東海大學統計研究所

碩士論文

指導教授：黃愉閔　博士

使用迴旋積分下概似函數估計時間序列迴歸模型

Estimating Time Series Regression Using Integrated Likelihood Function

研究生：林哲頡

中華民國一Ｏ二年七月

摘　要

本文針對時間序列線性迴歸，在我們假設的模型下估計出$\underset{\sim}{\beta}$參數值。透過兩種不同的估計方法進行分析，第一種是最小平方法，第二種是新方法，新方法是利用迴旋積分推導出Y的機率密度函數和 Likelihood Function，在獲得參數估計的過程中由於積分計算量太大所以我們利用了 Monte Carlo Methods 逼近，在本文第三章有詳細介紹新方法估計式的推導。但是在新方法的估計式中 Monte Carlo Method 所計算出來的 Likelihood 值在因變數Y及自變數X差異過大時其值會太小，所以我們進行資料上的篩選讓 Likelihood 值在現有給予的資料上盡可能靠近最佳值，進而使 Newton Raphson Method 估計出不錯的$\underset{\sim}{\beta}$參數值。最後，我們模擬資料應用於這兩種方法，比較他們的估計結果。

關鍵字：時間序列迴歸、蒙地卡羅、非高斯時間序列、概似函數

Abstract

The time series regression provides an explicit analysis, in which one time series (dependent variable) can be expressed linearly related to other time series variables (covariates), and often errors of the model are possibly correlated or simply white noises. The method of least squares is a naive approach to estimate the regression conditioned on the covariates. When the covariates are non-Gaussian stochastic time series, the least square estimators may not be quite efficient. We propose a new method taking into account the distribution properties. We estimate the parameters by maximizing the unconditional likelihood, which is obtained via convolution. The calculation of multi-fold convolution is insurmountable, so we approximate the unconditional likelihood using Monte Carlo, in which covariates are re-sampled and only selected probability weights are counted into the approximation. The maximum likelihood estimation is obtained applying the Newton-Raphson iterations on the approximated likelihood function. Simulation examples are given and the results are compared to the least squares estimates.

Keywords: Time series regression, Monte Carlo, Stochastic covariates, Non-Gaussian time series, Likelihood function.

謝　誌

美好的時光總是過得特別快，在這兩年的求學過程中，有歡笑有淚水，也受過許多人的幫助，使得哲頡能順利地完成學業。感謝系上師長們的細心指導與栽培、家人與同儕間不斷的鼓勵與幫忙，在研究所的學習路上才能如此順利，由衷地感謝你們。

論文能如期完成，最重要感謝的是指導教授黃愉閔老師，在學習的階段您扮演者亦師亦友的角色，細心教導我在學術上的提升，也關心我的日常生活，在此謹向恩師致上最深的敬意。

在論文撰寫過程中，還要感謝士嘉學長、孟樺助教、佳玲助教、雅文助教、淑姿助教、如君助教的建議與指導，給予我們許多協助。此外，仍要感謝與我們一起揮汗努力的同儕們，有了你們的陪伴與協助，使得我們在研究所的生活如此的多采多姿，謝謝你們，我會好好珍惜這一份難得的緣分。

林哲頡　敬上

中華民國 102 年 7 月

謹識於東海大學統計學系

目錄

第壹章　緒　論

第一節　研究背景與動機

線性迴歸是利用線性迴歸方程的最小平方函數，對一個或多個自變數和因變數之間關係進行建立模型的一種迴歸分析，這種函數是一個或多個稱為迴歸係數模型參數的線性組合。

模型$Y=\beta_1 X_1+\beta_2 X_2+\cdots+\beta_{p-1} X_{p-1}+\beta_p X_p+\varepsilon_t$，是一個線性迴歸模型，模型中的共變數不是一般常用的獨立資料，而是時間序列資料且並不特定依循常態，但在殘差項的部分假設為常態分配，在本研究中我以五個共變數為例：分別有 ARMA 時間序列兩列、EIAR(Exponential innovation)時間序列一列、EAR(Exponential stationary distribution)時間序列一列以及 BGAR (Beta-Gamma)時間序列一列。

ARMA(自迴歸移動平均模型)，是研究時間序列的重要方法，由 AR 模型（自迴歸模型）與 MA 模型（移動平均模型）為基礎「混合」構成。EIAR(ε_t是一個指數分配)、EAR(X_t是一個指數分配)，這兩種類型都是 AR 平穩的一個過程。BGAR 是由 Beta-Gamma 所組合而成的。

上述的組合無法得知Y的分配，但在有給定的條件下，可以假設Y的分配為常態，

$$Y \mid (\underset{\sim}{X_1}、\underset{\sim}{X_2}、\underset{\sim}{X_3}、\underset{\sim}{X_4}、\underset{\sim}{X_5}) \underset{\sim}{\overset{\text{iid}}{}} N(\beta_1 \underset{\sim}{X_1} + \beta_2 \underset{\sim}{X_2} + \beta_3 \underset{\sim}{X_3} + \beta_4 \underset{\sim}{X_4} + \beta_5 \underset{\sim}{X_5}, \sigma^2)$$

給予 $X_s^{'}$，$Y_s^{'}$ 為獨立，$\varepsilon_t \overset{iid}{\sim} N(0,\sigma_t^2)$，在這樣條件的情況下，我們可以估計出不錯的參數值。

但是在實務上可以發現，許多的時間序列資料並非符合理論上的情況，進而造成不是常態分配的結果，對於此種狀況傳統的最小平方法估計參數，並不是最適用的，所以我們找尋其他的估計方法來改善最小平方法的缺點。

為了改善這個缺點，我們提出了一個新的估計想法， 一開始模型是由五列時間序列(ARMA、EIAR、EAR 以及 BGAR)所組合而成的線性模型，模型中假設 $\varepsilon_t \overset{iid}{\sim} N(0,\sigma_t^2)$，然後在給定 $\underset{\sim}{X_1} \cdots \underset{\sim}{X_5}$ 的情況下，去推導 Y 的機率密度函數而獲得 likelihood function，在參數估計的過程中由於積分計算量太大所以我們利用了 Monte Carlo Methods 和 Newton-Raphson Method 來估計參數 $\underset{\sim}{\beta}$ 值，最後我們再經由電腦模擬將新的估計方法與最小平方法結果進行比較。

在比較的過程裡我們不只單單看估計結果上的數據，我們還重覆跑了多組結果，計算每一個參數的標準差，使結論更加確定。

第二節　研究目的

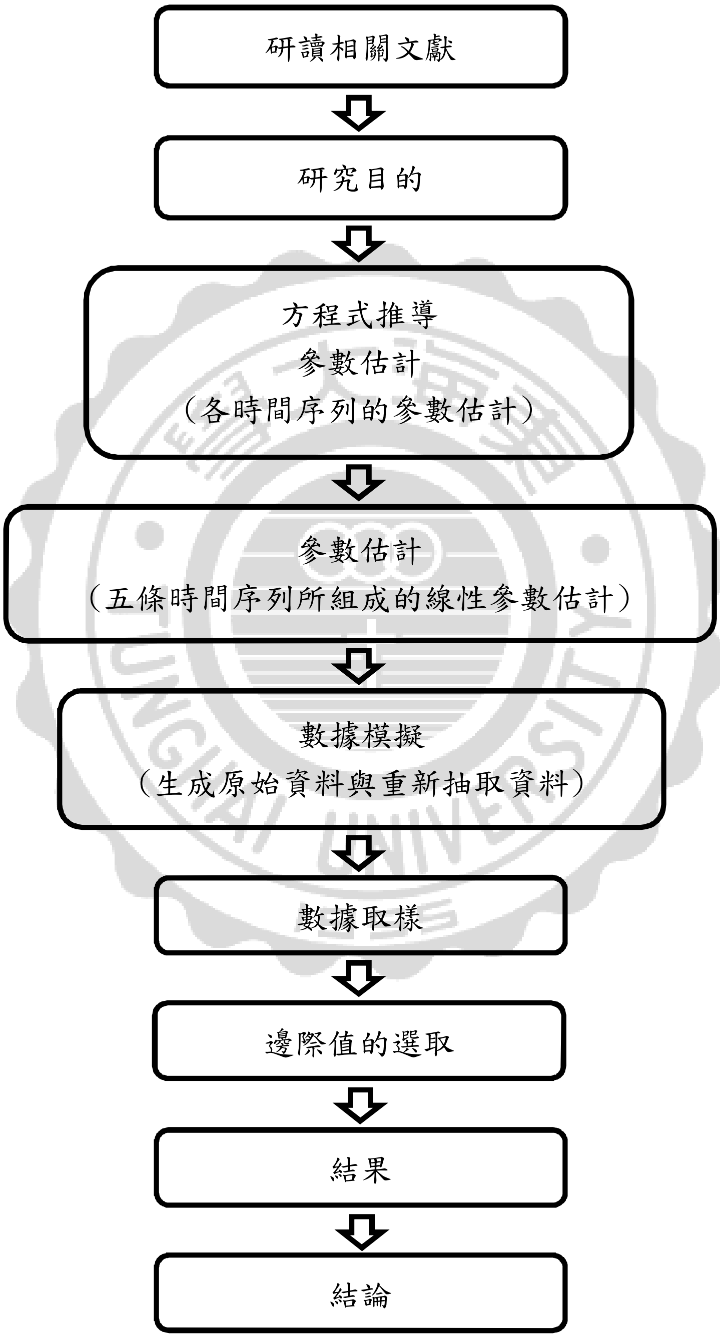
由於實務上所得到的資料，並非與理論上的情況相符合，所以我們提出一個新的研究方法，來與原本的最小平方方法相比較。

例如：線性模型中，變數不是一般常用的獨立資料，而是時間序列，所組成的線性迴歸模型，在估計模型係數上，我們使用了新的估計方法與最小平方法兩種方法，其結果做比較，最後還利用了變異數的大小來做更進一步的確認。

如果新方法比最小平方法估計結果好，也就表示說以後遇到這一種時間序列迴歸模型，可以使用這新方法估計，不用完全依靠最小平方法，並且加以推廣。

第三節　研究流程

經由研讀相關文獻，藉由研究目的，進而提出問題，再利用不同的方法研究，建構出研究架構：



第貳章　文獻回顧

Exponential Autoregressive and Beta-Gamma Autoregressive Model

第一節　EIAR、EAR

本文主要介紹 Georgiana Popovici(2010)文獻中所提出的具指數分配時間序列模型：

1.1　EIAR

$EIAR(1)$ 模型是 $X_t = \phi_1 X_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim \exp(\mu)$

對上述式子取期望值跟變異數：

$$E(X_t) = \phi_1 E(X_{t-1}) + E(\varepsilon_t) \Rightarrow \quad E(X_t) = \frac{\mu}{1-\phi_1}$$

$$Var(X_t) = \phi_1^2 Var(X_{t-1}) + Var(\varepsilon_t) \Rightarrow \quad Var(X_t) = \frac{\mu^2}{1-\phi_1^2}$$

特徵函數：

$$\varphi_X(t) = E(e^{itX_t}) = E(e^{it(\phi_1 X_{t-1}+\varepsilon_t)}) = E(e^{it\phi_1 X_{t-1}})E(e^{it\varepsilon_t})$$
$$\varphi_X(t) = \varphi_X(\phi_1 t)\varphi_{\varepsilon_t}(t) = \varphi_X(\phi_1 t)(1-it\mu)^{-1}$$

接下來我們利用 CLS(Conditional Least Squares)估計變數參數：

$EIAR(1)$ 條件估計式 $E(X_t | \mathrm{X}_{t-1} = x_{t-1}) = \phi_1 x_{t-1} + \mu$

$$Q_n(\psi) = \sum_{t=2}^{n}[X_t - (\phi_1 x_{t-1} + \mu)]^2 \quad \text{，其中} \psi = (\phi_1 \text{、} \mu)$$

$$\frac{\partial Q_n(\psi)}{\partial \mu}=0\text{得}\hat{\mu}=\frac{\sum_{t=2}^{n}X_t-\hat{\phi}_1\sum_{t=2}^{n}X_{t-1}}{n-1}$$

$$\frac{\partial Q_n(\psi)}{\partial \phi_1}=0\text{得}\hat{\phi}_1=\frac{\sum_{t=2}^{n}X_tX_{t-1}-\hat{\mu}\sum_{t=2}^{n}X_{t-1}}{\sum_{t=2}^{n}X_{t-1}^2}$$

將$\hat{\mu}$帶入式子 $\hat{\phi}_1=\dfrac{(\mathrm{n-1})\sum_{t=2}^{n}X_tX_{t-1}-\sum_{t=2}^{n}X_{t-1}\sum_{t=2}^{n}X_t}{(n-1)\sum_{t=2}^{n}X_{t-1}^2-(\sum_{t=2}^{n}X_{t-1})^2}$

接下來我們引用文獻結果，將過去的時間點推廣至兩個：

$EIAR(2)$ 模型是 $X_t=\phi_1X_{t-1}+\phi_2X_{t-2}+\varepsilon_t,\quad \varepsilon_t\sim\exp(\mu)$

對上述式子取期望值跟變異數：

$$E(X_t)=\phi_1E(X_{t-1})+\phi_2E(X_{t-2})+E(\varepsilon_t)=>\quad E(X_t)=\frac{\mu}{1-\phi_1-\phi_2}$$

$$r(0)=Var(X_t)=Cov(X_t、X_t)=\phi_1r(1)+\phi_2r(2)+\mu^2$$

$$r(1)=Cov(X_t、X_{t\text{-}1})=\phi_1r(0)+\phi_2r(1)$$

$$r(2)=Cov(X_t、X_{t\text{-}2})=\phi_1r(1)+\phi_2r(0)$$

$$r(0)=Var(X_t)=\frac{\mu^2}{1\text{-}\dfrac{\phi_1^2(1+\phi_2)+\phi_2^2(1\text{-}\phi_2)}{(1\text{-}\phi_2)}}$$

特徵函數：

$$\varphi_X(t)=E(e^{itX_t})=E(e^{it(\phi_1X_{t-1}+\phi_2X_{t-2}+\varepsilon_t)})=E(e^{it(\phi_1X_{t-1}+\phi_2X_{t-2})})E(e^{it\varepsilon_t})$$

$$\varphi_X(t)=E(e^{it(\phi_1X_{t-1}+\phi_2X_{t-2})})\varphi_{\varepsilon_t}(t)=E(e^{it(\phi_1X_{t-1}+\phi_2X_{t-2})})(1-it\mu)^{-1}$$

接下來我們利用 CLS 估計變數參數：

$EIAR(2)$ 條件估計式 $E(X_t|\ X_{t-1}=x_{t-1}\ X_{t-2}=x_{t-2})=\phi_1 x_{t-1}+\phi_2 x_{t-2}+\mu$

$$Q_n(\varphi)=\sum_{t=3}^{n}[X_t-(\phi_1 x_{t-1}+\phi_2 x_{t-2}+\mu)]^2$$

$$\frac{\partial Q_n(\varphi)}{\partial \mu}=0 \text{ 得 } \hat{\mu}=\frac{\sum_{t=3}^{n}X_t-\hat{\phi}_1\sum_{t=3}^{n}X_{t-1}-\hat{\phi}_2\sum_{t=3}^{n}X_{t-2}}{n-2}$$

$$\frac{\partial Q_n(\varphi)}{\partial \phi_1}=0 \text{ 得 } \hat{\phi}_1=\frac{\sum_{t=3}^{n}X_tX_{t-1}-\hat{\phi}_2\sum_{t=3}^{n}X_{t-1}X_{t-2}-\hat{\mu}\sum_{t=3}^{n}X_{t-1}}{\sum_{t=3}^{n}X_{t-1}^2}$$

$$\frac{\partial Q_n(\varphi)}{\partial \phi_2}=0 \text{ 得 } \hat{\phi}_2=\frac{\sum_{t=3}^{n}X_tX_{t-2}-\hat{\phi}_1\sum_{t=3}^{n}X_{t-1}X_{t-2}-\hat{\mu}\sum_{t=3}^{n}X_{t-2}}{\sum_{t=3}^{n}X_{t-2}^2}$$

再利用矩陣可解得 $\hat{\mu}$、$\hat{\phi}_1$、$\hat{\phi}_2$。

1.2 EAR

$EAR(1)$ 模型是 $X_t=\phi_1 X_{t-1}+(1-I_1)\varepsilon_t$

$X_t \sim \exp(\mu),\quad E(X_t)=\mu,\quad Var(X_t)=\mu^2$

特徵函數：

$$\varphi_X(t)=(\frac{1}{1+\mu t}),\quad \varphi_X(\phi_1 t)=(\frac{1}{1+\mu\phi_1 t})$$

$$\varphi_{\varepsilon_t}(t)=\frac{\varphi_X(t)}{\varphi_X(\phi_1 t)}=(\phi_1+(1-\phi_1)\frac{1}{1+\mu t})$$

接下來我們利用 CLS 估計變數參數：

$$Q_n(\psi)=\sum_{t=2}^{n}[x_t-(\phi_1 x_{t-1}+\mu(1-\phi_1))]^2$$ ，其中 ψ=（ϕ_1、μ)

$$\frac{\partial Q_n(\psi)}{\partial \phi_1}=0 \text{ 得 } \hat{\phi}_1=\frac{\sum_{t=2}^{n} X_t X_{t-1}-(n-1)^{-1}-\sum_{t=2}^{n} X_t \sum_{t=2}^{n} X_{t-1}}{\sum_{t=2}^{n} X_{t-1}^2-(n-1)^{-1}(\sum_{t=2}^{n} X_{t-1})^2}$$

$$\frac{\partial Q_n(\psi)}{\partial \mu}=0 \text{ 得 } \hat{\mu}=\frac{\sum_{t=2}^{n} X_t-\hat{\phi}_1 \sum_{t=2}^{n} X_{t-1}}{(n-1)(1-\hat{\phi}_1)}$$

接下來我們引用文獻結果，將過去的時間點推廣至兩個：

$EAR(2)$ 模型是 $X_t=\phi_1 X_{t-1}+\phi_2 X_{t-2}+(1-I_1 I_2)\varepsilon_t$

$X_t \sim \exp(\mu), \quad E(X_t)=\mu, \quad Var(X_t)=\mu^2$

$$\varphi_X(t)=(\frac{1}{1+\mu t}), \quad \varphi_X(\phi_1 t)=(\frac{1}{1+\mu\phi_1 t}), \quad \varphi_X(\phi_2 t)=(\frac{1}{1+\mu\phi_2 t})$$

$$\varphi_{\varepsilon_t}(t)=\frac{\varphi_X(t)}{\varphi_X(\phi_1 t)\varphi_X(\phi_2 t)}=\frac{(1+\mu\phi_1 t)(1+\mu\phi_2 t)}{(1+\mu t)}$$

在我們利用 CLS 估計變數的推導上遇到了一些困難，所以我們藉由電腦軟體（R 程式，最小平方法）估計出參數。

第二節　BGAR

本文主要介紹 Miroslav M.Ristic(2005)文獻中所提出的 BGAR(2)

時間序列模型：

BGAR(2)模型是

$$X_t = \begin{cases} B_{t1}X_{t-1} + G_{t1}\ , w.p\ \ \dfrac{\delta_2(1-\delta_2)}{\delta_1^2-\delta_2^2} \\ B_{t2}X_{t-2} + G_{t2}\ , w.p\ \ \dfrac{\delta_1^2-\delta_2}{\delta_1^2-\delta_2^2} \end{cases}$$

$$B_{t1} \sim Beta(k\delta_i, k(1-\delta_i)), \quad G_{t1} \sim Gamma(k(1-\delta_j), \beta)$$

如果$\delta_2 = \delta_1^2$，那麼$\{X_t\}$就是 BGAR(1)的過程。

參數估計的部分，我們直接引用文獻上的結果，其結果如下：

$$\tilde{\mu}_N = \frac{1}{N}\sum_{i=1}^{N} X_i, \quad \hat{\sigma}_N^2 = \frac{1}{N}\sum_{i=1}^{N}(X_i - \overline{X}_N)^2, \quad \tilde{k}_N = \frac{\tilde{\mu}_N^2}{\hat{\sigma}_N^2} \quad \tilde{\beta}_N = \frac{\tilde{\mu}_N}{\hat{\sigma}_N^2}$$

$$\tilde{\delta}_{1N} = \frac{\sum_{j=3}^{N}(X_j - \overline{X}_N)(X_{j-2} - \overline{X}_N)}{\sum_{j=2}^{N}(X_j - \overline{X}_N)(X_{j-1} - \overline{X}_N)}, \quad \tilde{\delta}_{2N} = \frac{\sum_{j=3}^{N}(X_j - \overline{X}_N)(X_{j-2} - \overline{X}_N)}{\sum_{j=1}^{N}(X_j - \overline{X}_N)^2}$$

X_t 經由轉換，我們可以求得它是服從 $Gamma(k$、$\beta)$ 的邊際分配。

第參章　研究方法

首先，本文經由研究目的，發展出研究架構。接著，了解研究架構後，再發展出模型假設、估計式推導、數據選取，最後為估計結果比較。

第一節　模型假設

一開始我們假設了一個線性迴歸模型，是由 p 個變數所組合而成的，但變數不是一般常用的分配，而是時間序列，在殘差項的部分假設為常態分配。

假設模型：

$$Y = \beta_1 \underset{\sim}{X}_1 + \beta_2 \underset{\sim}{X}_2 + \cdots + \beta_{p-1} \underset{\sim}{X}_{p-1} + \beta_p \underset{\sim}{X}_p + \varepsilon_t$$

$$where \quad \varepsilon_t \overset{iid}{\sim} N(0, \sigma_t^2)$$

本研究在模型中考慮五個時間序列的共變數，$\{\underset{\sim}{X}_1\}$、$\{\underset{\sim}{X}_2\}$、$\{\underset{\sim}{X}_3\}$、$\{\underset{\sim}{X}_4\}$、$\{\underset{\sim}{X}_5\}$分別為：

1. ARMA(2, 2)　時間序列兩列
2. EIAR(2)　時間序列一列
3. EAR(2)　時間序列一列
4. BGAR(2)　時間序列一列

第二節　估計式推導

本章節所要推導的是，我們提出估計參數新方法的過程。一開始我們在給定條件的情況下，假設$Y \mid (\underset{\sim}{X_1}、\underset{\sim}{X_2}、\underset{\sim}{X_3}、\underset{\sim}{X_4}、\underset{\sim}{X_5})$為常態分配，給予 $X_s^{'}$，$(Y_1 \cdots Y_n)$ 為獨立，再利用迴旋積分計算出Y的機率密度函數和 Likelihood Function，由於計算量太大，所以我們利用了 Monte Carlo Methods 去逼近，將結果再利用 Newton-Raphson Method 來估計參數$\underset{\sim}{\beta}$值。

推導式子：

$$
\begin{aligned}
f(y_1 \ldots y_n) &= \int \cdots \int f(y_1 \ldots y_n | \underset{\sim}{x_1} \ldots \underset{\sim}{x_5}) f(\underset{\sim}{x_1} \ldots \underset{\sim}{x_5}) d\underset{\sim}{x_1} \ldots d\underset{\sim}{x_5} \\
&= \int \cdots \int \prod_{i=1}^{n} f(y_i | \underset{\sim}{x_1} \ldots \underset{\sim}{x_5}) f(\underset{\sim}{x_1}) \ldots f(\underset{\sim}{x_5}) d\underset{\sim}{x_1} \ldots d\underset{\sim}{x_5} \\
&= \int \cdots \int \prod_{i=1}^{n} [\frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{1}{2\sigma^2}(y_i - (\beta_1 x_{i1} + \ldots + \beta_5 x_{i5}))^2\}] \\
&\quad \times f(x_{i1}) \ldots f(x_{i5}) dx_{i1} \ldots dx_{i5}
\end{aligned}
$$

因為在時間序列所組成的線型模型中，機率密度函數$f(y_1 \cdots y_n)$無法直接相乘，必須在有給定的條件下才可以拆開相乘，其原因，為$(y_1 \cdots y_n)$是一組時間序列資料變數觀察值之間有相關性。然而利用給予 $X_s^{'}$、$(Y_1 \cdots Y_n)$ 條件獨立下的性質獲得$f(y_1 \cdots y_n)$。

註 Monte Carlo Method

假設$X_1 \cdots X_N \sim f(y)$ ，$h(X)$是X的函數

其式子：$E(H(X)) = (1/N) \times \sum_{j=1}^{N} H(X_j)$

◆ Monte Carlo Method：

Likelihood Function：

$$f(y_1 \ldots y_n) = \int \cdots \int \underbrace{\prod_{i=1}^{n} [\frac{1}{\sqrt{2\pi\sigma}} \exp\{-\frac{1}{2\sigma^2}(y_i - (\beta_1 x_{i1} + \ldots + \beta_5 x_{i5}))^2\}]}_{H(X)}$$

$$\times f(x_{i1}) \ldots f(x_{i5}) dx_{i1} \ldots dx_{i5}$$

可表示成

$E(H(X))$=Likelihood Function=L

其 Monte Carlo 之逼近式

$$\widetilde{f}(y_1 \cdots y_n) = \hat{E}(H(X))$$

$$= \frac{1}{N}\sum_{j=1}^{N}[(\frac{1}{\sqrt{2\pi\sigma}})^n \exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - (\beta_1 \widetilde{x}_{j1} + \ldots + \beta_5 \widetilde{x}_{j5}))^2\}]$$

式子中的 $\widetilde{X}_{j1} \cdots \widetilde{X}_{j5}$ 是由原始資料 $\{\underset{\sim}{X}_1\} \cdots \{\underset{\sim}{X}_5\}$ 所估計出來的參數值，代入原來共變數的機率函數中，再重新抽取出來的資料，其如何抽取在第肆章會詳細說明。

Log-Likelihood Function=l

$$=> \log(\frac{1}{N}\sum_{j=1}^{N}[(\frac{1}{\sqrt{2\pi\sigma}})^n \exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i - (\beta_1 \widetilde{x}_{j1} + \ldots + \beta_5 \widetilde{x}_{j5}))^2\}]) = l$$

接下來

- 對每一個變數作一次偏微：

$$\frac{\partial l}{\partial \beta_k} = \frac{\partial \log L}{\partial \beta_k} = \frac{1}{L}\frac{\partial \mathrm{L}}{\partial \beta_k} \quad k = 1 \sim 5$$

$$\frac{\partial L}{\partial \beta_k}=\frac{1}{N}\sum_{j=1}^{N}(\frac{1}{\sqrt{2\pi}\sigma})^n\frac{\widetilde{x}_{jk}}{\sigma^2}\underbrace{\sum_{i=1}^{n}(y_i-(\beta_1\widetilde{x}_{j1}+\ldots+\beta_5\widetilde{x}_{j5}))}_{A}$$

$$\times\underbrace{\exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(y_i-(\beta_1\widetilde{x}_{j1}+\ldots+\beta_5\widetilde{x}_{j5}))^2\}}_{B}$$

其餘的變數與上述同理。

- 對每一個變數作兩次偏微：

$$\frac{\partial^2 l}{\partial \beta_k^2}=\frac{1}{L^2}[\frac{\partial^2 \mathrm{L}}{\partial \beta_k^2}L-(\frac{\partial L}{\partial \beta_k})^2]\quad k=1\cdots 5$$

$$\begin{aligned}\frac{\partial^2 L}{\partial \beta_k^2}&=\frac{1}{N}\sum_{j=1}^{N}[(\frac{1}{\sqrt{2\pi}\sigma})^n(-\frac{\widetilde{x}_{jk}^2}{\sigma^2})B]+\frac{1}{N}\sum_{j=1}^{N}[(\frac{1}{\sqrt{2\pi}\sigma})^n(\frac{\widetilde{x}_{jk}}{\sigma^2})A(\frac{\widetilde{x}_{jk}}{\sigma^2})A\times B]\\&=\frac{1}{N}\sum_{j=1}^{N}[(\frac{1}{\sqrt{2\pi}\sigma})^n(-\frac{\widetilde{x}_{jk}^2}{\sigma^2})B]+\frac{1}{N}\sum_{j=1}^{N}[(\frac{1}{\sqrt{2\pi}\sigma})^n(\frac{\widetilde{x}_{jk}}{\sigma^2})^2A^2\times B]\\&=\frac{1}{N}\sum_{j=1}^{N}[(\frac{1}{\sqrt{2\pi}\sigma})^n(-\frac{\widetilde{x}_{jk}^2}{\sigma^2})B]\times[\frac{1}{\sigma^2}A^2-1]\end{aligned}$$

其餘的變數與上述同理。

- 對兩兩變數作偏微：

$$\frac{\partial^2 l}{\partial \beta_k\partial \beta_{k^*}}=\frac{1}{L^2}[\frac{\partial^2 \mathrm{L}}{\partial \beta_k\partial \beta_{k^*}}L-(\frac{\partial L}{\partial \beta_k})(\frac{\partial L}{\partial \beta_{k^*}})]\quad k\neq k^*\quad k=1\cdots 5\quad k^*=1\cdots 5$$

$$\begin{aligned}&=\frac{1}{N}\sum_{j=1}^{N}[(\frac{1}{\sqrt{2\pi}\sigma})^n(-\frac{\widetilde{x}_{jk}}{\sigma^2})(x_{jk^*})B]+\frac{1}{N}\sum_{j=1}^{N}[(\frac{1}{\sqrt{2\pi}\sigma})^n(\frac{\widetilde{x}_{jk}}{\sigma^2})(\frac{\widetilde{x}_{jk^*}}{\sigma^2})A^2\times B]\\&=\frac{1}{N}\sum_{j=1}^{N}[(\frac{1}{\sqrt{2\pi}\sigma})^n(-\frac{\widetilde{x}_{jk}}{\sigma^2})(x_{jk^*})B]\times[\frac{1}{\sigma^2}A^2-1]\end{aligned}$$

其餘的變數與上述同理。

所有的變數都經過微分步驟後，再經由整理，藉由 Newton Raphson Method 來估計$\underset{\sim}{\beta}$參數值。

◆ Newton Raphson Method：

$$\beta^{(t+1)} = \beta^{(t)} - (H^{(t)})^{-1} u^{(t)}$$

$$where\ \beta^{(0)} = \text{起始值}$$

$$\underset{\sim}{\beta} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_5 \end{bmatrix} \quad \underset{\sim}{u} = \begin{bmatrix} \dfrac{\partial l}{\partial \beta_1} \\ \vdots \\ \dfrac{\partial l}{\partial \beta_5} \end{bmatrix} \quad \underset{\sim}{H} = \begin{bmatrix} \dfrac{\partial^2 l}{\partial^2 \beta_1} & \cdots & \dfrac{\partial^2 l}{\partial \beta_1 \partial \beta_5} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial^2 l}{\partial \beta_5 \partial \beta_1} & \cdots & \dfrac{\partial^2 l}{\partial^2 \beta_5} \end{bmatrix}$$

這兩個估計方法將由多次重覆後，會得到一組較收斂的$\underset{\sim}{\beta}$值，我們再利用這一組值與最小平方法所估計出來的參數值進行結果比較。

接下來，我們從資料重新取的部分，包括了 Monte Carlo Method、Newton Raphson Method，重覆多次的計算，估計出多組的收斂$\underset{\sim}{\beta}$參數值，再將由這一些參數值計算出標準差，再與最小平方法估計出來的標準差進行比較，加以確定結果。

第肆章　數據取樣

第一節　原始資料生成

在前面第壹章和第參章的部分均有提到，模型是由五條時間序列所組合而成的，我們將利用以下已知的參數生成出原始資料，用以分析。

在模型使用假設的真實參數值如下：

1. $ARMA(2,2)$
$$\begin{cases} AR\text{參數}0.5\text{、}0.2\text{，}MA\text{參數}-0.2\text{、}-0.3\text{，標準差}0.2 \\ AR\text{參數}-0.5\text{、}0.2\text{，}MA\text{參數}0.3\text{、}-0.4\text{，標準差}0.3 \end{cases}$$

2. $EAIR$ 模型　$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$

已知第一個時間點和第二個時間點的值分別為$X_1 = 0.3$、$X_2 = 0.2$

參數　$\phi_1 = 0.9$　$\phi_2 = -0.1$　$\varepsilon_t \sim \exp(1)$

3. EAR 模型　$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-1} + (1 - I_1 I_2)\varepsilon_t$

已知X_t的分配所求得的兩時間點的值$X_1 = 0.6627$、$X_2 = 0.1235$

參數　$\phi_1 = 0.2$　$\phi_2 = -0.1$

$$X_t \sim \exp\ (1) \quad I_1\text{、}I_2 \sim bernoulli(0.7) \quad \varepsilon_t \sim \exp(1)$$

4. $BGAR$ 模型　$X_t = \begin{cases} B_{t1} X_{t-1} + G_{t1}\ , w.p\ \dfrac{\delta_2(1-\delta_2)}{\delta_1^2 - \delta_2^2} \\ B_{t2} X_{t-2} + G_{t2}\ , w.p\ \dfrac{\delta_1^2 - \delta_2}{\delta_1^2 - \delta_2^2} \end{cases}$

參數　$B_{t1} \sim Beta(k\delta_i, k(1-\delta_i)),\quad G_{t1} \sim Gamma(k(1-\delta_j), \beta)$

$i, j = \{1,2\}, \quad \beta = 1, \quad k = 2, \quad \delta_1 = 0.9284, \quad \delta_2 = 0.6924$

where $0 < \delta_2 \le \delta_1^2 < 1$,

$\{B_{t1}\}$ 和 $\{G_{t1}\}$ 之間是相互獨立的。

每一條時間序列均生成 n 筆，當作原始資料，我們再利用這 n 筆資料，進行第貳章所推導出的估計式，來估計每一條時間序列的參數，將這一些估計出來的參數，重新放回時間序列，並且抽取新的資料，以利於我們接下來的計算。

第二節　重新抽取資料

在假設的模型中 $Y = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 + \beta_5 X_5 + \varepsilon_t$，對於 ARMA(2,2)的時間序列 $\{X_1\}$、ARMA(2,2)的時間序列 $\{X_2\}$、EIAR(2)的時間序列 $\{X_3\}$、EAR(2)的時間序列 $\{X_4\}$、BGAR(2)的時間序列 $\{X_5\}$ 依資料估計參數如下：

1. *ARMA*

$$\begin{cases} AR\text{參數}0.3303、0.2312，MA\text{參數}-0.0249、-0.2881，\text{標準差}0.2 \\ AR\text{參數}-0.6738、0.0772，MA\text{參數}0.4386、-0.3249，\text{標準差}0.3 \end{cases}$$

2. *EAIR*(2) 模型 $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \varepsilon_t$

已知第一個時間點和第二個時間點的值分別為 $X_1 = 0.3$、$X_2 = 0.2$

估計參數 $\phi_1 = 0.9304 \quad \phi_2 = -0.1293 \quad \varepsilon_t \sim \exp(0.9676)$

3. *EAR* 模型 $X_t = \phi_1 X_{t-1} + \phi_2 X_{t-1} + (1 - I_1 I_2)\varepsilon_t$

已知 X_t 的分配所求得的兩時間點的值 $X_1 = 0.6627$、$X_2 = 0.1235$

估計參數　$\phi_1 = 0.2265, \quad \phi_2 = -0.1173$

$$X_t \sim \exp\ (1), \quad I_1、I_2 \sim bernoulli(1,0.7), \quad \varepsilon_t \sim \exp(0.9923)$$

4. $BGAR$ 模型　$X_t = \begin{cases} B_{t1}X_{t1} + G_{t1}\ ,w.p\ \dfrac{\delta_2(1-\delta_2)}{\delta_1^2 - \delta_2^2} \\ B_{t2}X_{t2} + G_{t2}\ ,w.p\ \dfrac{\delta_1^2 - \delta_2}{\delta_1^2 - \delta_2^2} \end{cases}$

估計參數　$B_{t1} \sim Beta(k\delta_i, k(1-\delta_i)), \quad G_{t1} \sim Gamma(k(1-\delta_j), \beta)$

$$i, j = \{1,2\}, \quad \beta = 0.8652006, \quad k = 1.872567$$

$$\delta_1 = 0.8606307, \quad \delta_2 = 0.7401243$$

where $0 < \delta_2 \le \delta_1^2 < 1$,

$\{B_{t1}\}$ 和 $\{G_{t1}\}$ 之間是相互獨立的。

再重新抽取資料的部分，我們生成了 N 筆，為了不受起始值的影響，前面的 M 筆資料不使用。

接下來，我們使用原始資料和重新抽取資料，作資料上的篩選。

第三節　資料篩選

一開始，我將所有的資料都丟進去計算，但在 Likelihood Function 的計算上會出現值很小的情況，導致接下來的運算上出現了困難，甚至出現無法計算的狀況，所以我將資料做了特殊篩選，這篩選的目的是為了讓 Likelihood 值盡可能達到最大值，但在做這篩選的動作時已經有一點破壞了 Monte Carlo Method 原本的涵義了，即使這樣這一個

估計方法還是可以有效的估計出 $\underset{\sim}{\beta}$ 參數值。我們先估計共變數分配 $f(\underset{\sim}{x_1})\cdots f(\underset{\sim}{x_5})$ 的參數值代入後將估計的共變數分配函數寫成 $\hat{f}(\underset{\sim}{x_1})\cdots\hat{f}(\underset{\sim}{x_5})$。

$\widetilde{x}_{j1} \sim \hat{f}(\underset{\sim}{x_1}) \quad j = 1\cdots N$

$\widetilde{x}_{j2} \sim \hat{f}(\underset{\sim}{x_2}) \quad j = 1\cdots N$

$\vdots$

$\widetilde{x}_{j5} \sim \hat{f}(\underset{\sim}{x_5}) \quad j = 1\cdots N$

這裡 $\underset{\sim}{x}$ 是一組時間序列的資料 $\{x_{i1}\colon i = 1\cdots n\}$

將第 j 筆抽出的樣本以向量表示為 $\widetilde{\boldsymbol{x}}_j^T = (\widetilde{x}_{j1}\cdots\widetilde{x}_{j5})$。

在假設的模型中，我們針對真實值 $\beta_1\cdots\beta_5$，會有以下的線性關係 $\widetilde{y}_j = \beta_1\widetilde{x}_{j1} + \beta_2\widetilde{x}_{j2} + \beta_3\widetilde{x}_{j3} + \beta_4\widetilde{x}_{j4} + \beta_5\widetilde{x}_{j5} + \varepsilon_t$，然而 $\widetilde{y}_j$ 並未觀察到或模擬而得。如果我們將所有的重抽資料 $\widetilde{\boldsymbol{x}}_j^T \quad j = 1\cdots N$，丟進去 Monte Carlo 的式子中計算，基於 $\widetilde{y}_j$ 可能遠離原始資料 $y_1\cdots y_n$，則重新抽出的資料中有些值帶入後會使 Likelihood 值過低，因此可能會造成 $\widetilde{f}(y_i)$ 值過小，而使得參數值估計不理想的狀態，在第伍章會有詳細說明。因此，我們在 $\widetilde{\boldsymbol{x}}_j^T \quad j = 1\cdots N$ 中做了特別的選取。這樣的選取是為了使 Likelihood 值在現有給予的資料上盡可能靠近最佳值。

假設 $\boldsymbol{X}$ 為原始共變數的資料矩陣，$\widetilde{\boldsymbol{X}}$ 為重抽的共變數的資料矩陣：

$$\boldsymbol{X}=\begin{bmatrix} x_{11} & \cdots & x_{15} \\ \vdots & \ddots & \vdots \\ x_{n1} & \cdots & x_{n5} \end{bmatrix}=\begin{bmatrix} \boldsymbol{x}_1^T \\ \vdots \\ \boldsymbol{x}_5^T \end{bmatrix} \quad , \quad \widetilde{\boldsymbol{X}}=\begin{bmatrix} \widetilde{x}_{11} & \cdots & \widetilde{x}_{15} \\ \vdots & \ddots & \vdots \\ \widetilde{x}_{N1} & \cdots & \widetilde{x}_{N5} \end{bmatrix}=\begin{bmatrix} \widetilde{\boldsymbol{x}}_1^T \\ \vdots \\ \widetilde{\boldsymbol{x}}_5^T \end{bmatrix}$$

接下來將二資料矩陣對每一個共變數作標準化，標準化後的資料矩陣為：

$$\boldsymbol{X}_{sd}=[\boldsymbol{z}_1\cdots\boldsymbol{z}_1]=\begin{bmatrix} z_{11} & \cdots & z_{15} \\ \vdots & \ddots & \vdots \\ z_{n1} & \cdots & z_{n5} \end{bmatrix} \quad , \quad \widetilde{\boldsymbol{X}}_{sd}=[\widetilde{\boldsymbol{z}}_1\cdots\widetilde{\boldsymbol{z}}_1]=\begin{bmatrix} \widetilde{z}_{11} & \cdots & \widetilde{z}_{15} \\ \vdots & \ddots & \vdots \\ \widetilde{z}_{N1} & \cdots & \widetilde{z}_{N5} \end{bmatrix}$$

選取的過程是首先在步驟 1 步驟 2 中填補合適的$\widetilde{y}_j$於所對應的$\widetilde{\boldsymbol{x}}_j^T$，再由步驟 3 中選出$\widetilde{\boldsymbol{x}}_j^T \quad j=1\cdots N$其$\widetilde{y}_j$是接近$y_i$，步驟如下：

步驟 1：

對於一個$\widetilde{\boldsymbol{x}}_j^T$，選出最接近的$\widetilde{\boldsymbol{x}}_i^T=(\widetilde{\boldsymbol{x}}_{i1}^T\cdots\widetilde{\boldsymbol{x}}_{i5}^T) \quad i=1\cdots n$，依據標準差後的絕對差總和，即算出$\sum_i\left|z_{ip}-\widetilde{z}_{jp}\right|=A_{i(j)} \quad for\, i=1\cdots n$ $,p=1\cdots5$找出$\min(A_{1(j)}\cdots A_{n(j)})$，令$i^*=\arg\min(A_{1(j)}\cdots A_{n(j)})$給予一個邊際值$D_1>0$，如果$\min(A_{1(j)}\cdots A_{n(j)})<D_1$並且對於此選擇到的設$\widetilde{y}_j=y_{i^*}$則$\widetilde{y}_j^T$選擇入袋。

步驟 2：

共變數中原始資料的其中一個不考慮列入比較，只考慮其他 4 個變數與$\widetilde{\boldsymbol{x}}_j^T$的相對 4 的共變數值接近時之情況，然後我們填補 y 值。

以去掉X_5共變數為例：

依據標準化的資料計算$\sum_i\left|z_{ip^*}-\widetilde{z}_{jp^*}\right|=A_{i(j)} \quad for\, i=1\cdots n\, ,p=1\cdots4$找

出 $\min(A_{1(j)}\cdots A_{n(j)})$ ，令 $i^{*}=\arg\min(A_{1(j)}\cdots A_{n(j)})$ 給予兩個邊際值 $D_1, D_2>0$ ，如果 $\min(A_{1(j)}\cdots A_{n(j)})=A_{i^{*}(j)}<D_1$ ，找出 $A_{i(j)}$ 中符合 $\left|A_{i(j)}-A_{i^{*}(j)}\right|<D_2$ 的值，且令 $B(y)=\{y_i, i=1\cdots n：\left|A_{i(j)}-A_{i^{*}(j)}\right|<D_2\}$。在 $B(y)$ 這集合上假設有 c 個 y 值，分別計為 $y^{(1)}\cdots y^{(c)}$，計算

$$\widetilde{y}_j=\frac{\sum_{l=1}^{c} y_l}{c}+\left(y_{i^{*}}-\frac{\sum_{l=1}^{c} y_l}{c}\right)\times\frac{\widetilde{x}_{i^{*}p}}{x_{i^{*}p}}$$

，也將此 $\widetilde{y}_j$ 選擇入袋。

步驟 3：

我們利用步驟 1、2 所填補的 $\widetilde{y}_j$ 與原始的每一個 y 作比較，兩者相減小於一個常數(r)，則將小於 r 的 y 帶入計算式中，進行運算。

圖解說明第 3 步驟：

$\left.\begin{matrix}\widetilde{y}_1\\ \vdots\\ \widetilde{y}_N\end{matrix}\right\}$ 由步驟1、2所填補的 $\widetilde{y}_{\mathrm{j}}$ 設一個邊際值 $r>0$，將每一個 $\widetilde{y}_j$ ，算出 $\left|\widetilde{y}_j-y_i\right|$ $i=1\cdots n$ 若 $\left|\widetilde{y}_j-y_i\right|<r$ ，則將所對應的 $\widetilde{\boldsymbol{x}}_j$ 及 y_i 放入 $\widetilde{f}(y_1\cdots y_n)$ 進行運算。

第五章　邊際值的選取

根據我們所假設的時間序列迴歸模型，推導出Y的機率密度函數和 Likelihood Function，推導過程與結果我們利用了本文的第三章並從式子中說明如何選取邊際值。

$$f(y_1 \ldots y_n) = \int \cdots \int f(y_1 \ldots y_n | \underset{\sim}{x_1} \ldots \underset{\sim}{x_{i5}}) f(\underset{\sim}{x_1} \ldots \underset{\sim}{x_5}) d\,\underset{\sim}{x_1} \ldots d\underset{\sim}{x_5}$$

$$= \int \cdots \int \prod_{i=1}^{n} [\frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{1}{2\sigma^2}(y_i - \boldsymbol{x}_i^T \boldsymbol{\beta})^2\}]\ f(x_{i1}) \ldots f(x_{i5}) dx_{i1} \ldots dx_{i5}$$

$$= \int \cdots \int \prod_{i=1}^{n} [\frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{1}{2\sigma^2}(y_i - \boldsymbol{x}_i^T \boldsymbol{\beta})^2 \times \mathbf{1}\{\|y_i - \boldsymbol{x}_i^T \boldsymbol{\beta}\|^2 \le M_i\}\}]$$

$$\times f(x_{i1}) \ldots f(x_{i5}) dx_{i1} \ldots dx_{i5} + \underbrace{h(M_i)}_{small}$$

這裡$h(M_i)$會是一個很小的值

這裡$\boldsymbol{x}_i^T = (x_{i1} \ldots x_{i5}), \quad \beta^T = (\beta_1 \cdots \beta_5)$

接下來利用 Monte Carlo 去逼近 Likelihood Function，其逼近式：

$$\widetilde{f}(y_1 \cdots y_n) = \frac{1}{N} \sum_{j=1}^{N} \prod_{i=1}^{N(j)} [\frac{1}{\sqrt{2\pi}\sigma} \exp\{-\frac{1}{2\sigma^2}(y_i - \widetilde{\boldsymbol{x}}_j^T \boldsymbol{\beta})^2$$

$$\times \mathbf{1}\{(y_i - \boldsymbol{x}_j^T \boldsymbol{\beta})^2 \le M_i\}\}$$

這裡

$$N_{(j)} = \#\ \{\mathbf{1}\{(y_i - \boldsymbol{x}_j^T \boldsymbol{\beta})^2 \le M_i\}\} \quad for\ j = 1 \cdots N$$

$$N = \#\{j : N_{(j)} > 0\}$$

註：N是篩選步驟1、2所得的重抽筆數中又不使$N_{(j)}$為0的個數

$$\widetilde{f}(y_1 \cdots y_n) = \frac{1}{N}\sum_{j=1}^{N}[(\frac{1}{\sqrt{2\pi\sigma}})^{N_{(j)}} \exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N_{(j)}}(y_i - \widetilde{\boldsymbol{x}}_j^T \boldsymbol{\beta})^2$$

$$\times \mathbf{1}\{(y_i - \boldsymbol{x}_j^T \boldsymbol{\beta})^2 \le M_i\}\}$$

$$= \frac{1}{N}\sum_{j=1}^{N}[(\frac{1}{\sqrt{2\pi\sigma}})^{N_{(j)}} \exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N_{(j)}}(y_i - \widetilde{\boldsymbol{x}}_j^T \boldsymbol{\beta})^2\}$$

$$\widetilde{L}(\beta) = \widetilde{f}(y_1 \cdots y_n)$$

$$\widetilde{l}(\beta) = \log \widetilde{L}(\beta)$$

$$\frac{\partial \widetilde{l}(\beta)}{\partial \beta_k} = \frac{\partial \log \widetilde{L}(\beta)}{\partial \beta_k} = \frac{1}{\widetilde{L}(\beta)}\frac{\partial \widetilde{\mathrm{L}}(\beta)}{\partial \beta_k} \quad k = 1 \sim 5$$

$$\frac{\partial \widetilde{L}(\beta)}{\partial \beta_k} = \frac{1}{N}\sum_{j=1}^{N}(\frac{1}{\sqrt{2\pi\sigma}})^{N_{(j)}} \frac{\widetilde{x}_{jk}}{\sigma^2} d_j \times w_{ij}$$

這裡 $d_j = \sum_{i=1}^{N_{(j)}}(y_i - \boldsymbol{x}_j^T \boldsymbol{\beta})^2$， $w_{ij} = \exp\{-\frac{1}{2\sigma^2}\sum_{i=1}^{N_{(j)}}(y_i - \boldsymbol{x}_j^T \boldsymbol{\beta})^2\}$

$$\frac{\partial^2 \widetilde{l}(\beta)}{\partial \beta_k \beta_{k^*}} = \frac{1}{[\widetilde{L}(\beta)]^2}[\frac{\partial^2 \widetilde{\mathrm{L}}(\beta)}{\partial \beta_k \beta_{k^*}}\widetilde{L}(\beta) - (\frac{\partial \widetilde{L}(\beta)}{\partial \beta_k})(\frac{\partial \widetilde{L}(\beta)}{\partial \beta_{k^*}})] \quad k \ne k^*$$

$$\frac{\partial^2 \widetilde{\mathrm{L}}(\beta)}{\partial \beta_k \beta_{k^*}} = \frac{1}{N}\sum_{j=1}^{N}[(\frac{1}{\sqrt{2\pi\sigma}})^{N_{(j)}}(-\frac{\widetilde{x}_{jk}}{\sigma^2})(\widetilde{x}_{jk^*})w_{ij}][\frac{1}{\sigma^2}d_i^2 - 1]$$

$$\frac{\dfrac{\partial \widetilde{L}(\beta)}{\partial \beta_k}}{\widetilde{L}(\beta)} = \frac{\dfrac{1}{N}\displaystyle\sum_{j=1}^{N}(\frac{1}{\sqrt{2\pi\sigma}})^{N_{(j)}} \frac{\widetilde{x}_{jk}}{\sigma^2} d_j \times w_{ij}}{\dfrac{1}{N}\displaystyle\sum_{j=1}^{N}(\frac{1}{\sqrt{2\pi\sigma}})^{N_{(j)}} w_{ij}} = \sum_{j=1}^{N}\widetilde{x}_{jk} b_j^*$$

$$\frac{(\dfrac{\partial \widetilde{L}(\beta)}{\partial \beta_k})(\dfrac{\partial \widetilde{L}(\beta)}{\partial \beta_{k^*}})}{[\widetilde{L}(\beta)]^2} = \frac{\displaystyle\sum_{j=1}^{N}\frac{\widetilde{x}_{jk}}{\sigma^2} w_{ij} d_j \sum_{j=1}^{N}\frac{\widetilde{x}_{jk^*}}{\sigma^2} w_{ij} d_j}{(\displaystyle\sum_{j=1}^{N} w_j)^2} = \sum_{j=1}^{N}\widetilde{x}_{jk} b_j \times \sum_{j=1}^{N}\widetilde{x}_{jk^*} b_j$$

$$= \tilde{X}_k^T B \tilde{X}_{k^*} \quad \tilde{X}_k^T = (\tilde{x}_{1k} \cdots \tilde{x}_{N(j)k})$$

經由整理 $\dfrac{(\dfrac{\partial \tilde{L}(\beta)}{\partial \beta_k})(\dfrac{\partial \tilde{L}(\beta)}{\partial \beta_{k^*}})}{[\tilde{L}(\beta)]^2}$ 可整理為 $\tilde{X}^T B \tilde{X}$

$$\frac{\dfrac{\partial^2 \tilde{L}(\beta)}{\partial \beta_k \beta_{k^*}}}{\tilde{L}(\beta)} = \frac{\dfrac{1}{N}\sum_{j=1}^{N}[(\dfrac{1}{\sqrt{2\pi}\sigma})^{N(j)}(-\dfrac{\tilde{x}_{jk}}{\sigma^2})(\tilde{x}_{jk^*})w_{ij}][\dfrac{1}{\sigma^2}d_i^2-1]}{\dfrac{1}{N}\sum_{j=1}^{N}[(\dfrac{1}{\sqrt{2\pi}\sigma})^{N(j)}w_{ij}]} = \sum_{j=1}^{N}\tilde{x}_{jk}\tilde{x}_{jk^*}a_j$$

$$= \tilde{X}_k^T A \tilde{X}_{k^*}, \quad \tilde{X}_k^T = (\tilde{x}_{1k} \cdots \tilde{x}_{N(j)k})$$

經由整理 $\dfrac{\dfrac{\partial^2 \tilde{L}(\beta)}{\partial \beta_k \beta_{k^*}}}{\tilde{L}(\beta)}$ 可整理為 $\tilde{X}^T A \tilde{X}$

Newton Raphson Method：

$$\beta^{(t+1)} = \beta^{(t)} - (H^{(t)})^{-1}u^{(t)} \qquad where\ \beta^{(0)} = \text{起始值}$$

$$\underset{\sim}{H} = \left[\frac{\partial^2 l}{\partial \beta_k \partial \beta_{k^*}}\right]_{k,k^*} = \tilde{X}^T A \tilde{X} - \tilde{X}^T B \tilde{X}$$

如果選擇邊際值小，也表示這裡的M_i值小，會使得A、B矩陣中的元素值太小，而可能造成H^t的反矩陣無法運算；如果選擇邊際值大，也表示這裡的M_i值大，會使得估計範圍太大，所以我們必須選擇一個適當大小的邊際值，讓我們所要估計的$\underset{\sim}{\beta}$參數值達到最佳化。

第陸章　分析結果

本文，我們做了兩組模擬，使用了我們所篩選出來的數據資料，進行$\underset{\sim}{\beta}$參數估計，在估計式子的部分我們重覆了 20 次，讓值呈現收斂，接下來，再將重新抽取和估計式的部分重覆 30 次，並使用這 30 筆估計出來的$\underset{\sim}{\beta}$參數值，計算標準差，進而更進一步的了解新方法和最小平方法的優缺點。

其估計結果如下：

第一組：

表一　第一組估計值

解釋變數	X_1	X_2	X_3	X_4	X_5
分配	ARMA	ARMA	EAIR	EAR	BGAR
β已知	0.6	0.5	0.7	0.5	0.9
$\hat{\beta}$(LSE)	0.45272	0.31856	0.68179	0.53423	0.88838
Sd(LSE)	0.15955	0.13292	0.02452	0.09578	0.04970
$\hat{\beta}$(NEW)	0.64357	0.48948	0.60209	0.48671	0.91598
Sd(NEW)	0.05998	0.04130	0.00856	0.05476	0.02779

$\hat{\beta}$(LSE)是利用原始資料所估計出來的

$\hat{\beta}$(NEW)是利用原始資料與重新抽取資料所估計出來的

由表一，可以看出$\underset{\sim}{\beta}$參數值的估計上沒有很明顯的差異，但是在標準差來看，很明顯的看出，新方法的變異比最小平方法來的小很多，由此可見，新方法估計出來的結果比較接近原始參數值。

第二組：

表二　第二組估計值

解釋變數	X_1	X_2	X_3	X_4	X_5
分配	ARMA	ARMA	EAIR	EAR	BGAR
β 已知	1	2	3	4	5
$\hat{\beta}$(LSE)	0.85272	1.81856	2.98179	4.03423	4.98838
Sd(LSE)	0.15955	0.13292	0.02452	0.09578	0.04970
$\hat{\beta}$(NEW)	0.85289	1.81854	2.98145	4.03356	4.98814
Sd(NEW)	0.01130	0.00988	0.00550	0.01196	0.00538

$\hat{\beta}$(LSE)是利用原始資料所估計出來的

$\hat{\beta}$(NEW)是利用原始資料與重新抽取資料所估計出來的

由表二，可以看出$\underset{\sim}{\beta}$參數值的估計上沒有很明顯的差異，但是在標準差來看，很明顯的看出，新方法的變異比最小平方法來的小很多，由此可見，新方法估計出來的結果比較接近原始參數值。

第柒章　結論

本文主要探討時間序列線性迴歸模型，在新方法與最小平方法估計結果上的差異。在第貳章我們先介紹了 Georgiana Popovici (2010)指數自迴歸時間序列，和 Miroslav M.Ristic(2005) Beta-Gamma 自迴歸時間序列的式子推導，並且加以推廣至兩個時間點。第三章進行篩選，解決資料在估計式上的問題。第四章我們利用了 Monte Carlo Method 和 Newton Raphson Method 推導出 $\underset{\sim}{\beta}$ 參數值的估計式。第伍章比較資料分析結果。

由模擬結果顯示，在時間序列迴歸上，新方法所估計出來的 $\underset{\sim}{\beta}$ 參數值比最小平方法所估計出來的 $\underset{\sim}{\beta}$ 參數值來的更加接近原始參數值。針對時間序列迴歸，未來可以考慮在不同的模型下，不同實際資料選擇或應用更合理的模型，進一步推廣本文所提出的方法。

參考文獻

1. Anderson, E. C. (1999). Monte Carlo Method and Importance Sampling.*Lecture Notes for Stat 578C Statistical Genetics*.
2. Gaver, D. P. and Lewis, P. A. W. (1980). First-Order Autoregressive Gamma Sequences and Point Processes.*Adv. Appl. Prob.* 12, *pp.* 727-745
3. Jorgensen, B. (1987). Exponential Dispersion Models. *J.R. Statist. Soc.* B 49(2), *pp.*127-162
4. Jorgensen, B. (1998). Stationary Time Series Model with Exponential Dispersion Model Margins.*J. Appl. Prob.* 35, *pp*.78-92
5. Joe, H. (1996). Time Series Models with Univariate Margins in the Convolution-Closed Infinitely Divisible Class. *J. Appl. Prob.* 33, *pp.* 664-677
6. Ristic, M. M. (2005). A Beta-Gamma Autoregressive Process of the Second-Order(BGAR(2)). *Statistic & Probability Letters.* 73,*pp*.403-410
7. Popovici, G. (2010). On Three Classes of Time Series Involving Exponential Distribution.*Math. Reports* 12(62), 1,*pp.* 45-57
8. Song, P. X.- K. and Feng, D. (2001). On Parameter Estimation for Exponential Dispersion Arma Models.*Journal of Time Series Analysis Vol.* 26(6) ,*pp.*843-862