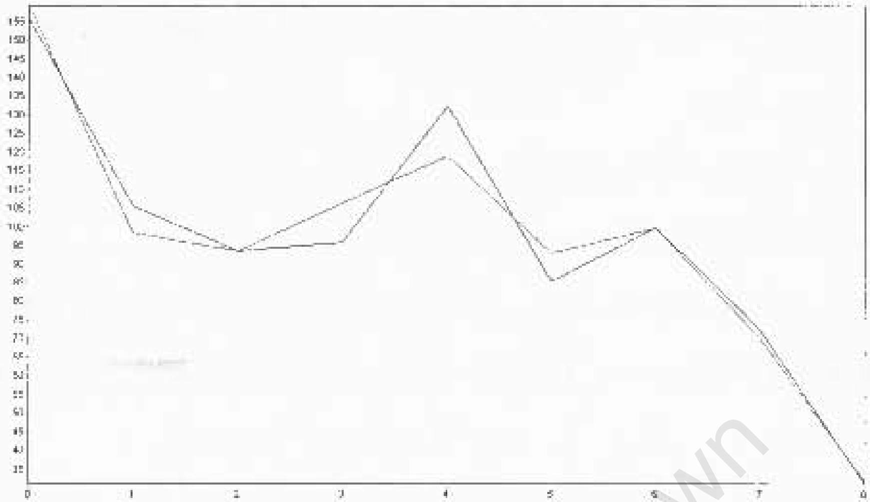

Figure 6.2.12 statistical-grey differential equation fitting: $t = \hat{T} + \hat{e}$ data sequence

(Levenberg-Marquardt) + General global optimization

(RMSE): 6.84453238417813

(SSE): 421.628612022568

(R): 0.978437796195042

(R^2): 0.95734052102301

(DC): 0.957340520427767

(Chi-Square): 1.94873467905206

(F-Statistic): 157.090142867794

Table 6.2.7 Best estimate parameters

Parameters	Best estimate
α	297.022936229672
δ	0.159869416687138
ω	0.172566172370957
$\overline{\omega}$	-43.244078717087
β	0.944387725482203
θ	0.999995222739004

Table 6.2.8 Observation data vs. fitted data

No	**Observation data** t	**Fitted data** $\hat{t}$
1	155.6555	159.2890905
2	105.5068	98.2735500
3	93.4127	106.4247326
4	95.8157	20.7732794
5	132.1958	118.9278467
6	85.155	93.0192455
7	99.9018	99.5302025
8	72.1465	69.9443415
9	31.0315	31.8434950

Step4. From the general result equation of statistical-grey consistent differential equation model we have:

$$x = c_1 e^{-\beta t} + A_0 e^{\delta t}\sin(\omega t+\varpi) + B_0 e^{\delta t}\cos(\omega t+\varpi) \tag{6.110}$$

$$\begin{cases} A_0 = \dfrac{(\delta+\beta)\alpha}{\omega^2+(\beta+\delta)^2} \\ B_0 = -\dfrac{\alpha\omega}{\omega^2+(\beta+\delta)^2} \end{cases} \tag{6.111}$$

According to:

$$z^{(1)}(t) = (1-\theta)*x^{(1)}(t)+\theta*x^{(1)}(t-1), t=2,3,\cdots,n$$

$\alpha =$ 297.022936229672 $\delta =$ 0.159869416687138 $\omega =$ 0.172566172370957 (6.112)

$\varpi =$ -43.244078717087 $\beta =$ 0.944387725482203 $\theta =$ 0.999995222739004

We will easily to have

$C_1 = -29.3501$ $A_0 = 262.5676$ $B_0 = -41.0324$

Then will easily to have the estimate value $\hat{X}^{(1)}$ on $X^{(1)}$ level:

$X^{(1)} =$ [54187 334 406 511 626 767 826 933 992 1028]

$\hat{X}^{(1)} =$ [54177.9472 278.9644 377.3467 479.5188 585.645 692.4145 793.9515 882.0097 945.9181 972.4772]

Accuracy=0.9363=93.63%

Then we are using ratio idea to have the $\hat{X}^{(0)}$ value on $X^{(0)}$ level

$X^{(0)}$ = [54 133 147 72 105 115 141 59 107 59 36];

$X^{(1)}$ = [54 187 334 406 511 626 767 826 933 992 1028]

$$\text{Ratio}=\frac{X^{(1)}}{X^{(0)}}=[1, 1.406, 2.2721, 5.6389, 4.8667, 5.4435, 5.4397, 14, 8.7196, 16.8136]$$

$$\hat{X}^{(0)}=\frac{\hat{X}^{(1)}}{ratio}=[54 \quad 126.5614 \quad 122.7777 \quad 66.9186 \quad 98.5313 \quad 107.5865 \quad 127.2887 \quad 56.7108 \quad 101.1522 \quad 56.2592 \quad 34.0556]$$

Accuracy=0.9363=93.63%

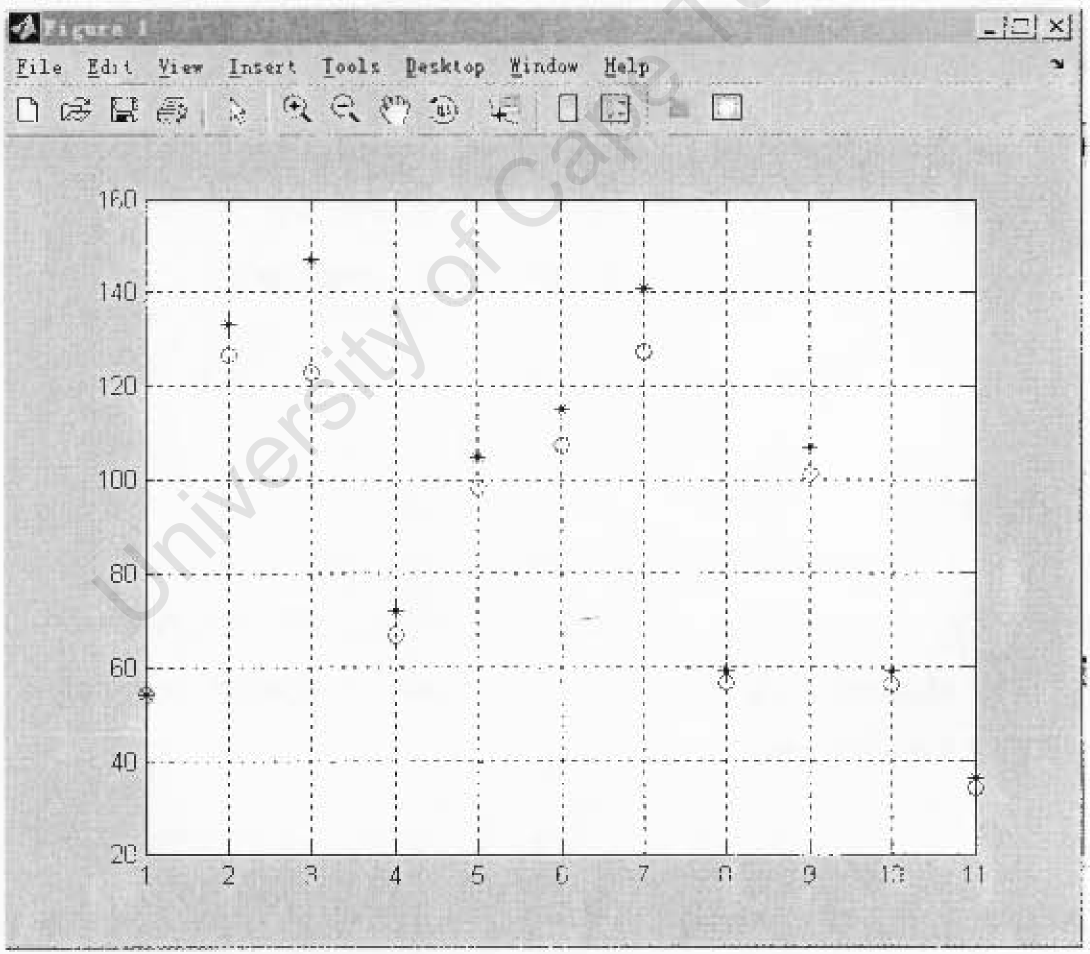

Figure 6.2.13 $\hat{X}^{(0)}$ data sequence compare with $X^{(0)}$ data sequence

6.3 Summary

We first introduced remnant methods to deal with the data sequence which not following the grey index and statistical index law in previously research, after that we represented our recently research outcome to build different formed grey differential equation models to deal with different kind of formed data sequence (For identify the data sequence utilize what kind of Types of differential equations, we should develop in future's research). The new model include: the differential equation with constant term on right side (Type I), the differential equation with exponential term on right side (Type II), the differential equation with polynomial term on right side (Type V*), the differential equation with sine and cosine term on right side (Type III), the differential equation with product of exponential and sine term on right side (Type IV). For easier to understand new statistical-grey differential equation models, we try to give each of them an example and utilize optimization algorithm help us to complete the computation of each model.

Chapter 7. MATLAB toolbox and VC++ program

7.1 Methods of Building GM (1, 1) modelling program

Based on GM (1, 1) model procedure, utilize visual C++ and Matlab language program to build software for GM (1, 1) modelling.

Program structure:

Import the sample size of the modelling data sequence (denotes as K) from keyboard

↓

Assume $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$ is the modelling data sequence, then $\sigma(k) = \dfrac{x^{(0)}(k)}{x^{(0)}(k-1)}$, $k = 1, 2, \cdots, n$, are called stepwise ratios of the sequence $X^{(0)}$ and $\exp(-2/(k+1)) \le \sigma(k) \le \exp(2/(k-1))$

↓

Import the modelling data sequence $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$ from keyboard

↓

Take 1-time accumulating generation operation of $X^{(0)}$, to generate the data sequence. $X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right)$

↓

Check whether the modelling data sequence $X^{(0)}$ obeys the stepwise ratio law. $\exp(-2/(k-1)) < \sigma(x) \le \exp(2/(k+1))$

↓

If the modeling data sequence doesn't obeys the ratio law, then warning the program user, and suggest to quite the running program, otherwise the program will be going on to next step.

↓

Generate $z^{(1)}$ data sequence from $z^{(1)} = \frac{1}{2}\left(x^{(1)}(k) + x^{(1)}(k-1)\right)$

↓

Utilizing least squares estimate the parameter sequence $\hat{a} = [a, b]^T = [B^T B]^{-1} B^T Y$, $B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}$, $Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}$

↓

Generate the grey model and the time response sequence from $\hat{x}^{(1)}(k+1) = \left[x^{(0)}(1) - \frac{b}{a} \right] e^{-ak} + \frac{b}{a}$

↓

Simulate value of $X^{(1)}$ and estimation value of $X^{(0)}$ from equation $\hat{x}^{(0)}(k) = \hat{x}^{(1)}(k) - \hat{x}^{(1)}(k-1)$

↓

Calculate the Errors and Relative Errors

↓

Type 'Y' from keyboard when finished typing keys

Example: 7.1.1: Let $X^{(0)} = (2.874, 3.278, 3.337, 3.390, 3.679)$, we strictly obey the GM (1, 1) modelling procedure law and utilize visual C++ program to simulate $X^{(0)}$, and evaluate the accuracy of $X^{(0)}$. (Visual C++ program code please check appendix A)

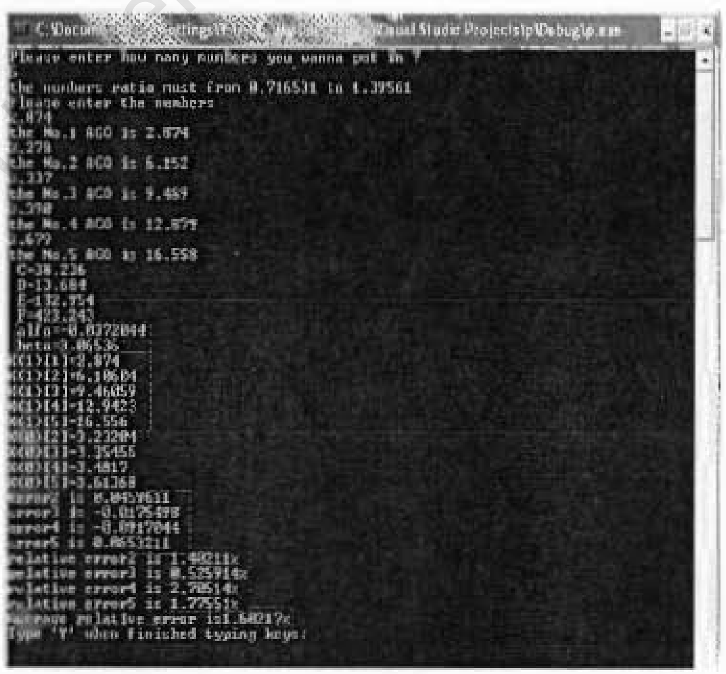

Figure 7.1.1 Visual C++ language program result of GM (1, 1) modelling

Example: 7.1.2: Let $X^{(0)} = (2.874, 3.278, 3.337, 3.390, 3.679)$, we strictly obey the GM (1, 1) modelling procedure law and utilize Matlab program to simulate $\hat{X}^{(0)}$, and evaluate the accuracy of $\hat{X}^{(0)}$. (Matlab program code please check appendix B)

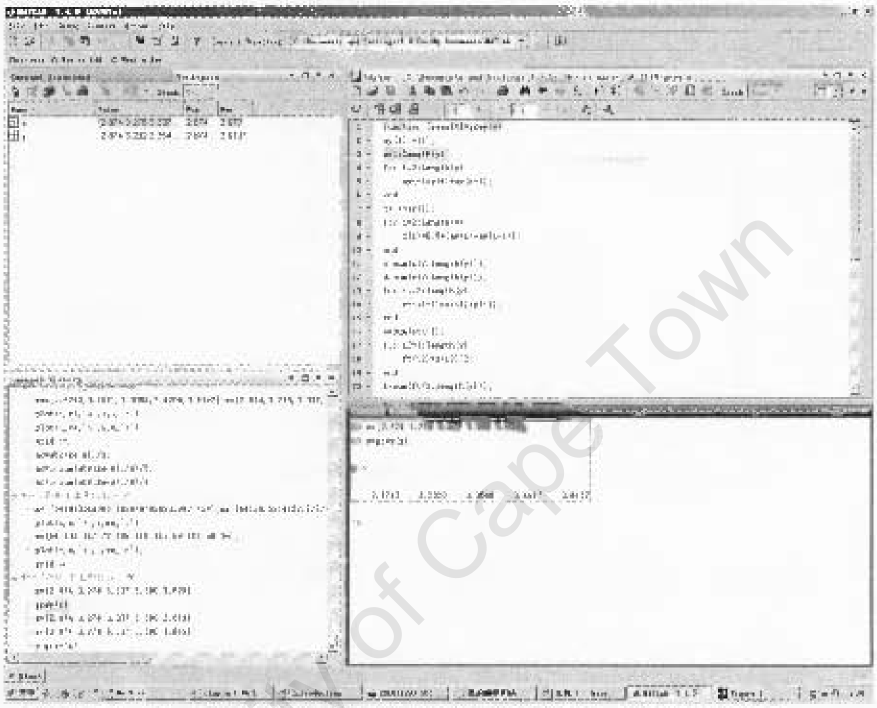

Figure 7.1.2 Matlab program result of GM (1, 1) modelling

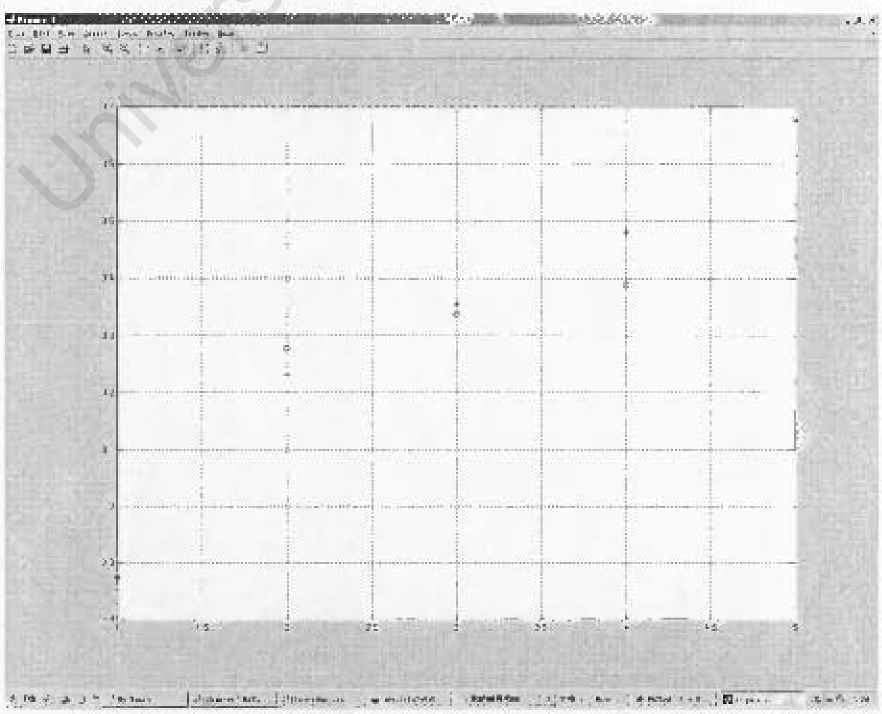

Figure 7.1.3 Matlab program image result of GM (1, 1) modelling

7.2 Methods of Building GM (1, 1) prediction modelling program

Based on GM (1, 1) prediction model procedure, utilize visual C++ language program to build software for GM (1, 1) model prediction.

Program structure:

Import the sample size of the modelling data sequence (denotes as K) from keyboard

↓

Assume $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$ is the modelling data sequence, then $\sigma(k) = \dfrac{x^{(0)}(k)}{x^{(0)}(k+1)}$, $k = 1,2,\cdots,n$, are called stepwise ratios of the sequence $X^{(0)}$ and $\exp(-2/(k+1)) \le \sigma(x) \le \exp\left(2/(k+1)\right)$

↓

Import the modelling data sequence $X^{(0)} = \left(x^{(0)}(1), x^{(0)}(2), \cdots, x^{(0)}(n)\right)$ from keyboard

↓

Take 1-time accumulating generation operation of $X^{(0)}$, to generate the data sequence. $X^{(1)} = \left(x^{(1)}(1), x^{(1)}(2), \cdots, x^{(1)}(n)\right)$

↓

Check whether the modelling data sequence $X^{(0)}$ obeys the stepwise ratio law. $\exp(-2/(k+1)) \le \sigma(x) \le \exp(2/(k+1))$

↓

Generate $z^{(1)}$ data sequence from $z^{(1)} = \dfrac{1}{2}\left(x^{(1)}(k) + x^{(1)}(k-1)\right)$

↓

Utilizing least squares estimate the parameter sequence $\hat{a} = [a,b]^T = \left[B^T B\right]^{-1} B^T Y$, $B = \begin{bmatrix} -z^{(1)}(2) & 1 \\ -z^{(1)}(3) & 1 \\ \vdots & \vdots \\ -z^{(1)}(n) & 1 \end{bmatrix}$, $Y = \begin{bmatrix} x^{(0)}(2) \\ x^{(0)}(3) \\ \vdots \\ x^{(0)}(n) \end{bmatrix}$

Enter how many data you want to predict from keyboard

Generate the grey model and the time response sequence from
$$\hat{x}^{(1)}(k+1)=\left[x^{(0)}(1)-\frac{b}{a}\right]e^{-ak}+\frac{b}{a}$$

Simulate value of $X^{(1)}$ and estimation value of $X^{(0)}$ from equation $\hat{x}^{(0)}(k)=\hat{x}^{(1)}(k)-\hat{x}^{(1)}(k-1)$

Simulate value of predict data

Calculate the Errors and Relative Errors

Type 'Y' from keyboard when finished typing keys

Example: 7.2.1: Let $X^{(0)}=(2.874,3.278,3.337,3.390,3.679)$, we follow the GM (1, 1) predict modelling procedure law and utilize visual C++ program to simulate predict value of $X^{(0)}$, and evaluate the accuracy of $\hat{X}^{(0)}$. (Visual C++ program code please check appendix C)

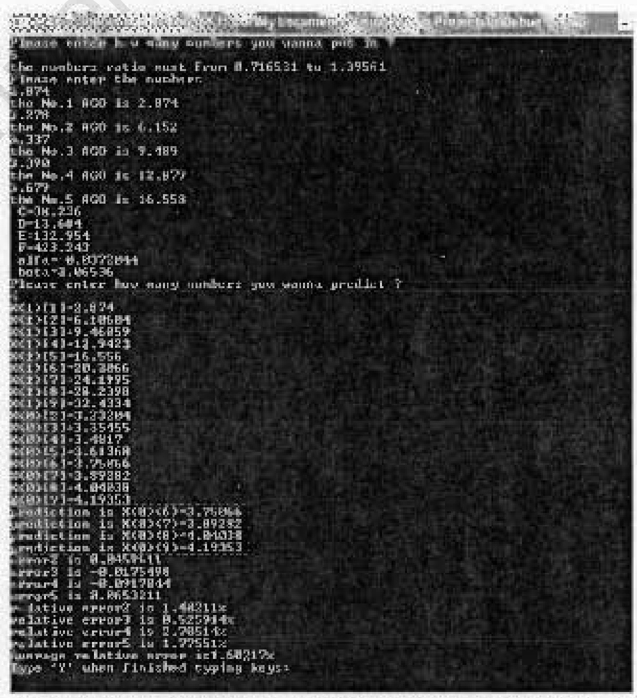

Figure 7.2.1 Visual C++ language program result of GM (1, 1) predict modelling

7.3 Methods of Building Statistical-Grey Consistency GM (1, 1) modelling program

Based on Statistical-Grey Consistency GM (1, 1) model procedure, utilize Matlab program to build Matlab toolbox.

Program structure:

The program structure of Statistical-Grey Consistency GM (1, 1) model base on the program structure of GM (1, 1) modelling procedure and combined with lot of necessary information of modelling export, such as: 1-AGO data sequence value $X^{(1)}$; the parameters of regression model α, β; the estimation value of 1-AGO data sequence value $\hat{X}^{(1)}$; $z^{(1)}$ data sequence value from $z^{(1)} = \frac{1}{2}\left(x^{(1)}(k) + x^{(1)}(k-1)\right)$; R^2 value of regression model; R^2 value of Statistical-Grey Consistency model; error term of $X^{(1)}$ level; error term of Statistical-Grey Consistency model; average error of Statistical-Grey Consistency model; Confidence of Statistical-Grey Consistency model.

Example: 7.3.1: Let $X^{(0)} = (2.874, 3.278, 3.337, 3.390, 3.679)$, we follow Statistical-Grey Consistency GM (1, 1) modelling procedure law and utilize Matlab program to simulate $\hat{X}^{(0)}$, and evaluate the accuracy of $\hat{X}^{(0)}$. (Matlab program code please check appendix D)

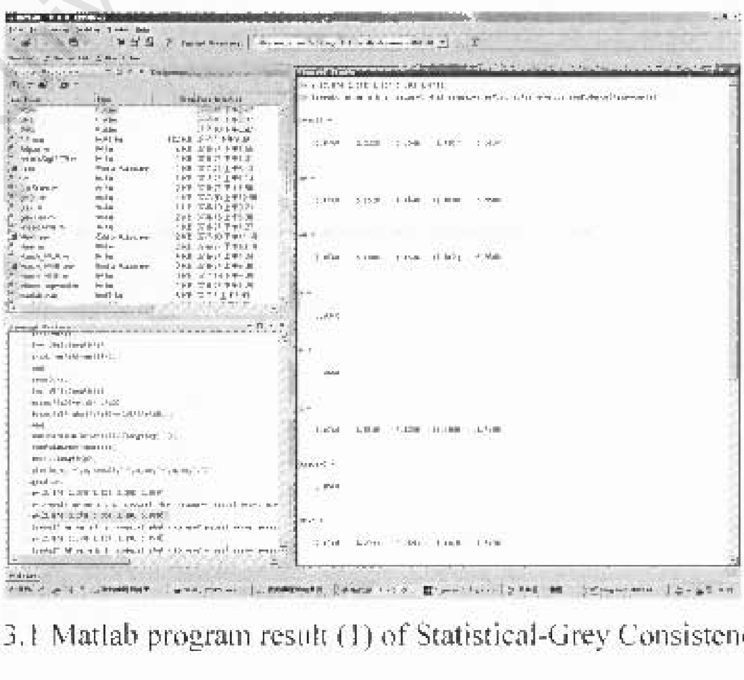

Figure 7.3.1 Matlab program result (1) of Statistical-Grey Consistency model

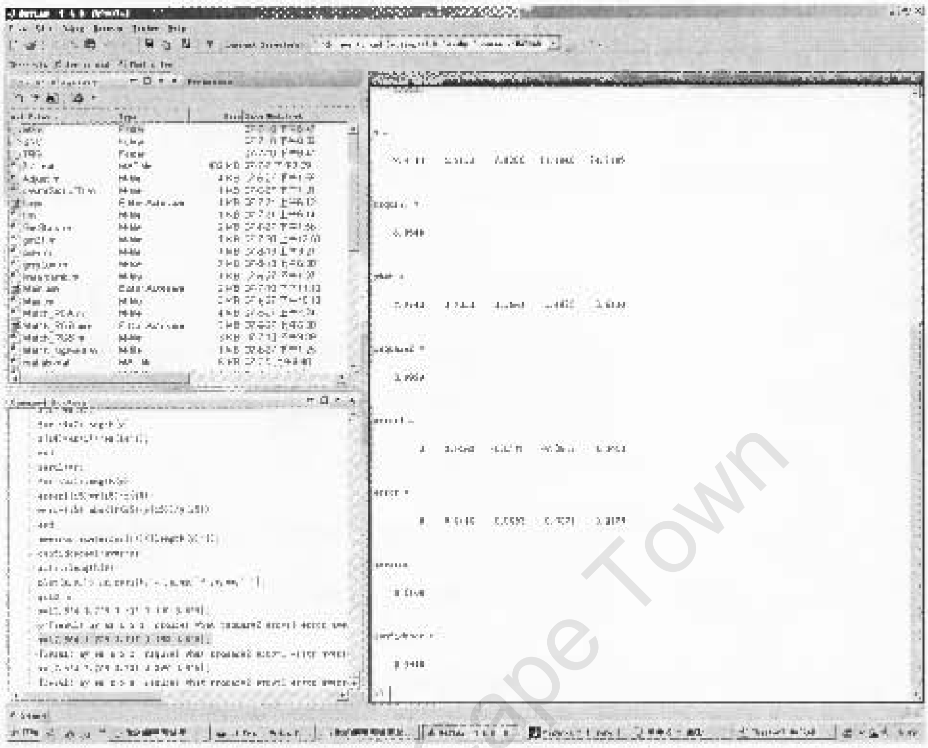

Figure 7.3.2 Matlab program result (2) of Statistical-Grey Consistency model

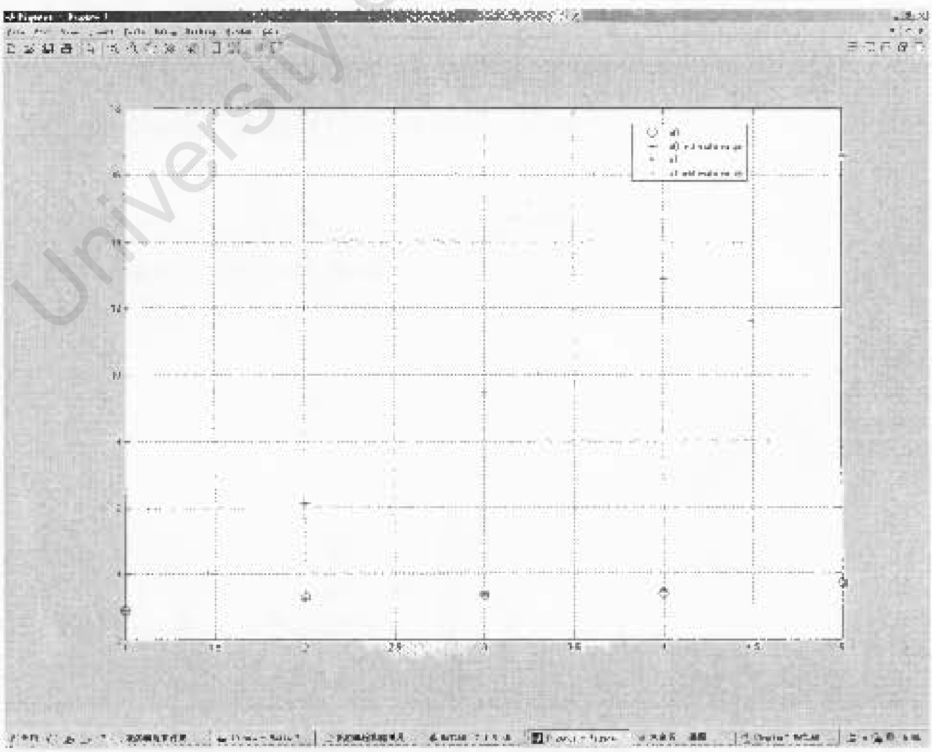

Figure 7.3.3 Matlab image result of Statistical-Grey Consistency model

7.4 Methods of Building Statistical-Grey Consistency with new ratio idea GM (1, 1) modelling program

Based on Statistical-Grey Consistency GM (1, 1) model procedure and combined with new ratio idea, utilize Matlab program to build Matlab toolbox.

Program structure:

The structure is as same as the Statistical-Grey Consistency GM (1, 1) model, but with new ratio idea instead of traditional 1-AGO operation.

Example: 7.4.1: Let $X^{(0)} = (2.874, 3.278, 3.337, 3.390, 3.679)$, we follow Statistical-Grey Consistency with new ratio idea GM (1, 1) modelling procedure law and utilize Matlab program to simulate $\hat{X}^{(0)}$, and evaluate the accuracy of $\hat{X}^{(0)}$. (Matlab program code please check appendix E)

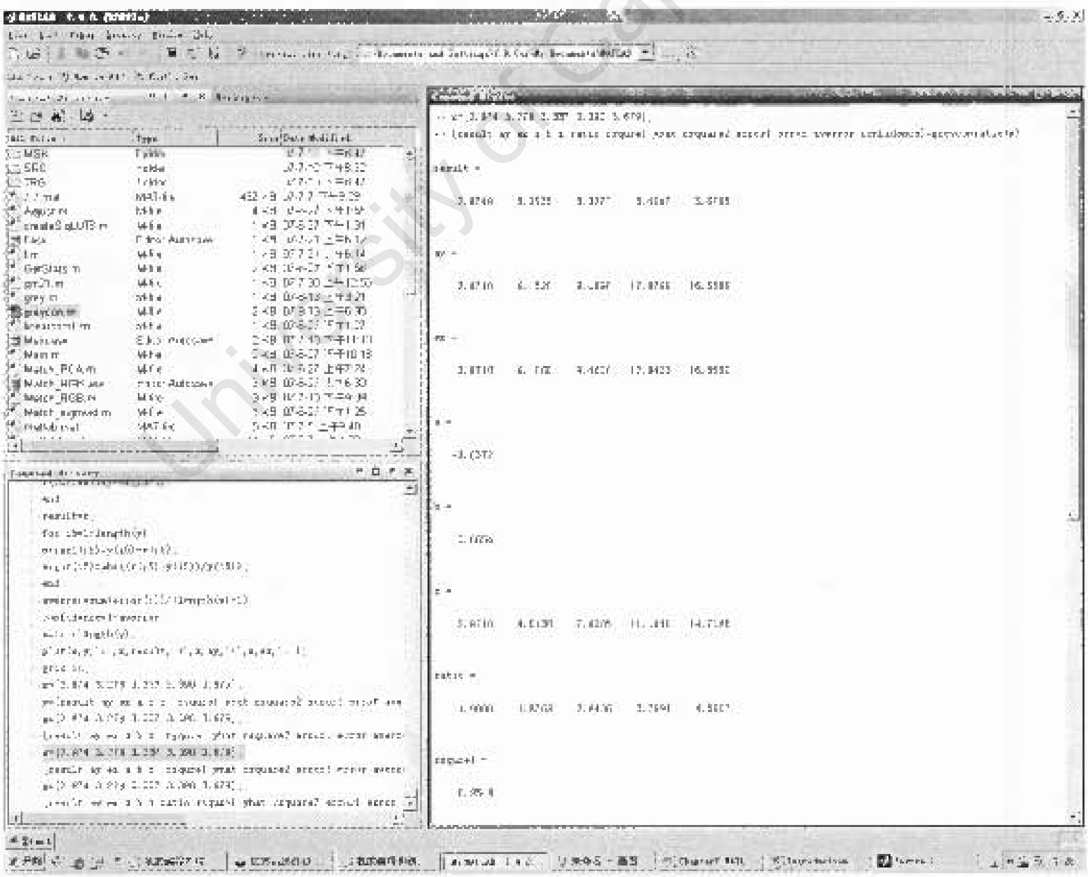

Figure 7.4.1 Matlab program result (1) of Statistical-Grey Consistency model with new ratio idea

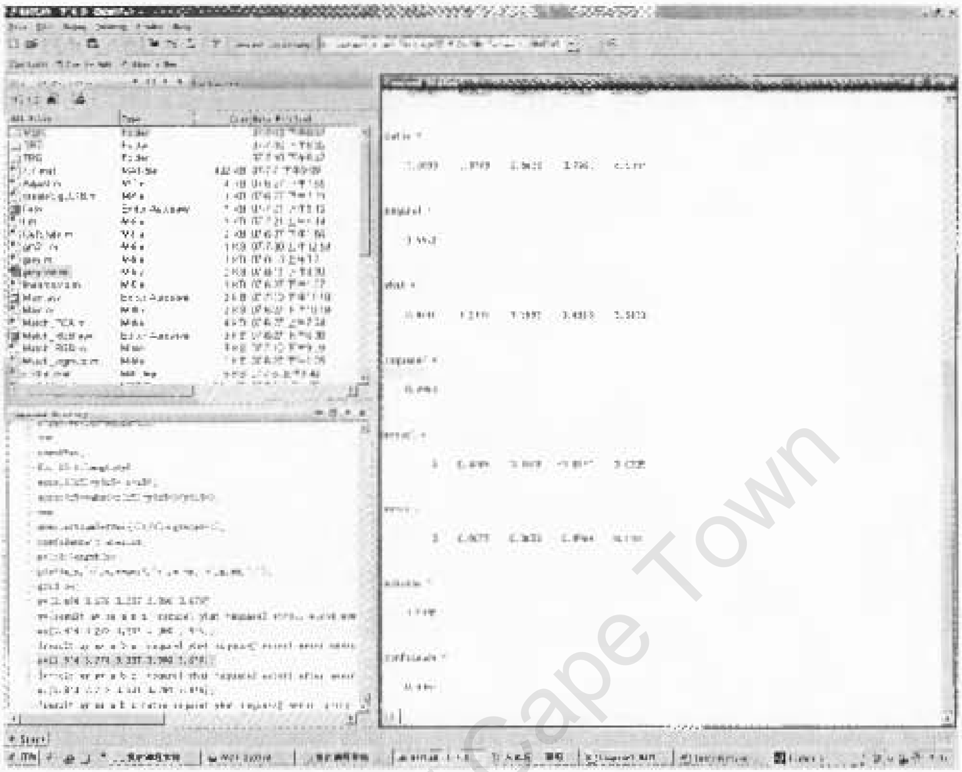

Figure 7.4.2 Matlab program result (2) of Statistical-Grey Consistency model with new ratio idea

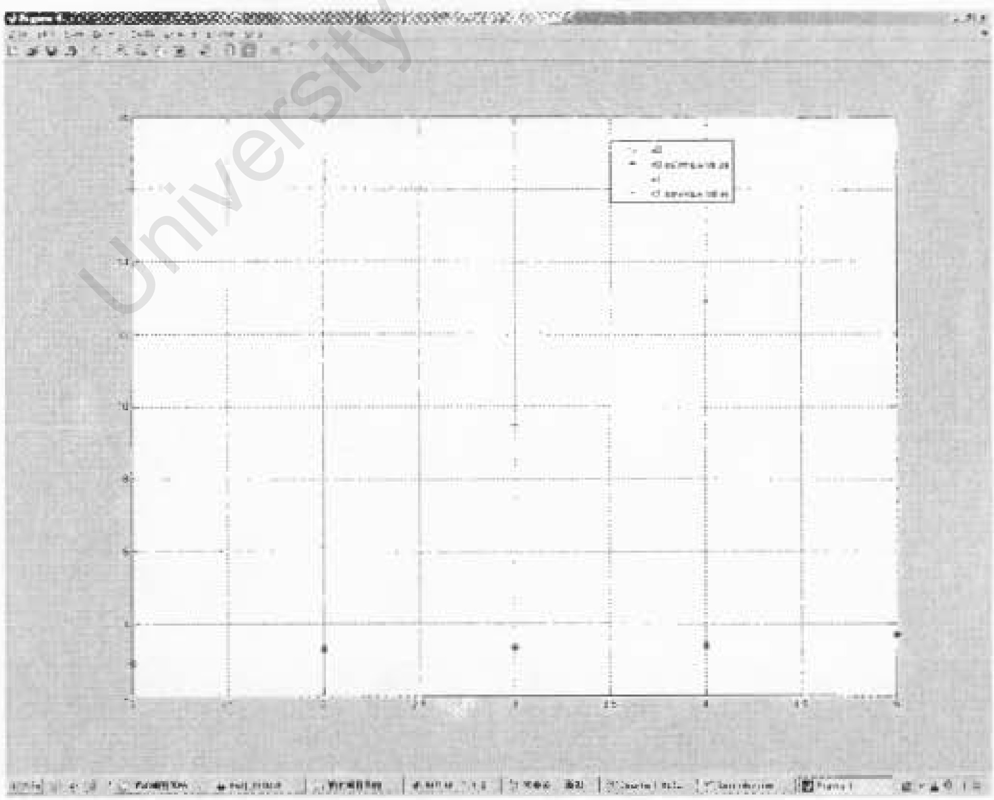

Figure 7.4.3 Matlab image result of Statistical-Grey Consistency model with new ratio idea

7.5 Methods of building statistical-grey consistency with new flowchart method of GM (1, 1) differential equation modelling with constant term on right side program

Based on Statistical-Grey Consistency GM (1, 1) model procedure and combined with new flowchart idea, utilize Matlab program to build Matlab toolbox.

Program structure:

The structure is as same as the Statistical-Grey Consistency GM (1, 1) model, but with new flowchart method instead of traditional half-weighted operation.

Example: 7.5.1: Let $X^{(0)} = (2.874, 3.278, 3.337, 3.390, 3.679)$, we follow Statistical-Grey Consistency with new flowchart method GM (1, 1) modelling procedure law and utilize Matlab program to simulate $\hat{X}^{(0)}$, and evaluate the accuracy of $\hat{X}^{(0)}$. (Matlab program code please check appendix F)

Steps: we rewrite equation (6.24) as

$$\underset{\min}{objectfunction} = x^{(0)}(k) - \left(\alpha + \beta\left(-z^{(1)}(k)\right)\right) \tag{7.1}$$

Where

$$x^{(1)}(k) = \sum_{i=1}^{k} x^{(0)}(i), k = 2,3,4 \cdots n \tag{7.2}$$

$$z^{(1)}(k) = \omega * x(k) + (1-\omega) * x(k+1), 0 \le \omega \le 1 \tag{7.3}$$

After Matlab genetic toolbox optimization we could estimate $\hat{\alpha}, \hat{\beta}$,

$\hat{\alpha} = 3.0327, \hat{\beta} = -0.0386, \omega = 0.005$

From equation (6.23) we could generate $\hat{x}^{(0)} = (2.874, 3.2051, 3.3313, 3.4624, 3.5987)$, the accuracy of model is $accuracy = 1 - 0.0168 = 98.32\%$

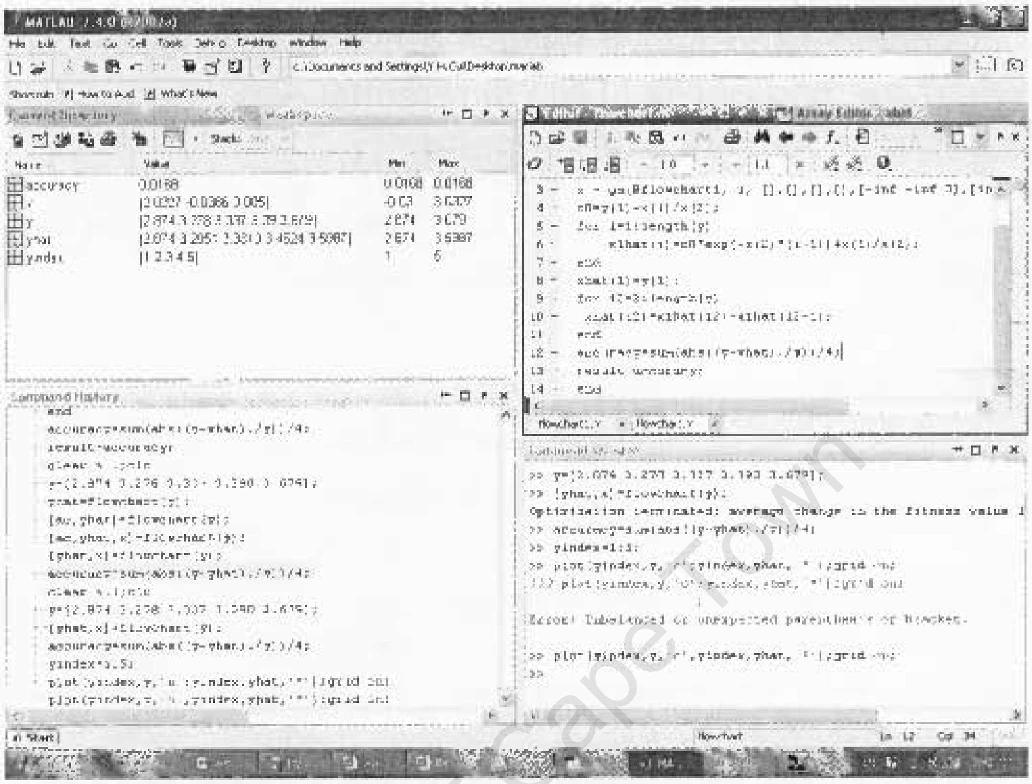

Figure 7.5.1 Matlab program result of flowchart model calculating

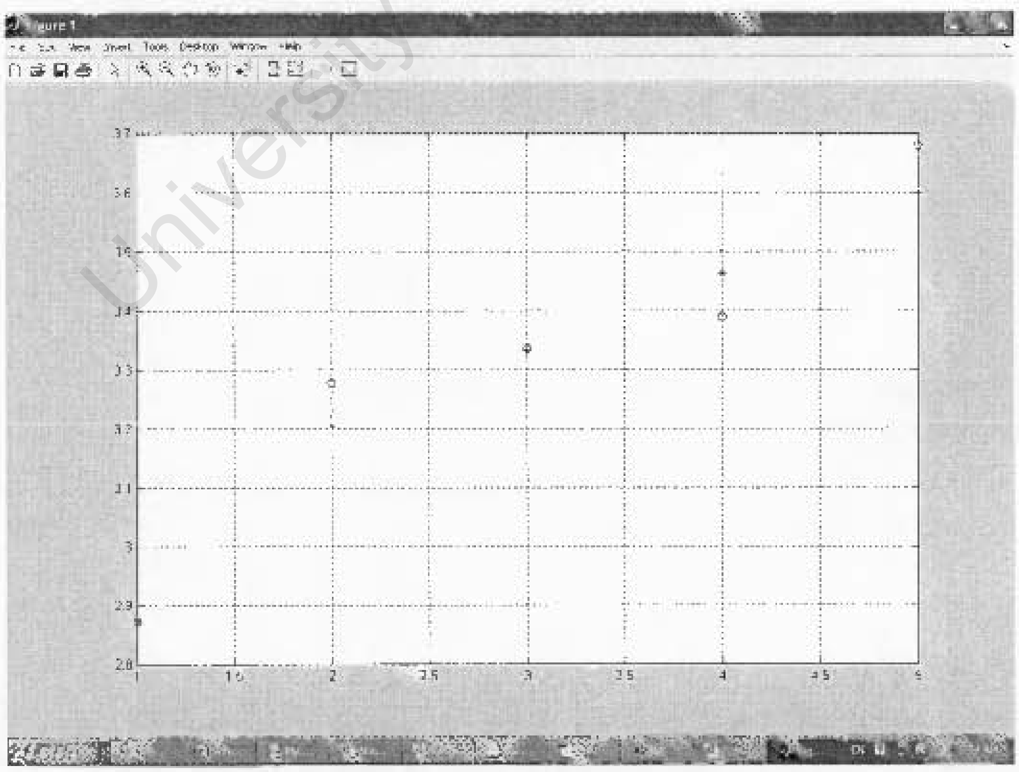

Figure7.5.2 Plot of $x^{(0)}$ data sequence and $\hat{x}^{(0)}$ data sequence

7.6 Methods of building statistical-grey consistency with new flowchart method of GM (1, 1) differential equation modelling with exponential term on right side program

Based on Statistical-Grey Consistent differential equation model with exponential term on right side procedure and combined with new flowchart idea, utilize Matlab program to build Matlab toolbox.

Program structure:

The structure is as same as the Statistical-Grey Consistent differential equation model with exponential term on right side procedure, and with new flowchart method instead of traditional half-weighted operation.

Example 7.6.1: let the discrete positive data sequence $X^{(0)} = [2.874, 3.278, 3.337, 3.390, 3.679]$, and then we use new differential equation (6.35) to calculate the estimated value $\hat{X}^{(0)}$ data sequence.

Setp1. Rewrite the equation (6.39) as (7.4), then made an M-file named "nlin.m" to express the modified equation (7.4). Then we use Matlab optimization toolbox to simulate parameters $\alpha, \beta\ \delta$ (denoted as p (1), p (2), and p (3), respectively) to minimize the object function (7.4):

$$\underset{nlin}{objectfun} = x^{(0)}(k) - \left(\alpha e^{\delta k} - \beta\left(-z^{(1)}(k)\right)\right) \tag{7.4}$$

```
function nlin=nlin(p)
x=[2.874;3.278;3.337;3.390;3.679];
xx=[3.278;3.337;3.390;3.679];
x1=cumsum(x);
for i=2:length(x1)
z(i-1)=0.5*(x1(i-1)+x1(i));
end
k=[1;2;3;4];
nlin=xx-(p(1)*exp(p(3)*k)+p(2)*(-z'));
```

Step2. Using Matlab optimization toolbox, Nonlinear least-square method, medium scale-Gauss-Newton algorithm to estimate the parameter α, β, δ.

Then we can easily estimate the parameter

$$\alpha = 3.211; \beta = -0.16; \delta = -0.224$$

And the minimized object function value equal to 0.003

$$\underset{\min}{objectfun} = 0.003$$

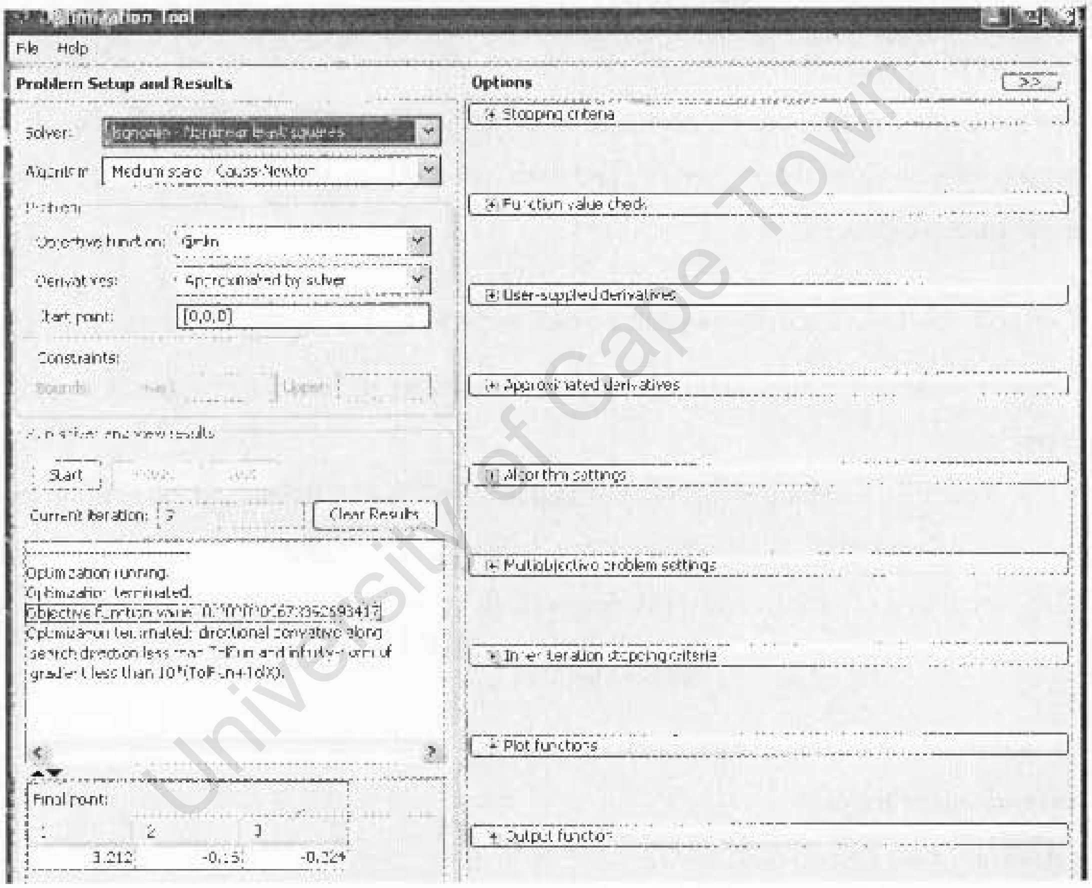

Figure 7.6.1 Utilize Optimization tool to estimate Parameter α, β, δ

Step3. Put the first initial value 2.874 into general solution to calculate constant value C_0

$$\begin{aligned} 2.874 &= c_0 e^{-\beta \cdot 0} + \frac{\alpha}{(\delta + \beta)} e^{\delta \cdot 0} \\ c_0 &= 2.874 - \frac{\alpha}{(\delta + \beta)} = 11.2386 \end{aligned} \tag{7.5}$$

Step4. Using general solution

$$x = c_0 e^{-\beta t} + \frac{\alpha}{(\delta + \beta)} e^{\delta t} \tag{7.6}$$

To calculate the estimate value

$$\hat{X}^{(1)} = (2.874, 6.5027, 10.1328, 13.8907, 17.8993) \tag{7.7}$$

From $\hat{X}^{(1)}(i) = \sum_{i=1}^{k} \hat{X}^{(0)}(i)$, we could obtain

$$\hat{X}^{(0)} = (2.874, 3.6287, 3.6301, 3.7579, 4.0086) \tag{7.8}$$

$$\Delta_k = \frac{\left|x^{(0)}(i) - \hat{x}^{(0)}(i)\right|}{x^{(0)}(i)} = (0.1070, 0.0878, 0.1085, 0.0896) \tag{7.9}$$

$$\Delta = \frac{1}{4} \sum_{i=1}^{4} \Delta_4 = 0.0982 = 9.8\% \tag{7.10}$$

The relative average error is 9.8%.

The accuracy of this model is

$$1 - 9.8\% = 100\% - 9.8\% = 90.2\% \tag{7.11}$$

7.7 Methods of building statistical-grey consistency with new flowchart method of GM (1, 1) differential equation modelling with sine and cosine term on right side program

Based on Statistical-Grey Consistent differential equation model with sine and cosine term on right side procedure and combined with new flowchart idea, utilize Matlab program to build Matlab toolbox.

Program structure:

The structure is as same as the Statistical-Grey Consistent differential equation model with sine and cosine term on right side procedure, and with new flowchart method instead of traditional half-weighted operation.

Example 7.7.1: Let discrete positive data sequence $X^{(0)}$=[1.6000 0.0976 0.7552 0.3181 1.0327 0.6674 1.4724 1.2207 2.1688 2.0973 3.2721 3.4858 5.0197 5.6853 7.7880 9.1695 12.1732], then we use new differential equation (6.49) to calculate the estimated value of $X^{(0)}$ sequence.

Step1 Rewrite the equation (6.58) as (7.12), then made an M-file named "nlinearfit.m" to express the modified equation (7.12). Then we use Matlab genetic algorithm toolbox to simulate parameter, α, β, ω (denoted as p (1), p (2), and p(3), respectively) to minimize the object function:

$$\underset{\min}{objectfun} = x^{(0)}(k) - \left(\alpha \sin(\omega k) + \beta\left(-z^{(1)}(k)\right)\right) \tag{7.12}$$

```
function nlinearfit=nlinearfit(p)
x= [1.6000, 0.0976, 0.7552, 0.3181, 1.0327, 0.6674, 1.4724, 1.2207,
2.1688, 2.0973, 3.2721, 3.4858, 5.0197, 5.6853, 7.7880, 9.1695, 12.1732];
xx=[0.0976 ,0.7552 ,0.3181 ,1.0327 ,0.6674 ,1.4724 ,1.2207 ,2.1688,2.0973 ,3.2721,3.4858
,5.0197 ,5.6853 ,7.7880 ,9.1695 ,12.1732];
x1=cumsum(x);
for i=2:length(x1)
z(i-1)=0.5*(x1(i-1)+x1(i));
end
```

```
k=1:length(xx);
f=0;
for t=1:length(xx)
n(t)=abs(xx(t)-(p(1)*sin(p(3)*k(t))-p(2)*(-z(t))));
f=f+n(t);
end
nlinearfit=f;
```

Step2. Using Matlab genetic algorithm toolbox to estimate the parameter α, β, ω.

Then we can easily estimate the parameter

$$\alpha = \ 1.01331270; \beta = -0.2293; \omega = -2.99239$$

And the minimized object function value equal to 1.0133127

$$objectfun_{min} = 1.0133127$$

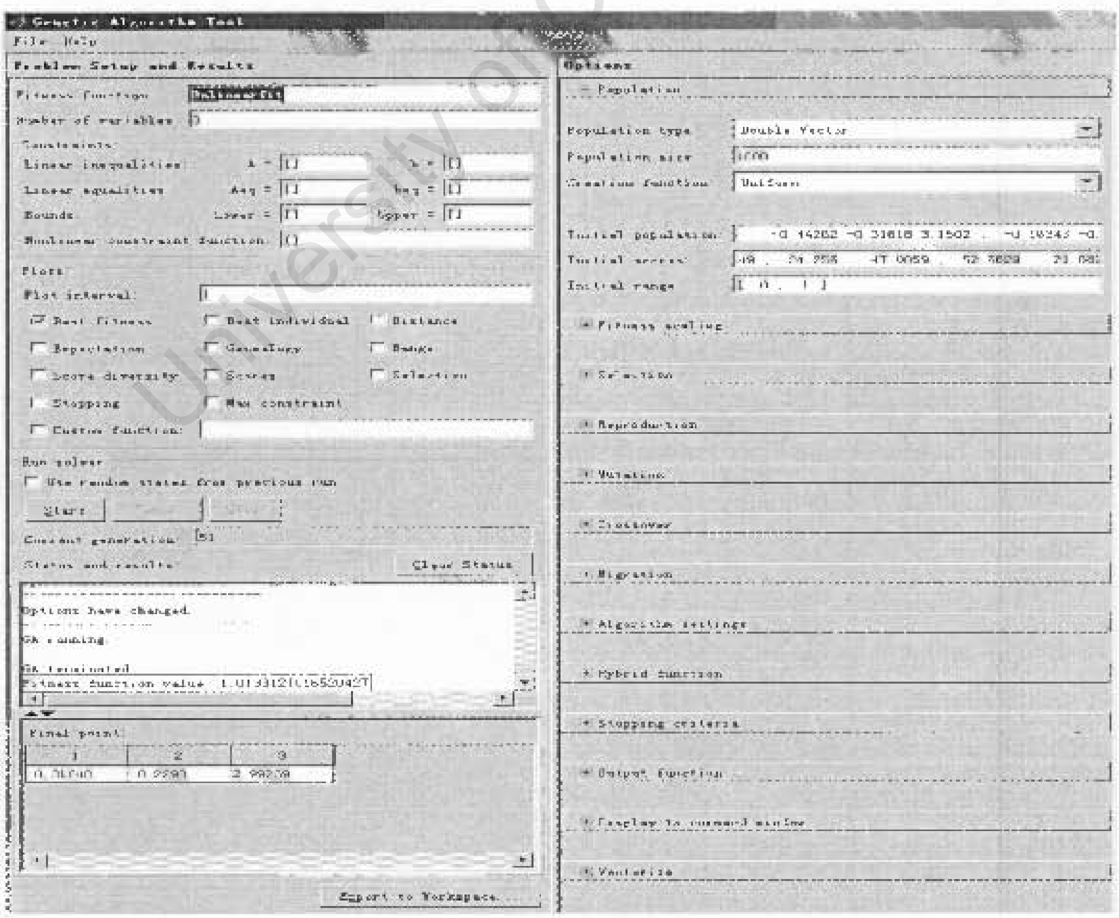

Figure 7.7.1 Utilize Genetic Algorithm tool to estimate Parameter α, β, δ

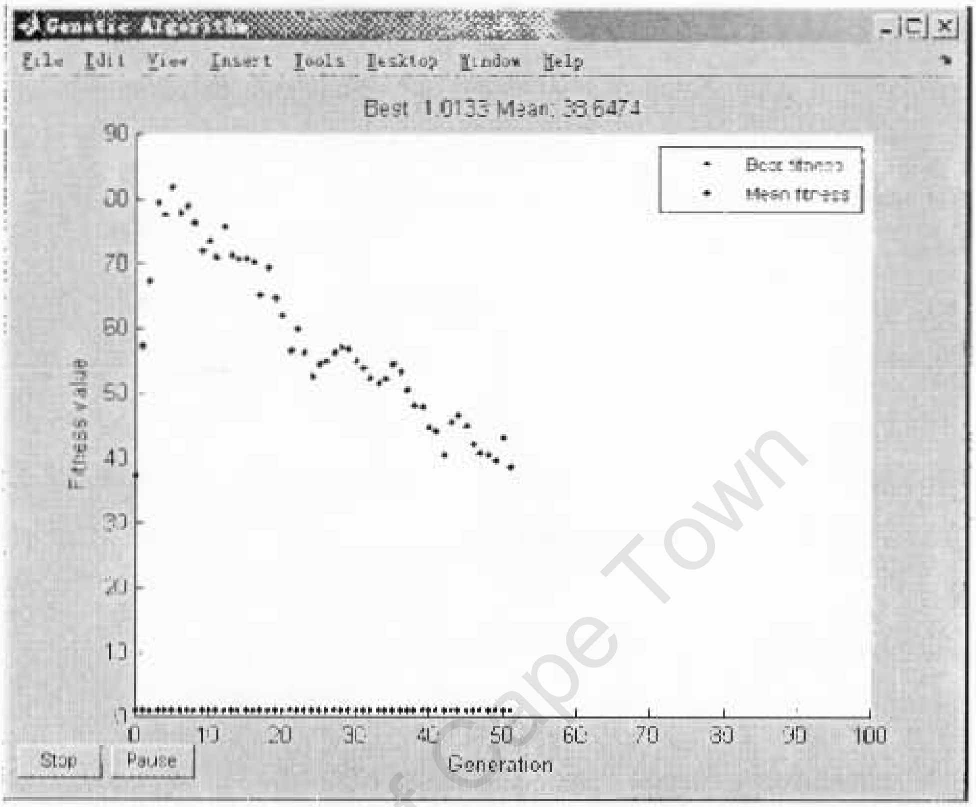

Figure 7.7.2 Genetic Algorithm running procedure

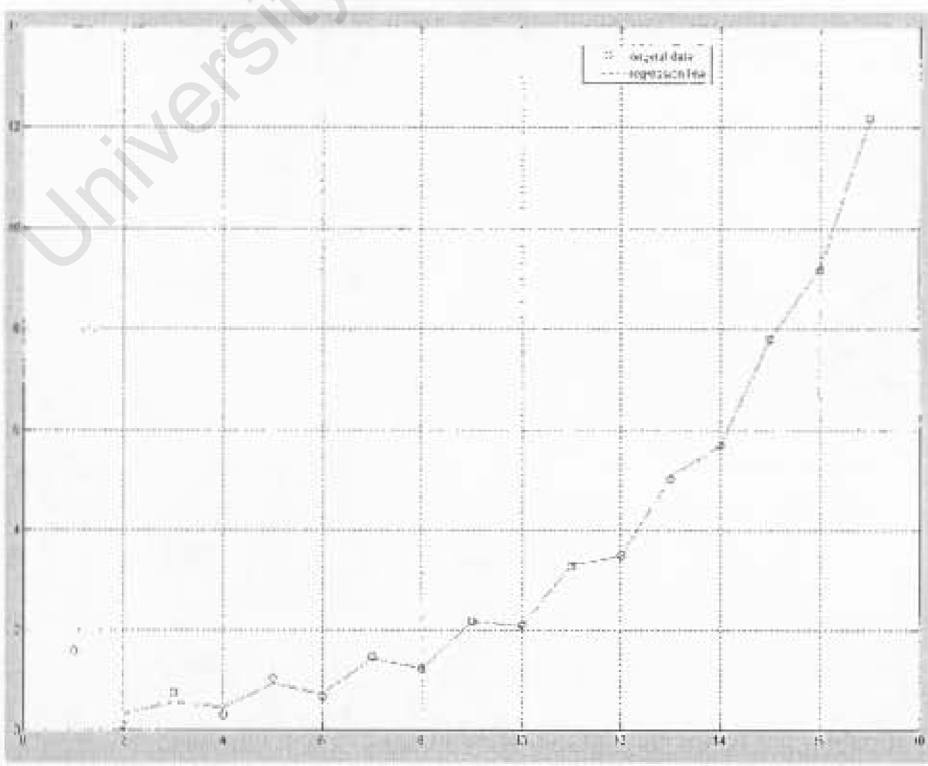

Figure 7.7.3 Coupling regression line

Step3. Put the first initial value 1.6 into general solution to calculate constant value C_0

$$
\begin{aligned}
1.6 &= C_0 - \frac{\alpha\omega}{\omega^2+\beta^2} \\
C_0 &= 1.6 - \frac{\alpha\omega}{\omega^2+\beta^2} = 1.4968
\end{aligned}
\tag{7.13}
$$

Step4. Using general solution

$$x = c_0 e^{-\beta t} + \frac{\alpha\beta}{\omega^2+\beta^2}\sin\omega t - \frac{\alpha\omega}{\omega^2+\beta^2}\cos\omega t \tag{7.14}$$

To calculate the estimate value

$$
\begin{aligned}
\hat{X}^{(1)} = (&1.6000 \quad 1.7818 \quad 2.4641 \quad 2.8885 \quad 3.8265 \quad 4.6405 \quad 5.9833 \quad 7.4070 \quad 9.4032 \\
&11.7725 \quad 14.8264 \quad 18.6624 \quad 23.4229 \quad 29.5420 \quad 37.0417 \quad 46.7311 \quad 58.6062)
\end{aligned}
\tag{7.15}
$$

From $\hat{X}^{(1)}(t) = \sum_{i=1}^{t} \hat{X}^{(0)}(i)$, $X^{(1)}(t) = \sum_{i=1}^{t} X^{(0)}(i)$, $\frac{\hat{X}^{(0)}}{\hat{X}^{(1)}} = \frac{X^{(0)}}{X^{(1)}} = constant$ we could obtain

$constant = \frac{X^{(1)}}{X^{(0)}} = (1.0000$ 17.3934 3.2479 8.7108 3.6832 6.6991 4.0365 5.8688 4.3033 5.4500 4.4932 5.2178 4.6233 5.0821 4.7100 5.0003 4.7665

$\hat{X}^{(0)} = (1.6$ 0.1818 0.6823 0.4244 0.938 0.814 1.3428 1.4237 1.9962 2.3693 3.0539 3.836 4.7675 6.1121 7.4997 9.6894 11.8751)

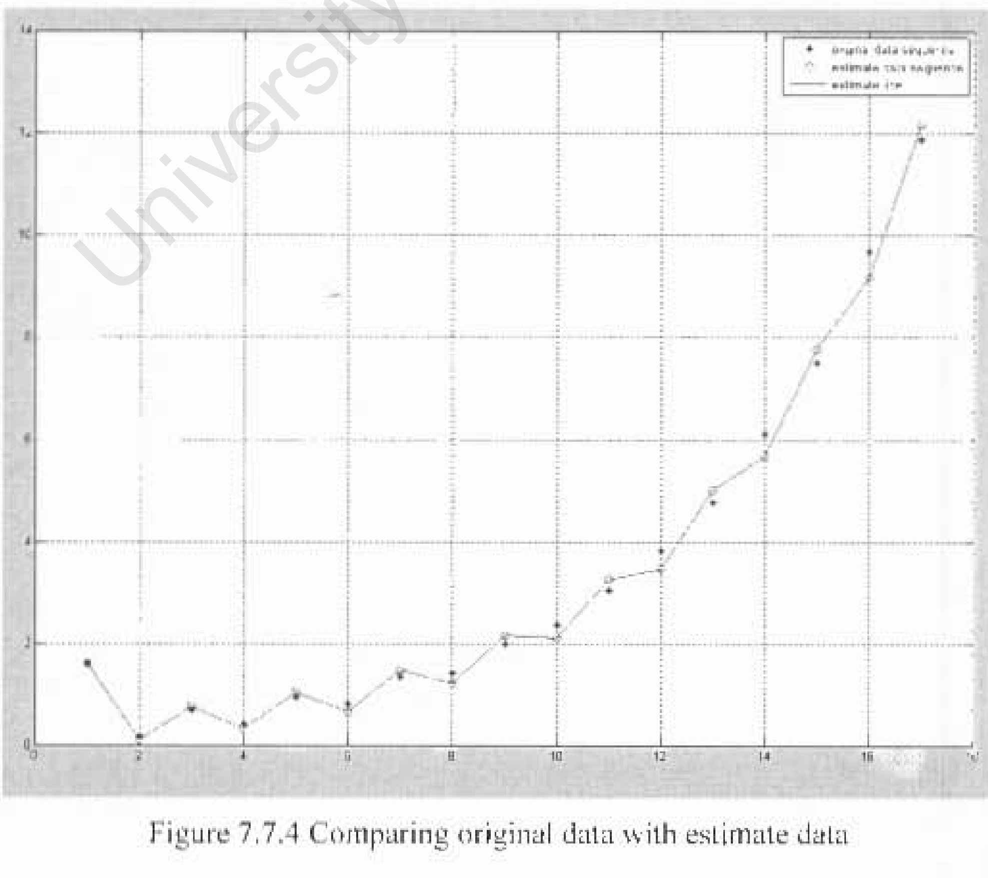

Figure 7.7.4 Comparing original data with estimate data

$$\Delta_k = \frac{\left|x^{(0)}(i) - \hat{x}^{(0)}(i)\right|}{x^{(0)}(i)} = (0 \quad 0.0496 \quad 0.0046 \quad 0.0424 \quad 0.0060 \quad 0.0379 \quad 0.0067 \quad 0.0339$$
$$0.0075 \quad 0.0299 \quad 0.0084 \quad 0.0261 \quad 0.0093 \quad 0.0225 \quad 0.0098 \quad 0.0192 \quad 0.0100)$$

$$\Delta = \frac{1}{16}\sum_{i=1} \Delta_{16} = 0.0202 = 2.02\%$$

The relative average error is 2.02%.

The accuracy of this model is

$1 - 2.02\% = 100\% - 2.02\% = 97.98\%$

7.8 Methods of building statistical-grey consistency with new flowchart method of GM (1, 1) differential equation modelling with polynomial term on right side program

Example 7.8.1: Let discrete positive data sequence $X^{(0)}$ = [50.9779 42.4028 45.8524 61.3186 88.796 128.2808 179.7707 243.2639 318.7593 406.2562 505.7542 617.2528 740.7519 876.2513 1023.7508 1183.2506], then we use new differential equation (6.74) to calculate the estimated value of $X^{(0)}$ sequence.

Step1 Rewrite the equation (6.74) as (7.16), then made an M-file named "nlinpoly.m" to express the modified equation (7.16). Then we use Matlab genetic algorithm toolbox to simulate parameter, $\alpha_0, \alpha_1, \alpha_2, \beta$ (denoted as p (1), p (2), p (3) and p(4) respectively) to minimize the object function:

$$\underset{\min}{objectfun} = x^{(0)}(k) - \left(\alpha_0 + a_1 k + a_2 k^2 + \beta\left(-z^{(1)}(k)\right)\right) \tag{7.16}$$

```
function nlinpoly=nlinpoly(p)
x=[50.9779,42.4028,45.8524,61.3186,88.796,128.2808,179.7707,243.2639,318.7593,406.2562,505.7542,617.2528,740.7519,876.2513,1023.7508,1183.2506];
xx=[42.4028,45.8524,61.3186,88.796,128.2808,179.7707,243.2639,318.7593,406.2562,505.7542,617.2528,740.7519,876.2513,1023.7508,1183.2506];
x1=cumsum(x);
for i=2:length(x1)
z(i-1)=p(5) *x1(i-1)+(1-p(5))*x1(i);
end
k=1:length(xx);
f=0;
for t=1:length(xx)
n(t)=abs(xx(t)-(p(1)+p(2)*k(t)+p(3)*(k(t)^2)+p(4)*(-z(t))));
f=f+n(t);
end
```

nlinpoly=f;

Step2. Using Matlab genetic algorithm toolbox to estimate the parameter $\alpha_0, \alpha_1, \alpha_2, \beta$.

Then we can easily estimate the parameter

$$\alpha_0 = 50.7035; \alpha_1 = -14.4504; \alpha_2 = 5.9917; \beta = -0.0309 \tag{7.17}$$

And the minimized object function value equal to 1.0133127

$$objectfun_{min} = 0.46073 \tag{7.18}$$

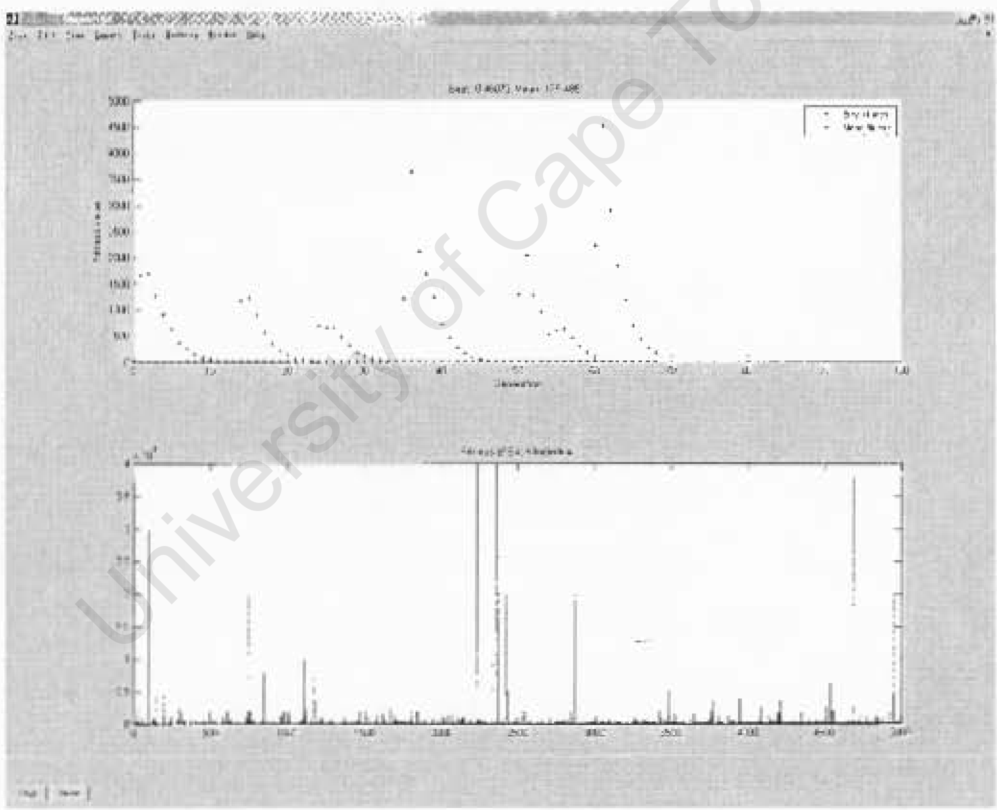

Figure 7.8.1 Genetic Algorithm running procedure

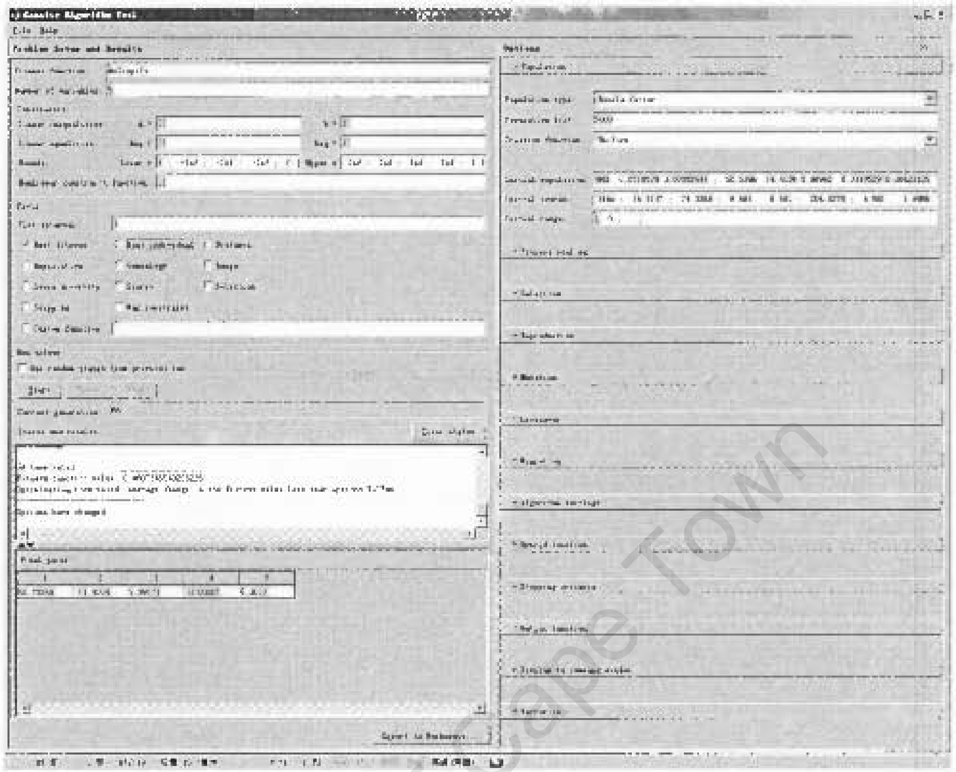

Figure 7.8.2 Utilize Genetic Algorithm tool to estimate Parameter $\alpha_0, \alpha_1, \alpha_2, \beta$

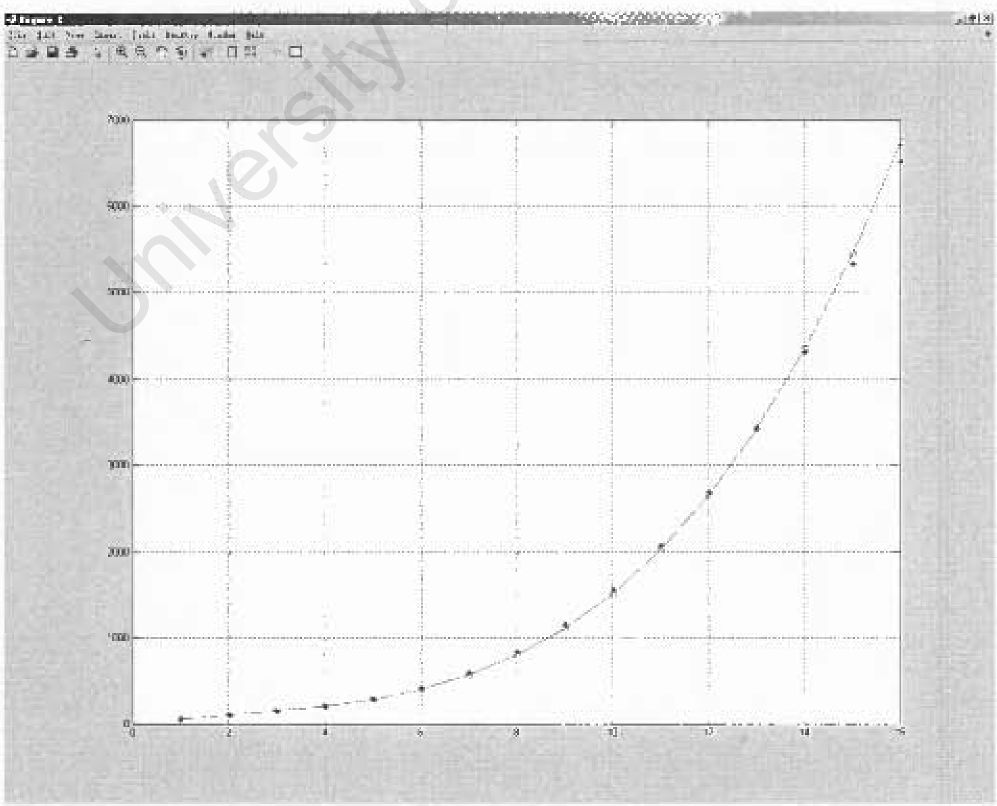

Figure 7.8.3 Coupling regression line

Step3. Put the first initial value 50.9779 into general solution to calculate constant value C_0

$$
\begin{aligned}
&50.9779 = C_0 - \frac{a_0\beta^2 - a_1\beta + 2a_2}{\beta^3} \\
&C_0 = 50.9779 - \frac{a_0\beta^2 - a_1\beta + 2a_2}{\beta^3} = 3.9369e+005
\end{aligned}
\tag{7.19}
$$

Step4. Using general solution

$$
x = c_0 e^{-\beta t} + \frac{a_0\beta^2 - a_1\beta + 2a_2}{\beta^3} + \frac{a_1\beta - 2a_2}{\beta^2} t + \frac{a_2}{\beta} t^2 \tag{7.20}
$$

To calculate the estimate value

$$
\begin{aligned}
\hat{X}^{(1)} = &(50.9779, 98.7833, 145.551, 203.4187, 284.9048, 402.9203, 570.7808, \\
&802.2197, 1111.4009, 152.9324, 2021.8805, 2653.7835, 3424.6673, 4351.0601, \\
&5450.0085, 6739.0936)
\end{aligned}
\tag{7.21}
$$

From $\hat{X}^{(1)}(t) = \sum_{i=1}^{k} \hat{X}^{(0)}(i) \cdot X^{(1)}(t) = \sum_{i=1}^{k} X^{(0)}(i)\ \frac{\hat{X}^{(0)}}{\hat{X}^{(1)}} = \frac{X^{(0)}}{X^{(1)}} = constant$ we could obtain

$$
\begin{aligned}
constant = \frac{X^{(1)}}{X^{(0)}} = (&1.0000\quad 2.2022\quad 3.0365\quad 3.2707\quad 3.2586\quad 3.2556\quad 3.3231\quad 3.4558 \\
&3.6373\quad 3.8539\quad 4.0957\quad 4.3559\quad 4.6297\quad 4.9138\quad 5.2058\quad 5.5041)
\end{aligned}
\tag{7.22}
$$

$$
\begin{aligned}
\hat{X}^{(0)} = (&50.9779\quad 44.856\quad 47.933\quad 62.1952\quad 87.4325\quad 123.763\quad 171.7606\quad 232.1395 \\
&305.5568\quad 392.5698\quad 493.6557\quad 609.2402\quad 739.7215 \\
&885.4848\quad 1046.9113\quad 1224.3844)
\end{aligned}
\tag{7.23}
$$

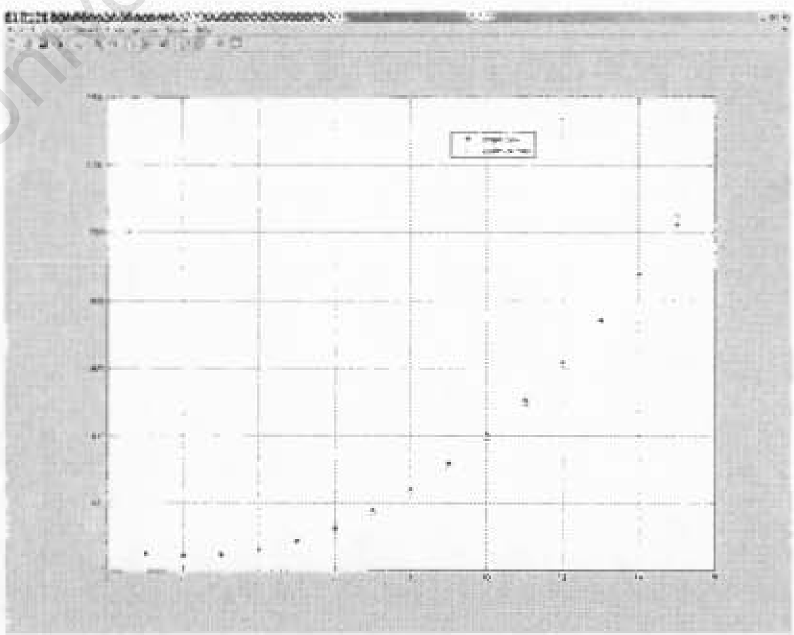

Figure 7.8.4 original data value compare with estimate one's

$$\Delta_k = \frac{\left|x^{(0)}(i) - \hat{x}^{(0)}(i)\right|}{x^{(0)}(i)} = (0\ 0.0579\ 0.0454\ 0.0143\ 0.0154\ 0.0352\ 0.0446\ 0.0457 \quad (7.24)$$
$$0.0414\ \ 0.0337\ 0.0239\ 0.0130\ 0.0014\ 0.0105\ 0.0226\ 0.0348)$$

$$\Delta = \frac{1}{16}\sum_{i=1}^{16}\Delta_{16} = 0.0275 = 2.75\% \quad (7.25)$$

The relative average error is 2.75%.

The accuracy of this model is

$1 - 2.75\% = 100\% - 2.75\% = 97.25\%$

References

Ashby, W.R. (1964). Methuen, London: An Introduction to Cybernetics.

Beckett, A. (2003) 8 September. "How an inventor from West By fleet revolutionized Chile – Santiago dreaming", The Guardian, Vol. G2 pp.1-3.

Bertalanffy, K.L.V., (1951), General System Theory - A new approach to unity of science (Symposium), Human Biology, Dec 1951, Vol. 23, p. 303-361

Buckland, M. K., (1991), Information and Information systems, New York, Greenwood Press, ISBN: 0313274630 978

Caudill, M.,(1989) Neural Networks Primer, San Francisco, CA: Miller Freeman Press

Chen, S.Y. (1998). "Engineering Fuzzy Set Theory and Application", in Engineering Theory Trends. Editors: B. Liu, R. Zhao, and G. Wang. Beijing: Defence Industry Publishing House.

Cui, Y.H. and Guo, R. (2006a).Spline Function and First-Order One-Variable Grey Differential Equation Model. Proceedings of 11th International Conference on Industrial Engineering, Theory, Applications and Practice, Nagoya, Japan, pp. 1308-1313.

Cui, Y. H., Guo, R. and Dunne, T. (2006b). Spline Functions in Grey Relation Analysis and Applications in System Reliability Improvements. Proceedings of the 1st International Conference on System Integration and Reliability Improvements (SIRI 2006), Hanoi, Vietnam.

Deng, J. L. (1982), "Control problems of grey system," Systems and Control Letters, vol. 5, pp. 288-294.

Deng, J. L. (2002). The Foundation of Grey Theory. Wuhan, China: Huazhong University of Technology Press.

Guo, R. (2005) "Repairable System Modelling Via Grey Differential Equations", Journal of Grey System, Vol. 8, No.1, pp 69-92.

Guo, R. (2007a). "Grey Differential Equation GM (1, 1) Model in Reliability Engineering," Computational Intelligence in Reliability Engineering – Vol. 2. New Metaheuristics, Neural and Fuzzy Techniques in Reliability, pp 387-413, Springer, Edited by G.Levitin.

Guo, R. (2007b). "A Univariate DEAR Modelling on repair Effects". Proceedings of Summer Safety and Reliability Seminars, Gdańsk-Sopot, Poland, pp 157-166, Editors: Krzysztof Kolowrocki and Enrico Zio. ISBN Volume 1: 978-83-925436-0-2.

Guo, R. (2007c) "Modelling Imperfectly Repaired System Data via Grey Differential Equations with Unequal-gapped Times," Reliability Engineering and Systems Safety, Vol. 92, No. 3, pp.378-391.

Guo, R. (2007d) "Differential Equation Motivated Regression Modelling on Repair Effects". Accepted by International Journal of Materials and Structural Reliability.

Guo, R. and Cui, Y.H. (2006a). Grey Reliability Analysis under L1 Norm. Proceedings of the Asian International Workshop on Advanced Reliability Modelling II: Reliability Testing and Improvement, Pusan, Korea, pp. 645-653.

Guo, R. and Cui, Y. H. (2006b). Foundations for GM (1, 1) Modelling of System Lifetimes. Proceedings of the International Conference on Maintenance Engineering (ICME), Cheng Du, China, pp 572-579.

Guo, R., Cheng, C.Y. and Cui, Y.H. (2006). L1 Normed GM (1, 1) Models and Reliability Analysis. Proceedings of 2006 IEEE International Conference on Systems, Man and Cybernetics, Taipei, Taiwan.

Guo, R., Cui, Y.H. and Dunne, T. (2006). Dynamic Grey Quality Control Chart with Sparse Data. Proceedings of the 1st International Conference on System Integration and Reliability Improvements (SIRI 2006), Hanoi, Vietnam.

Guo, R. and Dunne, T. (2006) "Grey Predictive Control Charts," Communications in Statistics–Theory and Methods, Vol. 35, No. 10, pp.1857-1868.

Guo, R., Dunne T. and Cui, Y. H. (2005).The Investigation on the Rationale of GM (1, 1) Model. Proceedings of the 10th International Conference on Industrial Engineering – Theory, Applications and Practice, Clearwater, Florida, U.S.A., pp. 438-443.

Guo, R. and Guo, D. (2007). "Multivariate DEMR Model – The Formation and Foundation". Proceedings of IEEE International Conference on Grey Systems and Intelligent Services, Nanjing, China, Editors: Sifeng Liu.

Guo, D., Guo, R. and Cui, Y.H. (2006). Rationale of GM (1, 1) Modelling: A Variational Approach. Proceedings of the International Conference on Maintenance Engineering (ICME), Cheng Du, China, pp 603-609.

Guo, R., Guo, D. and Cui, Y.H. (2006).Rationale Exploration of First-Order One-Variable Grey Differential Equation Model via Variational Approach. Proceedings of 11th International Conference on Industrial Engineering, Theory, Applications and Practice, Nagoya, Japan, pp. 1302-1307.

Guo, D., Guo, R. and Thiart, C. (2006) "The Coupling of Regression Modelling and Differential Equation Model in GM (1, 1) Modelling and Extended GM (1, 1) Models". Journal of Grey System, Vol. 9, No. 2, December, pp 143-154.

Guo, D., Guo, R., Thiart, C. and Cui, Y.H. (2005).Optimal Data Transformations in Grey Analysis. Proceedings of the 10th International Conference on Industrial Engineering – Theory, Applications and Practice, Clearwater, Florida, U.S.A., pp. 426-431.

Guo, R. and Love, C. E.(2005a) "Grey Repairable System Analysis" (Plenary Lecture), In Advances in Safety and Reliability, Proc. European Safety and Reliability Conference, Tri City, Poland, pp.753-766.

Guo, R. and Love, C. E. (2005b). "Fuzzy Set-valued and Grey Filtering Statistical Inferences on A System Operation Data," International Journal of Quality in Maintenance Engineering–advanced Reliability Modelling, Vol. 11, No. 3, pp.267-278.

Guo, R. and Love, C. E. (2006) "Grey Repairable System Analysis". International Journal of Automation and Computing, Vol. 3. No. 2, pp.131-144.

Huang, N., Shao, H. H. and Zhang Z. J. (1995). Systems & Control Letters, Volume 25 Issue 1, Publisher: Elsevier Science Publishers B. V.

Li, F. (1992). Grey System Theory and Applications, Technology and Science Literature Press, (in Chinese).

Li, G. D., Yamaguchi, D. and Nagai, M. (2005). "New methods and accuracy improvement of GM according to Laplace transform", Journal of Grey System., Vol. 8 No.1, pp.13-25.

Liu, S., and Lin, Y. (2006), Grey Information Theory and Practical Applications. London: Springer Press.

Liu, B. and Liu, Y. K. (2002) "Fuzzy Systems", IEEE Transactions on, ieeexplore.ieee.org.

Liu, B., Zhao, R., and Wang, G. (2003). Uncertain Programming with Applications. Beijing: Tsinghua University Press (in Chinese).

Raber, D. and Budd, J.M. (2003) "Information as sign: semiotics and information science", Journal of Documentation, Vol. 59 No. 5, pp. 507-22.

Shannon, C. E. (1948) ,A mathematical theory of coammunication. Bell System Tech. J. 27, 379-423, 623-656.

Wen, K. L. (2004), Grey Systems: Modelling and Prediction, USA: Yang's Scientific Research Institute Press.

Zadeh, L. A.(1965) ,"Fuzzy Sets and Systems", Information and Control, 8, 338-353.

Zhang, W. N., Du, Z. D., and Xue, B. (2007) Beijing, China: Ordinary Differential Equations, Higher Education Press.

Software:

MATLAB R2007a Copyright 1984-2007, The Mathworks, Inc. Protected by U.S. patents

Microsoft Development Environment 2003, Version 7.1.3088 Copyright 1987-2002 Microsoft Corporation. All rights reserved.

Microsoft.NET Framework 1.1 Version 1.1.4322 Copyright 1998-2002 Microsoft Corporation. All rights reserved

Appendix A: Visual C++ program for Chapter 2 (GM (1, 1) modelling procedure)

```
#include "stdafx.h"
#include <stdio.h>
#include <math.h>
#include <iostream>
#include <conio.h>
#include <ctype.h>

main(void)
{
using namespace std;
int i,j,l,g,h,m,n,t,y;
double
d[50],b[50],c[50],c1[50],cc,t1,e[50],f[50],b1[50],x[50],k,alfa,beta,C,D,E,F,m1,hat[50],hat1[50],er[
50],ler[50];
int ch;
do
{
cout << "Please enter how many numbers you wanna put in ?" <<endl;
cin >>k;
cout << "the numbers ratio must from "<<exp((-2)/(k+1))<<" to "<<exp(2/(k+1))<<endl;
cout << "Please enter the numbers" <<endl;
x[0]=0.0;
for(i=0;i<k;i++)
{
cin >>b[i];
x[i+1]=x[i]+b[i];
```

```
cout << "the No."<<i+1<<" AGO is "<<x[i+1]<<endl;
}
c[0]=0.0;
for(j=0;j<k-1;j++)
{
b1[j]=b[j]/b[j+1];
c1[j]=(x[j+1]+x[j+2])/2;
c[j+1]=c[j]+(x[j+1]+x[j+2])/2;
if(b1[j]<exp((-2)/(k+1))||b1[j]>exp(2/(k+1)))
{
cout<< "the No."<<j+1<<" sequence ratio can't be accepted"<<endl;
}
if(j>k-3)
{
C=c[j+1];
cout << " C=" << C <<endl;
}
}
d[0]=0.0;
for(l=0;l<k-1;l++)
{
d[l+1]=d[l]+b[l+1];
if(l>k-3)
{
D=d[l+1];
cout << " D=" <<D<<endl;
}
}
e[0]=0.0;
for(g=0;g<k-1;g++)
{
e[g+1]=e[g]+(c1[g]*b[g+1]);
```

```
if(g>k-3)
{
E=e[g+1];
cout << " E=" <<E<<endl;
}
}
f[0]=0.0;
for(h=0;h<k-1;h++)
{
f[h+1]=f[h]+pow(c1[h],2.0);
if(h>k-3)
{
F=f[h+1];
cout << " F=" <<F<<endl;
}
}
alfa=(C*D-(k-1.0)*E)/((k-1.0)*F-C*C);
cout << " alfa=" << alfa <<endl;
beta=(D*F-C*E)/((k-1.0)*F-C*C);
cout << " beta=" << beta <<endl;
m1=0.0;
for(m=0;m<k;m++)
{
hat1[m]=(b[0]-(beta/alfa))*exp((-alfa)*(m1))+(beta/alfa);
m1=m1+1.0;
cout<<"X(1)["<<m+1<<"]="<<hat1[m]<<endl;
}
for(n=0;n<k-1;n++)
{
hat[n]=hat1[n+1]-hat1[n];
cout<<"X(0)["<<n+2<<"]="<<hat[n]<<endl;
}
```

```
for(y=0;y<k-1;y++)//error from 3-k
{ler[y]=b[y+1]-hat[y];
cout<<"error"<<y+2<<" is "<<ler[y]<<endl;}
cc=0.0;
t1=0.0;
for(t=0;t<k-1;t++)//error from 3-k
{er[t]=abs((b[t+1]-hat[t])/b[t+1]);
cc=er[t]+cc;
t1=t1+1;
cout<<"relative error"<<t+2<<" is "<<er[t]*100<<"%"<<endl;}
cout<<"Average relative error is"<<(cc/t1)*100<<"%"<<endl;

      _cputs( "Type 'Y' when finished typing keys: \n" );
      ch = _getch();
      ch = toupper( ch );
   } while( ch != 'Y' );

   _putch( ch );
   _putch( '\r' );    /* Carriage return */
   _putch( '\n' );    /* Line feed       */
return 0;
}
```

Appendix B: Matlab program for chapter2 (GM(1,1) modelling procedure)

```
function [result]=grey(y)
ay(1)=y(1);
x=1:length(y);
for t=2:length(y)
   ay(t)=y(t)+ay(t-1);
end
z(1)=ay(1);
for i=2:length(y)
   z(i)=0.5*(ay(i)+ay(i-1));
end
c=sum(z(2:length(y)));
d=sum(y(2:length(y)));
for i1=2:length(y)
   et(i1-1)=z(i1)*y(i1);
end
e=sum(et(:));
for i2=1:length(y)
   ft(i2)=z(i2)^2;
end
f=sum(ft(2:length(y)));
a=(c*d-(length(y)-1)*e)/((length(y)-1)*f-c^2);
b=(d*f-c*e)/((length(y)-1)*f-c^2);
ex(1)=y(1);
for i3=2:length(y)
   ex(i3)=(y(1)-b/a)*exp(-a*(i3-1))+b/a;
end
r(1)=ex(1);
```

```
for i4=2:length(y)
   r(i4)=ex(i4)-ex(i4-1);
end

result=r;
for i5=1:length(y)
   error(i5)=abs((r(i5)-y(i5))/y(i5));
end
averror=sum(error(:))/(length(y)-1);
confidence=1-averror;
plot(x,y,'o',x,r,'*');
grid on;
```

Appendix C: Visual C++ program for chapter 2 (GM(1,1) for grey prediction)

Visual C++ program for Chapter 2

(GM (1, 1) for grey prediction)

```
#include "stdafx.h"
#include <stdio.h>
#include <math.h>
#include <iostream>
#include <conio.h>
#include <ctype.h>

main(void)
{
using namespace std;
int i,j,l,g,h,m,n,t,y;
double
d[50],b[50],c[50],c1[50],cc,t1,e[50],f[50],b1[50],x[50],k,alfa,beta,C,D,E,F,m1,hat[50],hat1[50],er[50],ler[50];
int ch,k1,n1;
do
{
cout << "Please enter how many numbers you wanna put in ?" <<endl;
cin >>k;
cout << "the numbers ratio must from "<<exp((-2)/(k+1))<<" to "<<exp(2/(k+1))<<endl;
cout << "Please enter the numbers" <<endl;
x[0]=0.0;
for(i=0;i<k;i++)
{
cin >>b[i];
```

```
x[i+1]=x[i]+b[i];
cout << "the No."<<i+1<<" AGO is "<<x[i+1]<<endl;
}
c[0]=0.0;
for(j=0;j<k-1;j++)
{
b1[j]=b[j]/b[j+1];
c1[j]=(x[j+1]+x[j+2])/2;
c[j+1]=c[j]+(x[j+1]+x[j+2])/2;
if(b1[j]<exp((-2)/(k+1))||b1[j]>exp(2/(k+1)))
{
cout<< "the No."<<j+1<<" sequence ratio can't be accepted"<<endl;
}
if(j>k-3)
{
C=c[j+1];
cout << " C=" << C <<endl;
}
}
d[0]=0.0;
for(l=0;l<k-1;l++)
{
d[l+1]=d[l]+b[l+1];
if(l>k-3)
{
D=d[l+1];
cout << " D=" <<D<<endl;
}
}
e[0]=0.0;
for(g=0;g<k-1;g++)
{
```

```
e[g+1]=e[g]+(c1[g]*b[g+1]);
if(g>k-3)
{
E=e[g+1];
cout << " E=" <<E<<endl;
}
}
f[0]=0.0;
for(h=0;h<k-1;h++)
{
f[h+1]=f[h]+pow(c1[h],2.0);
if(h>k-3)
{
F=f[h+1];
cout << " F=" <<F<<endl;
}
}
alfa=(C*D-(k-1.0)*E)/((k-1.0)*F-C*C);
cout << " alfa=" << alfa <<endl;
beta=(D*F-C*E)/((k-1.0)*F-C*C);
cout << " beta=" << beta <<endl;
m1=0.0;
cout << "Please enter how many numbers you wanna predict ?" <<endl;
cin >>k1;
for(m=0;m<k+k1;m++)
{
hat1[m]=(b[0]-(beta/alfa))*exp((-alfa)*(m1))+(beta/alfa);
m1=m1+1.0;
cout<<"X(1)["<<m+1<<"]="<<hat1[m]<<endl;
}
for(n=0;n<k+k1-1;n++)
{
```

```
hat[n]=hat1[n+1]-hat1[n];
cout<<"X(0)["<<n+2<<"]="<<hat[n]<<endl;
}
for(n1=k1;n1<2*k1;n1++)
{
cout<<"prediction is X(0)("<<n1+2<<")="<<hat[n1]<<endl;
}
for(y=0;y<k-1;y++)//error from 3-k
{ler[y]=b[y+1]-hat[y];
cout<<"error"<<y+2<<" is "<<ler[y]<<endl;}
cc=0.0;
t1=0.0;
for(t=0;t<k-1;t++)//error from 3-k
{er[t]=abs((b[t+1]-hat[t])/b[t+1]);
cc=er[t]+cc;
t1=t1+1;
cout<<"relative error"<<t+2<<" is "<<er[t]*100<<"%"<<endl;}
cout<<"Average relative error is"<<(cc/t1)*100<<"%"<<endl;
    _cputs( "Type 'Y' when finished typing keys: \n" );
    ch = _getch();
    ch = toupper( ch );
  } while( ch != 'Y' );

  _putch( ch );
  _putch( '\r' );   /* Carriage return */
  _putch( '\n' );   /* Line feed      */
return 0;
}
```

Appendix D: Matlab program for chapter4 (Checking statistical-grey consistency GM(1,1))

Matlab toolbox program for chapter 4

(Checking statistical-grey consistency GM (1, 1))


```
function [result ay ex a b z ratio rsqure1 yhat rsquare2 error1 error averror confidence]=greycon(y)
ay(1)=y(1);
x=1: length(y);
for t=2:length(y)
   ay(t)=y(t)+ay(t-1);
end
z(1)=ay(1);
for i=2:length(y)
   z(i)=0.5*(ay(i)+ay(i-1));
end
c=sum(z(2:length(y)));
d=sum(y(2:length(y)));
for i1=2:length(y)
   et(i1-1)=z(i1)*y(i1);
end
e=sum(et(:));
for i2=1:length(y)
   ft(i2)=z(i2)^2;
end
f=sum(ft(2:length(y)));
a=(c*d-(length(y)-1)*e)/((length(y)-1)*f-c^2);
b=(d*f-c*e)/((length(y)-1)*f-c^2);
yhat(1)=y(1);
```

```
for t1=2:length(y)
   yhat(t1)=b+(-a)*z(t1);
end
rsqure1=1-sum((y-yhat).^2)/sum((y-mean(y)).^2);
ex(1)=y(1);
for i3=2:length(y)
   ex(i3)=(y(1)-b/a)*exp(-a*(i3-1))+b/a;
end
rsquare2=1-sum((ay-ex).^2)/sum((ay-mean(ay)).^2);
r(1)=ex(1);
for i4=2:length(y)
   r(i4)=ex(i4)-ex(i4-1);
end

result=r;
for i5=1:length(y)
   error1(i5)=y(i5)-r(i5);
   error(i5)=abs((r(i5)-y(i5))/y(i5));
end
averror=sum(error(:))/(length(y)-1);
confidence=1-averror;
x=1:1:length(y);
plot(x,y,'o',x,result,'*',x,ay,'+',x,ex,'.');
grid on;
```

Appendix E: Matlab program for chapter4 (Checking statistical-grey consistency with new ratio idea GM(1,1) model)

Matlab toolbox program for chapter 4

(Checking statistical-grey consistency with new ratio idea GM (1, 1) model)

```
function [result ay ex a b z ratio rsqure1 yhat rsquare2 error1 error averror
confidence]=greyconratio(y)
ay(1)=y(1);
x=1: length(y);
for t=2:length(y)
   ay(t)=y(t)+ay(t-1);
end
for t1=1:length(y)
   ratio(t1)=ay(t1)/y(t1);
end
z(1)=ay(1);
for i=2:length(y)
   z(i)=0.5*(ay(i)+ay(i-1));
end
c=sum(z(2:length(y)));
d=sum(y(2:length(y)));
for i1=2:length(y)
   et(i1-1)=z(i1)*y(i1);
end
e=sum(et(:));
for i2=1:length(y)
   ft(i2)=z(i2)^2;
end
```

```
f=sum(ft(2:length(y)));
a=(c*d-(length(y)-1)*e)/((length(y)-1)*f-c^2);
b=(d*f-c*e)/((length(y)-1)*f-c^2);
yhat(1)=y(1);
for t1=2:length(y)
   yhat(t1)=b+(-a)*z(t1);
end
rsqure1=1-sum((y-yhat).^2)/sum((y-mean(y)).^2);
ex(1)=y(1);
for i3=2:length(y)
   ex(i3)=(y(1)-b/a)*exp(-a*(i3-1))+b/a;
end
rsquare2=1-sum((ay-ex).^2)/sum((ay-mean(ay)).^2);
for i4=1:length(y)
   r(i4)=ex(i4)/ratio(i4);
end
result=r;
for i5=1:length(y)
   error1(i5)=y(i5)-r(i5);
   error(i5)=abs((r(i5)-y(i5))/y(i5));
end
averror=sum(error(:))/(length(y)-1);
confidence=1-averror;
x=1:1:length(y);
plot(x,y,'o',x,result,'*',x,ay,'+',x,ex,'.');
grid on;
```

Appendix F: Matlab program for section 7.5 (Methods of building statistical-grey consistency with new flowchart method of GM(1,1) differential equation modelling with constant term on right side program)

```
function [xhat,x] =flowchart(y)
options = gaoptimset('PopulationSize', 100);
x = ga(@flowchart1, 3, [],[],[],[],[-inf -inf 0],[inf inf 1],[],options);
c0=y(1)-x(1)/x(2);
for i=1:length(y)
    x1hat(i)=c0*exp(-x(2)*(i-1))+x(1)/x(2);
end
xhat(1)=y(1);
for i2=2:length(y)
 xhat(i2)=x1hat(i2)-x1hat(i2-1);
end
accuracy=sum(abs((y-xhat)./y))/4;
result=accuracy;
end


function result=flowchart1(p)
x=[2.874 3.278 3.337 3.390 3.679];
x1=cumsum(x);
for i=2:length(x1)
   z(i-1)=p(3)*x1(i-1)+(1-p(3))*x1(i);
end
f=0;
for t=1:length(z)
    n(t)=abs(x(t+1)-(p(1)+p(2)*(-z(t))));
    f=f+n(t);
end
result=f;
end
```

Appendix G: Matlab toolbox program for chapter 7.6 (Methods of building statistical-grey consistency with new flowchart method of GM (1, 1) differential equation modelling with exponential term on right side program)

```
function [xhat,x]=flowchart(y)
options = gaoptimset('PopulationSize', 100);
x = ga(@nlin, 4, [],[],[],[],[-inf -inf  inf 0],[inf inf inf 1],[],options);
c0=y(1)-x(1)/x(2);
for i=1:length(y)
    x1hat(i)=c0*exp(-x(2)*(i-1))+x(1)/x(2);
end
xhat(1)=y(1);
for i2=2:length(y)
 xhat(i2)=x1hat(i2)-x1hat(i2-1);
end
accuracy=sum(abs((y-xhat)./y))/4;
result=accuracy;
end

function nlin=nlin(p)
x=[2.874;3.278;3.337;3.390;3.679];
xx=[3.278;3.337;3.390;3.679];
x1=cumsum(x);
for i=2:length(x1)
z(i-1)=p(4)*x1(i-1)+(1-p(4))x1(i);
end
k=[1;2;3;4];
nlin=xx-(p(1)*exp(p(3)*k)-p(2)*(-z'));
```

Appendix H: Matlab toolbox program for chapter 7.7 (Methods of building statistical-grey consistency with new flowchart method of GM (1, 1) differential equation modelling with sine and cosine term on right side program)

```
function [xhat,x] flowchart(y)
options   gaoptimset('PopulationSize', 100);
x = ga(@nlinearfit, 4, [],[],[],[],[-inf -inf -inf 0],[inf inf inf
1],[],options);
c0 y(1)-x(1)/x(2);
for i=1:length(y)
    x1hat(i)=c0*exp(-x(2)*(i-1))+x(1)/x(2);
end
xhat(1)=y(1);
for i2 2:length(y)
 xhat(i2)=x1hat(i2)-x1hat(i2-1);
end
accuracy=sum(abs((y-xhat)./y))/4;
result=accuracy;
end


function nlinearfit=nlinearfit(p)
x= [1.6000, 0.0976, 0.7532, 0.3181, 1.0327, 0.6674, 1.4724, 1.2207,
2.1688, 2.0973, 3.2721, 3.4858, 5.0197, 5.6853, 7.7880, 9.1695, 12.1732];
xx=[0.0976 ,0.7532 ,0.3181 ,1.0327 ,0.6674 ,1.4724 ,1.2207 ,2.1688,2.0973 ,3.27
21,3.4858 ,5.0197 ,5.6853 ,7.7880 ,9.1695 ,12.1732];
x1=cumsum(x);
for i=2:length(x1)
z(i-1) p(4)*(x1(i-1)+(1-p(4))x1(i));
end
k=1:length(xx);
f=0;
for t=1:length(xx)
n(t)=abs(xx(t)-(p(1)*sin(p(3)*k(t))+p(2)*(-z(t))));
f=f+n(t);
end
nlinearfit=f;
```

Appendix I: Matlab toolbox program for chapter 7.8 (Methods of building statistical-grey consistency with new flowchart method of GM (1, 1) differential equation modelling with polynomial term on right side program)

```
function [xhat,x]=flowchart(y)
options = gaoptimset('PopulationSize', 100);
x = ga(@nlinpoly, 5, [],[],[],[],[-inf -inf -inf -inf 0],[inf inf inf inf
1],[],options);
c0=y(1)-x(1)/x(2);
for i=1:length(y)
    x1hat(i)=c0*exp(-x(2)*(i-1))-x(1)/x(2);
end
xhat(1)=y(1);
for i2=2:length(y)
 xhat(i2)=x1hat(i2)-x1hat(i2-1);
end
accuracy=sum(abs((y-xhat)./y))/4;
result=accuracy;
end

function nlinpoly=nlinpoly(p)
x=[50.9779,42.4028,45.8524,61.3166,88.796,128.2806,179.7707,243.2639,318.7593,4
06.2562,505.7542,617.2528,740.7519,876.2513,1023.7508,1183.2506];
xx=[42.4028,45.8524,61.3166,88.796,128.2806,179.7707,243.2639,318.7593,406.2562
,505.7542,617.2528,740.7519,876.2513,1023.7508,1183.2506];
x1=cumsum(x);
for i=2:length(x1)
z(i-1)=p(5)*x1(i-1)+(1-p(5))*x1(i);
end
k=1:length(xx);
f=0;
for t=1:length(xx)
n(t)=abs(xx(t)-(p(1)+p(2)*k(t)+p(3)*(k(t)^2)+p(4)*(-z(t))));
f=f+n(t);
end
nlinpoly=f;
```

Appendix J: Matlab toolbox program for chapter 6 (GM (2, 1) program)

```
function [result] = gm21v(y)
options = gaoptimset('PopulationSize', 500);
x = ga(@gm21, 3, [],[],[],[],[-inf -inf -inf],[inf inf inf],[],options);
c1=(y(2)+y(1)-(y(1)-x(1)/x(3))*exp((-x(2)-sqrt(square(x(2))-4*x(3)))/2)-
x(1)/x(3))/(exp((-x(2)-sqrt(square(x(2))-4*x(3)))/2)-exp((-x(2)-
sqrt(square(x(2))-4*x(3)))/2));
c2=y(1)-c1-x(1)/x(3);
for i=1:length(y)
    xhat1(i)=c1*exp(((-x(2)+sqrt(square(x(2))-4*x(3)))/2)*(i-1))+c2*exp(((-
x(2)-sqrt(square(x(2))-4*x(3)))/2)*(i-1))-x(1)/x(3);
end
result=xhat1;
end

function result=gm21(p)
x=[2.874 3.278 3.337 3.390 3.679 3.907 4.132];
x1=cumsum(x);
for i=2:length(x1)
   z(i-1)=0.5*x1(i-1)+(1-0.5)*x1(i);
end
for i2=2:length(x)
    r(i2-1)=x(i2)-x(i2-1);
end
f=0;
for t=2:length(x)
    n(t-1)=abs(r(t-1)-(p(1)+p(2)*(-x(t))+p(3)*(-z(t-1))));
    f=f+n(t-1);
end
result=f;
end
```