

國立政治大學應用數學系
碩士學位論文

區間迴歸與模糊資訊分析及應用
Interval regression analysis with fuzzy data



碩士班學生：蔡皓旭 撰
指導教授：陳政輝 博士

中華民國 105年 6月16日

國立政治大學應用數學系
蔡皓旭 君所撰之碩士學位論文

區間迴歸與模糊資訊分析及應用

Interval regression analysis with fuzzy data

業經本委員會審議通過
論文考試委員會委員：

指導教授：

系主任：

中華民國 105年6月16日

摘要

動機與目的：傳統的統計迴歸模式假設觀測值的不確定性來自於隨機現象，而模糊迴歸則考慮不確定性來自於多重隸屬現象。不同的模型建構所得到的估計值也不一致。如何衡量模型的優劣程度，至今仍沒有一套嚴謹的標準。

研究方法：本研究以區間模糊數建構模糊迴歸模式，如此一來對樣本的解釋方式將更為貼近現實，並提出一套區間模糊數距離測度，以衡量估計值與實際值之間的差距。實證分析中（懸浮微粒 PM_{10} 濃度預測、台灣加權股價指數預測），我們藉由此距離測度衡量二維模糊迴歸與傳統二項最小平方法對於樣本的配適性。

創新與推廣：提出區間模糊數距離衡量估計值與原樣本之差異程度。在符合傳統統計迴歸精神之下，當距離最小就是差異最小的估計，最能符合所抽取的樣本，也是最佳估計。

重要發現：利用本區間模糊數距離測度，我們發現二維模糊迴歸方法比起傳統二項最小平方法更有效率且廣義殘差（generalized residual）將更小。

結論：過去以來，我們對於模糊迴歸架構一直都沒有完整的衡量標準。文中我們定義區間模糊數區間距離與平均距離，並推導賦距空間等性質。結合實例分析及應用，建構一合適模糊迴歸模式，以利統計決策分析參考。

關鍵字：區間模糊迴歸、懸浮微粒 PM_{10} 、台灣加權股價指數、區間模糊數

Abstract

Objective: This study concerns how to develop effective fuzzy regression models. In the literature, little is addressed on how to evaluate the effectiveness of fuzzy regression models developed with different regression methods. We consider this issue in this work and present a framework for such evaluation.

Method: We consider fuzzy regression models developed with different regression approaches. A method to evaluate the developed models is proposed. We then show that the proposed method possesses desirable mathematical properties and it is applied to compare the two-dimensional regression method and the traditional least square based regression method in our case studies: predicating the concentration of PM_{10} and the volatility of the weighted price index of the Taiwanese stock exchange.

Innovation: We propose a new metric to define a distance between two fuzzy numbers. This metric can be used to evaluate the performance of different fuzzy regression models. When a prediction from one model is closest to the sample data measured in terms of the proposed metric, it can be recognized as the optimal predication.

Results: Based on the proposed metric, it can be obtained that the two-dimensional fuzzy regression method is better than the traditional least square based regression method. Especially, its resulting generalized residual is smaller.

Conclusion: In the literature, no unified framework has been previously proposed in evaluating the effectiveness of developed fuzzy regression models. In this work, we present a metric to achieve this goal. It facilitates the work to determine whether a fuzzy regression model suitably fits obtained samples and whether the model has potential to provide sufficient accuracy for follow-up analysis in a considered problem.

Keyword: Fuzzy regression、 PM_{10} 、TAIEX、Interval fuzzy number



目錄

1 前言.....	1
2 研究方法.....	4
2.1 模糊迴歸模式架構.....	4
2.2 模糊迴歸參數估計.....	6
2.3 殘差分析.....	8
2.4 在財務金融上的應用.....	11
2.5 模糊區間距離與賦距空間.....	12
2.6 其它模糊迴歸方法.....	15
3 實證.....	19
3.1 懸浮微粒 PM_{10} 濃度預測.....	19
3.2 台灣加權股價指數價量分析.....	25
4 結論.....	30
5 參考文獻.....	32
6 附錄.....	36



1 前言

在大數據資料的分析和研究中，人們逐漸體會到大量資料並非就是大量的資訊，資料分析與萃取扮演重要的角色。雖然資料採礦的興起，解決了不少資料分析的問題。但是對於如何處理非實數樣本資料（如：區間模糊資料），傳統統計方法已無法有效的分析與掌控。

機率上的隨機性是代表著發生的不確定，由機率測度函數來表示；模糊理論的模糊性是隸屬程度的不確定，由隸屬函數表示 [3][13]。日常生活中，很多現象充滿著不確定性，而面對不確定性資訊，過去皆是以隨機的概念加以描述。直到西元 1960 年，人們開始質疑隨機的概念用在具有不確定性資訊的適用性。為了有效描述人文思維的不確定性，扎德 (Zadeh L.A.) 教授於 1965 年提出模糊集合理論 (Fuzzy Set Theory) 的概念。有別於傳統的布林邏輯，模糊理論以模糊邏輯為基礎，藉由隸屬度的概念較能反映出生活中實際的情況。到了 1970 年，Zadeh 與 Bellman 又區分出隨機(Randomness)與模糊(Fuzziness)之間的差異。

傳統迴歸分析假設觀察值是由實際應變數影響加上隨機誤差而產生的。亦即，應變數是一種帶有不確定性的隨機變數。然而，在很多的實務應用上，此假設無法成立。例如，有些應變數的觀察值並不是以單一數值的形式存在，而是帶有多重隸屬的特性。這些觀察值雖然具有不確定性，但這種不確定的特性是來自於多重隸屬現象，而非隨機現象。在處理具模糊性的樣本時，如果利用假性的「精確值」，可能誤導模型的建構，也可能擴大預測結果和實際狀態之間的誤差。傳統的數字資料有需求過度及過度解釋的危險，如果應用模糊的數值資料，較能避免這樣的危險發生。因此為了改變傳統統計資料的隨機觀點，以另一種角度來看待

不確定性的問題，我們考慮用模糊觀點來取代隨機觀點，而由模糊觀點套用迴歸分析來處理模糊資料問題，就稱為模糊迴歸分析[1]。

模糊迴歸最早是由 Tanaka, H. & Uejima, S. & Asai, K. (1980) 提出，文中將模糊理論與線性迴歸模型結合，提出具有模糊性的線性迴歸模型，並利用線性規劃來對模糊迴歸模型的參數做估計[14]。Diamond (1988) 利用傳統最小平方方法的觀念，求出一個模糊迴歸模式以估算應變數的觀察值，並利用應變數的觀察值與估計值的模糊數值之最小值、最可能值與最大值，定義兩者之間的距離，以衡量迴歸模式與模糊資料之間的最佳配適度情形[16]。導出類似傳統迴歸分析中的正規方程式(Normal Equation) 以求解迴歸係數。其後，Chang 與 Lee (1996) 則將 Diamond (1988)的方法加以擴充，將迴歸係數的最可能值與分展度 (Spread) 分別估計，該方法的特點在於可將極端值加以處理，並降低影響程度。Wu 與 Tseng (2002)提出改善模糊迴歸模式之最小平方方法，其模式主要觀念是利用最小平方方法將所有樣本的左端點及右端點各自配置一條迴歸線，然後再將其左右迴歸線統合成一條迴歸線[12]。

對於模糊迴歸參數的估計方法有很多，但是如何衡量模型對樣本的配適性，一直以來都沒有一個嚴謹的標準[24][34]。文中我們以區間模糊數距離衡量估計與實際之間的差異程度。我們除了定義區間模糊數區間距離與平均距離外，並推導出賦距空間等性質，使得距離定義更加嚴謹。此一衡量模型對樣本的配適性的距離定義，亦為本文最重要的貢獻。

近年來模糊統計與模糊相關性日漸受到重視，應用的範圍非常廣泛，從工程科技到社會人文科學都可以發現模糊理論研究的蹤跡與成果。本文將以區間模糊數建構模糊迴歸模式，如此一來對樣本的解釋方式將更貼近現實。並將模型應用在台灣加權股價指數的波動預測與懸浮微粒 (Suspended Particulate Matter；簡稱

PM) 濃度的變化，並以本文所定義的距離來衡量模型與樣本的配適性。

在財務金融等相關文獻中，「價量關係」經常被拿出來討論，量是價的先行指標，亦即「量」與「價」關係的配合好壞，會對股價的上漲或下跌形成重要的影響，因為成交量代表了有多少的籌碼被交換，而價格所代表的成本則明白揭示出該批籌碼交換的價位。藉由價量來觀察台灣加權股價指數，不能單以一兩天的成交量衡量，需要長期的資料作分析。對於長期價量資訊，若以單一數值表示，將無法詮釋樣本的模糊性。甚至會因為取決的頻率而影響估計值的結果[4]。

懸浮微粒是指懸浮在空氣中之固、液態微小粒子，粒徑大小從 0.001 至 500 微米 (μm) 都有可能，但大部分在 0.1~10 μm 之間。主要來源分為原發性污染物與衍生性污染物。懸浮微粒為原發性污染物是由污染源排放至大氣中，例如：大自然風吹塵土、工廠或垃圾燃燒、汽機車排放等。衍生性污染物是大氣中經由光化學反應產生的二次衍生性氣膠，主要參與作用的有硫化物 SO_x 、氮化物 NO_x 、揮發性有機物 VOC_s 、氨 NH_3 等。這些懸浮微粒有的本身就是有害物質，有的會吸附溶解其他有害物，如水霧吸收。當含有有害物質的微粒經呼吸作用進入人體後，便可能造成呼吸系統疾病，或是經循環系統輸送至體內其他器官，造成各種類型的病變。台灣十大死因中，除了意外事故，其餘九項都與空汙有關[11]。近年來，許多流行病學研究指出，長期暴露於懸浮微粒濃度高的環境中，會增加罹患呼吸與心血管疾病的機率，亦可能間接增加死亡的風險[8]。因此懸浮微粒對人體之危害逐漸受到重視。然而污染物的濃度，隨時都會發生變化，本研究藉由模糊的概念，將濃度以區間範圍的形式呈現，以保留樣本的模糊性[8]。

2 研究方法

本章先在 2.1 模糊迴歸模式架構中，以區間模糊數建構模糊迴歸模式。而在 2.2 模糊迴歸參數估計中，採用中心與分展度二維度技術，並藉由最小平方法 (Least Squares Estimation) 來進行參數估計。2.3 殘差分析 (Residual Analysis) 中，結合傳統迴歸殘差分析概念，以距離大小來呈現兩區間的差異程度。並提出區間數區間距離與平均距離的概念。2.4 在財務金融上的應用中，介紹波動率的特性，及模糊迴歸在財務金融上的相關文獻。2.5 模糊區間距離與賦距空間，驗證區間模糊數平均距離具有賦距空間等性質。2.6 其它模糊迴歸方法，介紹其它模糊迴歸參數估計方法與區間數加法和乘法。

2.1 模糊迴歸模式架構

迴歸分析預測法，具有描述、預測與控制獨立變數和應變數之間因果關係的功能，透過建立變數之間的迴歸方程，可將迴歸方程作為預測模型。因此，迴歸分析預測是一種重要的市場預測方法。

傳統迴歸分析假設觀察值是由實際應變數影響加上隨機誤差而產生的。亦即，應變數是一種帶有不確定性的隨機變數。不過在很多實務應用上，此假設無法成立。以匯率為例，統計數字只能表示單一時間點的資訊，不能完全準確地反應時間區間下匯率變化。如果利用假性的「精確值」，可能誤導模型的建構，也可能擴大預測結果和實際狀態之間的誤差。本研究中，將以區間模糊數建構模糊迴歸模式，如此一來對樣本的解釋方式將更貼近現實。

傳統的統計迴歸模式假設觀測值的不確定性來自於隨機現象，而模糊迴歸則考慮不確定性來自於多重隸屬現象。當考慮具有模糊性資料時，參數本身具有不

確定性與模糊性，線性模糊迴歸常表示成：

$$Y(\mathbf{X}_i) = \mathbf{A}_0 + \mathbf{A}_1 X_{i1} + \mathbf{A}_2 X_{i2} + \cdots + \mathbf{A}_p X_{ip} \quad (2.1)$$

$X_i = (X_{i1}, X_{i2}, \dots, X_{ip})'$ 為獨立變數向量

向量元素 X_{im} 皆為區間模糊數， $m = 1, 2 \dots p$

A_k 為區間模糊參數， $k = 1, 2 \dots p$

$Y(X_i)$ 為區間模糊應變數



2.2 模糊迴歸參數估計

我們考慮應用最小平方法 (Least Squares Estimation) 來進行參數估計，由於 $Y(X_i)$ ， X_{im} ， A_K 為區間模糊數，因此採用中心點半徑二維度之技術，將 $Y(X_i)$ ， X_{im} ， A_K 化為中心半徑形式，表示如下：

$$Y(X_i) = [Y_C(X_i) - Y_S(X_i), Y_C(X_i) + Y_S(X_i)] = \langle Y_C(X_i), Y_S(X_i) \rangle \quad (2.2)$$

$$X_{im} = [C_{im} - S_{im}, C_{im} + S_{im}] = \langle C_{im}, S_{im} \rangle \quad (2.3)$$

$$A_k = [C_{A_k} - S_{A_k}, C_{A_k} + S_{A_k}] = \langle C_{A_k}, S_{A_k} \rangle \quad (2.4)$$

並將方程式(2.1)，改寫如下：

$$\begin{aligned} Y(X_i) &= A_0 + A_1 X_{i1} + A_2 X_{i2} + \cdots + A_p X_{ip} \\ &= \langle C_{A_0}, S_{A_0} \rangle + \langle C_{A_1}, S_{A_1} \rangle \langle C_{i1}, S_{i1} \rangle + \langle C_{A_2}, S_{A_2} \rangle \langle C_{i2}, S_{i2} \rangle + \cdots \\ &+ \langle C_{A_p}, S_{A_p} \rangle \langle C_{ip}, S_{ip} \rangle = \langle Y_C(X_i), Y_S(X_i) \rangle \end{aligned} \quad (2.5)$$

並藉由二維度技術，將模糊迴歸模式(2.2)改寫成中心迴歸模式與分展度迴歸模式如下：

中心迴歸模式

$$Y_C(X_i) = C_{A_0} + C_{A_1} C_{i1} + C_{A_2} C_{i2} + \cdots + C_{A_p} C_{ip} \quad (2.6)$$

半徑迴歸模式

$$Y_S(X_i) = S_{A_0} + S_{A_1}S_{i1} + S_{A_2}S_{i2} + \cdots + S_{A_p}S_{ip} \quad (2.7)$$

藉由中心迴歸模式(2.6)與半徑迴歸模式(2.7)分別預測出模糊應變數 $Y(X_i)$ 的應變數中心 $Y_C(X_i)$ 與應變數半徑 $Y_S(X_i)$ 。

下一節中，我們將提出以距離衡量估計值與實際值之間的誤差程度，並定義嚴謹的區間數距離測度，以定義的區間數距離測度來衡量樣本與模型的配適性。



2.3 殘差分析

傳統實數值線性迴歸分析之假設 (Assumptions) 如下：

1. 應變數 (Response Variable) 是常態分佈 (Normal Distribution) ，且來自隨機取樣(Random Sampling)
2. 獨立變數 (Independent Variable) 是固定的常數或實驗數據，且其量測沒有誤差
3. 任意兩獨立變數是獨立事件 (Independent Event)
4. 應變數之誤差值 (Error) 在母群體是常態分佈，變異數是一個常數，且平均值是 0

傳統統計中，殘差是指觀察值與預測值之間的差距，其定義如下式：

$$e = Y_i - \hat{Y}_i \quad (2.8)$$

在區間模糊迴歸模式下，觀察值與預測值皆為區間數。為保有傳統統計精神，我們以距離大小來呈現兩個區間之間的差異程度，其定義如下：

定義 2.1 區間模糊數區間距離

假設 $\tilde{A} = [A_1, A_2]$ ， $\tilde{B} = [B_1, B_2]$ 為兩區間模糊數，則我們定區間模糊數區間距離 d 如下：

$$d(\tilde{A}, \tilde{B}) = \langle L_{AB}, M_{AB}; C_{AB} \rangle \quad (2.9)$$

其中 $L_{AB} = \max\{A_1 - B_2, B_1 - A_2, 0\}$ 為兩區間最小距離

$M_{AB} = \max\{A_2 - B_1, B_2 - A_1\}$ 為兩區間最大距離

$C_{AB} = -\min\{\max[A_1 - B_2, B_1 - A_2, B_1 - B_2, A_1 - A_2], 0\}$ 為兩區間重疊距離

由定義 2.1 區間模糊數區間距離，兩區間的距離並非單一實數值，而是最小距離與最大距離的區間數。類似的概念，就好比是各行星與太陽之間的距離，隨著時間週期變化。離太陽最近時，稱為近日點；離太陽最遠時，稱為遠日點。西元 1596 年，天文學家克卜勒曾寫過一本書〈世界的和諧〉，書中敘述天球音階的和諧，以及天體變化與人體的關係。書中有一頁紀載克卜勒第三定律的內容—行星平均軌道半徑的三次方與其公轉週期的平方比值是固定的。藉由平均軌道半徑的概念，我們可將區間距離（區間數）轉換成平均距離（實數）。

定義 2.2 區間模糊數平均距離

假設 $\tilde{A} = [A_1, A_2]$ ， $\tilde{B} = [B_1, B_2]$ 為兩區間模糊數，則我們定區間模糊數平均距離 D 如下：

$$D(\tilde{A}, \tilde{B}) = f \circ d(\tilde{A}, \tilde{B}) = f(\langle L_{AB}, M_{AB}; C_{AB} \rangle) = \frac{L_{AB} + M_{AB} - C_{AB}}{2} \quad (2.10)$$

其中 D 為 f 、 d 的合成函數

$$f: \mathbb{R} \rightarrow [0, \infty), \quad d: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}, \quad D: \mathbb{R} \times \mathbb{R} \rightarrow [0, \infty)$$

\mathbb{R} 表示實數區間數集合

$L_{AB} = \max\{A_1 - B_2, B_1 - A_2, 0\}$ 為兩區間最小距離

$M_{AB} = \max\{A_2 - B_1, B_2 - A_1\}$ 為兩區間最大距離

$C_{AB} = -\min\{\max[A_1 - B_2, B_1 - A_2, B_1 - B_2, A_1 - A_2], 0\}$ 為兩區間重疊距離

例 1 假設 $\tilde{A} = [4, 10]$ ， $\tilde{B} = [16, 20]$ 為兩區間模糊數，則區間距離、平均距離為

$$L_{AB} = \max\{4 - 20, 16 - 10, 0\} = 6$$

$$M_{AB} = \max\{10 - 16, 20 - 4\} = 16$$

$$C_{AB} = -\min\{\max[4 - 20, 16 - 10, 16 - 20, 4 - 10], 0\} = 0$$

$$D(\tilde{A}, \tilde{B}) = f \circ d(\tilde{A}, \tilde{B}) = \frac{6 + 16 - 0}{2} = 11$$

$$\text{區間距離 } d(\tilde{A}, \tilde{B}) = \langle 6, 16; 0 \rangle \quad \text{平均距離 } D(\tilde{A}, \tilde{B}) = 11$$

例 2： $\tilde{A} = [3, 10]$ ， $\tilde{B} = [5, 20]$ 為兩區間模糊數，則區間距離、平均距離為

$$L_{AB} = \max\{3 - 20, 5 - 10, 0\} = 0$$

$$M_{AB} = \max\{10 - 5, 20 - 3\} = 17$$

$$C_{AB} = -\min\{\max[3 - 20, 5 - 10, 5 - 20, 3 - 10], 0\} = 5$$

$$D(\tilde{A}, \tilde{B}) = f \circ d(\tilde{A}, \tilde{B}) = \frac{0 + 17 - 5}{2} = 6$$

$$\text{區間距離 } d(\tilde{A}, \tilde{B}) = \langle 0, 17; 5 \rangle \quad \text{平均距離 } D(\tilde{A}, \tilde{B}) = 6$$

由以上距離的定義，若估計區間與實際區間的平均距離 $D(Y_i, \hat{Y}_i)$ 越大，表示估計出來的模糊應變數 \hat{Y}_i 與原本樣本 Y_i 差距越大。在符合傳統統計迴歸精神下，當距離越小就是差異最小的估計，最能符合所抽取的樣本，也是最佳估計。

2.4 在財務金融上的應用

在金融領域中，波動率 (Volatility) 常被用來量化資產的風險程度。估計波動率的模型，大約分為三種。最簡單的模型是由該標的過去的指數或價格等歷史訊息，所計算得到的數據，稱歷史波動率 (Historical Volatility)，這是屬於無條件的波動率；ARCH 及其延伸模型 (GARCH、EGARCH、IGARCH、GARCH-M 等) 則是捕捉報酬叢聚性的特徵，屬於條件波動率；而隱含波動率 (Implied Volatility) 是將選擇權市價代入理論評價模型中，所計算出來的波動率。

隱含波動率所代表的是投資人對標的物未來不確定性的意見和看法，且反應了其它不在定價公式中的因子。例如流動性、買賣價差、錯誤資料、稅負、交易成本等。因此，一般認為相較於其它波動率模型而言，隱含波動率隱含著較多的訊息，所以對於預測未來的真實波動率 (Realized Volatility) 而言，隱含波動率是個預測市場波動較好的代理變數 (Proxy)。然而，Huang (2002) 綜合一些研究指出隱含波動率有一些衡量的偏誤。因不同模型各有其優缺點，所以模型的爭論一直都存在，也各有學者支持。

隨著模糊統計與模糊相關性日漸受到重視，近年來也有許多以模糊迴歸預測價格市場的研究，如：Reza Ghodsi, Mohammadsaleh Zakerinia, Mahdi Joka (2010) 等由差分整合移動平均自迴歸模型 (ARIMA)、人工神經網絡 (ANN)、模糊迴歸 (FR) 三種方法預測電力市場價值。研究中，FR 與 ANN 預測結果較佳[30]；Ghodratollah Emamverdi, Ebrahim Siami Araghi, Fatemeh Fahimifar (2013) 結合 ARIMA 和 FR 兩種方法預測股票市場價值[32]。

2.5 模糊區間距離與賦距空間

賦距空間是指一集合上可定義距離的概念。距離是指集合上定義一函數，集合中任兩元素可依此賦予一函數值，且此函數滿足某些性質，更具體地說：設 X 為一集合，若存在一非負函數 d

$$d: X \times X \rightarrow [0, \infty)$$

滿足

(a) $d(x, y) \geq 0 \quad \forall x, y \in X$

(b) $d(x, y) = 0 \Leftrightarrow x = y$

(c) $d(x, y) = d(y, x)$

(d) 三角不等式： $d(x, y) \leq d(x, z) + d(z, y) \quad \forall x, y, z \in X$

我們稱 (X, d) 為一賦距空間 (Metric Space)， d 為 X 上的一個度規 (Metric)。

本章節中，我們將證明若以模糊區間距離測度 D 定義距離，則區間模糊數可視為一賦距空間。

設 $\tilde{A}, \tilde{B}, \tilde{C} \in \mathbb{IR}$ 為三個區間模糊數。 $\tilde{A} = [A_1, A_2]$ ， $\tilde{B} = [B_1, B_2]$ ， $\tilde{C} = [C_1, C_2]$

$D: \mathbb{IR} \times \mathbb{IR} \rightarrow [0, \infty)$ 為區間數集合 \mathbb{IR} 上的一個測度。

$$D(\tilde{A}, \tilde{B}) = f \circ d(\tilde{A}, \tilde{B}) = \frac{L_{AB} + M_{AB} - C_{AB}}{2}$$

其中 $d: \mathbb{IR} \times \mathbb{IR} \rightarrow \mathbb{IR}$ 定義如下：

$$d(\tilde{A}, \tilde{B}) = \langle L_{AB}, M_{AB}, C_{AB} \rangle$$

其中 $f: \mathbb{IR} \rightarrow \mathbb{R}$ 定義如下：

$$f(\langle L_{AB}, M_{AB}, C_{AB} \rangle) = \frac{L_{AB} + M_{AB} - C_{AB}}{2}$$

L_{AB}, M_{AB}, C_{AB} 定義如下：

$$M_{AB} = \max\{A_2 - B_1, B_2 - A_1\}$$

$$L_{AB} = \max\{A_1 - B_2, B_1 - A_2, 0\}$$

$$C_{AB} = -\min\{\max[A_1 - B_2, B_1 - A_2, B_1 - B_2, A_1 - A_2], 0\}$$

Lemma 1. $D(\tilde{A}, \tilde{B}) \geq 0 \quad \forall \tilde{A}, \tilde{B} \in \mathbf{IR}$

<pf> **Case 1 :** Suppose that $\tilde{A} \cap \tilde{B} = \emptyset$.

This implies that $L_{AB} \geq 0$, $M_{AB} \geq 0$ and $C_{AB} = 0$.

Therefore, $\frac{L_{AB} + M_{AB} - C_{AB}}{2} \geq 0$.

Case 2 : Suppose that $\tilde{A} \cap \tilde{B} \neq \emptyset$.

This implies that $L_{AB} = 0$ and $M_{AB} \geq 0 \wedge M_{AB} \geq C_{AB}$.

Therefore, $\frac{L_{AB} + M_{AB} - C_{AB}}{2} \geq 0$.

Lemma 2. $D(\tilde{A}, \tilde{B}) = 0 \Leftrightarrow \tilde{A} = \tilde{B}$

<pf> (\Leftarrow) Suppose that $\tilde{A} = \tilde{B}$.

This implies that

$$L_{AB} = 0, \quad M_{AB} = A_2 - A_1 = B_2 - B_1 \quad \text{and} \quad C_{AB} = A_2 - A_1 = B_2 - B_1.$$

$$\text{Thus, } D(\tilde{A}, \tilde{B}) = \frac{L_{AB} + M_{AB} - C_{AB}}{2} = 0.$$

(\Rightarrow) Suppose that $D(\tilde{A}, \tilde{B}) = 0$

This implies that

$$L_{AB} = 0, \quad M_{AB} = A_2 - A_1 = B_2 - B_1 \quad \text{and} \quad C_{AB} = A_2 - A_1 = B_2 -$$

B_1

Thus, $\tilde{A} = \tilde{B}$.

Lemma 3. $D(\tilde{A}, \tilde{B}) = D(\tilde{B}, \tilde{A})$

<pf> Since $d(\tilde{A}, \tilde{B}) = \langle L_{AB}, M_{AB}, C_{AB} \rangle = d(\tilde{B}, \tilde{A})$,

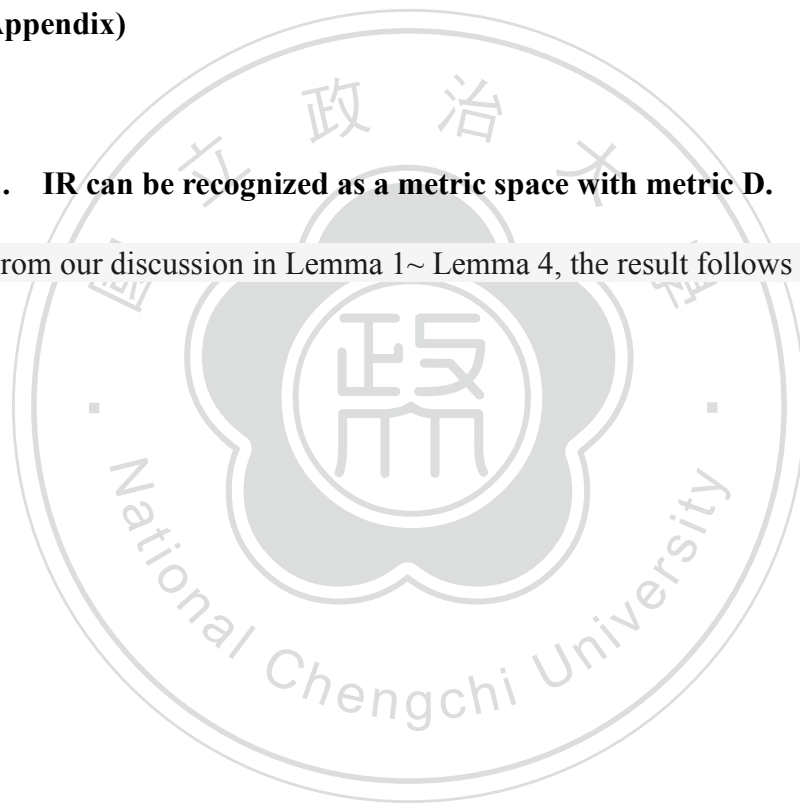
we have $D(\tilde{A}, \tilde{B}) = \frac{L_{AB} + M_{AB} - C_{AB}}{2} = D(\tilde{B}, \tilde{A})$. (by definition)

Lemma 4. Triangle inequality : $D(\tilde{A}, \tilde{C}) \leq D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \quad \forall \tilde{A}, \tilde{B}, \tilde{C} \in \mathbf{IR}$

(See Appendix)

Theorem 1. \mathbf{IR} can be recognized as a metric space with metric D .

<pf> From our discussion in Lemma 1~ Lemma 4, the result follows directly.



2.6 其它模糊迴歸方法

估計模糊迴歸參數的方法大略可分為兩種，一種是線性規劃，一種是最小平方法。線性規劃，主要是尋求一組能將實數值樣本或模糊數樣本全包括的區間，再用三角形隸屬函數套用在此區間上。最小平方法，除了二維度技術（中心、半徑）外，Wu 與 Tseng (2002) 提出改善模糊迴歸模式之最小平方法，其模式主要觀念是利用最小平方法將所有樣本的左端點及右端點各自配置一條迴歸線，然後再將其左右迴歸線統合成一條迴歸線[12]。陳孝煒、吳柏林 (2007) 所撰寫的區間迴歸與模糊樣本分析，文中提出中心與半徑迴歸帶入模糊參數，並藉由區間模糊數乘法與區間數加法定義，計算出模糊區間應變數[2]。

區間模糊數乘法與加法定義如下：

定義 2.3 區間模糊數乘法 1

假設 $\tilde{A} = \langle C_1, S_1 \rangle$ ， $\tilde{B} = \langle C_2, S_2 \rangle$ 為兩區間模糊數，則我們定義區間模糊數乘法 1 如下：

$$\tilde{A} \times \tilde{B} = \langle C_1, S_1 \rangle \times \langle C_2, S_2 \rangle = \langle C_1 \times C_2, \min\{S_1, S_2\} \rangle \quad (2.11)$$

（見 陳孝煒、吳柏林 (2007)，區間迴歸與模糊樣本分析）[2]

定義 2.4 區間模糊數乘法 2

假設 $\tilde{A} = \langle C_1, S_1 \rangle$ ， $\tilde{B} = \langle C_2, S_2 \rangle$ 為兩區間模糊數，則我們定義區間模糊數乘法 2 如下：

$$\tilde{A} \times \tilde{B} = \langle C_1, S_1 \rangle \times \langle C_2, S_2 \rangle = \langle C_1 \times C_2, \max\{S_1, S_2\} \rangle \quad (2.12)$$

（見 陳孝煒、吳柏林 (2007)，區間迴歸與模糊樣本分析）[2]

定義 2.5 區間模糊數乘法 3

假設 $\tilde{A} = [A_1, A_2]$, $\tilde{B} = [B_1, B_2]$ 為兩區間模糊數，則我們定義區間模糊數乘法 3 如下：

$$\begin{aligned}\tilde{A} \times \tilde{B} &= [A_1, A_2] \times [B_1, B_2] \\ &= [\min\{A_1B_1, A_2B_1, A_1B_2, A_2B_2\}, \max\{A_1B_1, A_2B_1, A_1B_2, A_2B_2\}] \quad (2.13)\end{aligned}$$

(見 吳柏林 (2015), 模糊統計導論—方法與應用, 五南出版社) [1]

定義 2.6 區間模糊數加法

假設 $\tilde{A} = [A_1, A_2]$, $\tilde{B} = [B_1, B_2]$ 為兩區間模糊數，則我們定義區間模糊數加法如下：

$$\tilde{A} + \tilde{B} = [A_1, A_2] + [B_1, B_2] = [A_1 + B_1, A_2 + B_2] \quad (2.14)$$

陳孝煒、吳柏林 (2007) 所撰寫的區間迴歸與模糊樣本分析，文中藉由 2.3 ~ 2.5 三種不同的乘法定義，得到不同的區間模糊數相乘結果。我們在估計 $Y_C(X_i)$, $Y_S(X_i)$ 過程中，會將此三種定義分別使用在估計上，預期得到三種估計值 $\widehat{Y(X_i)}_{2.2}$, $\widehat{Y(X_i)}_{2.3}$, $\widehat{Y(X_i)}_{2.4}$, 再分別用之與原本的 $Y(X_i)$ 計算其距離，從中獲得較優的估計解釋。

例 我們以 81~95 年度歲入歲出淨額為例，探討收入與支出間的關聯性，建構迴歸模式如下：

$$Y_{\text{支出}} = A_0 + A_1 X_{\text{收入}} \quad (2.15)$$

其中 $Y_{\text{支出}}$ 為年支出

$X_{\text{收入}}$ 為年收入

並估計模糊迴歸參數 A_0 、 A_1 。首先，我們以二維度技術建立中心與半徑迴歸模式

如下：

中心迴歸模式：

$$C_{\text{支出}} = -13.50 + 1.269C_{\text{收入}} \quad (2.16)$$

半徑迴歸模式：

$$S_{\text{支出}} = -1.613 + 1.23S_{\text{收入}} \quad (2.17)$$

將迴歸參數代入模糊迴歸模式(2.15)式，改寫如下：

$$\begin{aligned} Y_{\text{支出}} &= \langle -13.5, -1.613 \rangle + \langle 1.269, 1.23 \rangle X_{\text{收入}} \\ &= [-15.11, -11.89] + [0.04, 2.50] X_{\text{收入}} \end{aligned} \quad (2.18)$$

接著分別使用乘法定義 2.3~2.5 求得三種估計值，並計算估計值與實際值的區間距離與平均距離：

(1) 乘法定義 2.3 之估計值與實際值距離加總

區間距離為： $\langle 5.07, 245.64; 55.49 \rangle$

平均距離為：97.61

(2) 乘法定義 2.4 之估計值與實際值距離加總

區間距離為： $\langle 0, 345.19; 162.62 \rangle$

平均距離為：91.29

(3) 乘法定義 2.5 之估計值與實際值距離加總

區間距離為： $\langle 0, 2220.87; 234 \rangle$

平均距離為：993.44

由上述結果可知，定義 2.4 之估計值與實際值區間平均距離加總最小，因此樣本的最佳區間相乘定義為乘法定義 2.4

模糊迴歸模型為

$$Y_{\text{支出}} = [-15.11, -11.89] + [0.04, 2.50]X_{\text{收入}}$$

表 1. 81~95 年度歲入歲出淨額 (單位:百億元) $n = 15$

年	Y(年支出)	Y(中心,半徑)	X(年收入)	X(中心,半徑)
81	[42,48]	<45,3>	[40,46]	<43,3>
82	[48,50]	<49,1>	[45,53]	<49,4>
83	[49,55]	<52,3>	[46,54]	<50,4>
84	[50,62]	<56,6>	[51,59]	<55,4>
85	[54,64]	<59,5>	[56,64]	<60,4>
86	[62,72]	<66,6>	[58,76]	<67,9>
87	[78,110]	<94,16>	[68,92]	<80,12>
88	[73,109]	<91,18>	[79,109]	<94,15>
89	[118,128]	<120,8>	[92,104]	<98,6>
90	[116,136]	<126,10>	[111,125]	<118,7>
91	[130,172]	<151,21>	[105,141]	<123,18>
92	[159,179]	<169,10>	[126,150]	<138,12>
93	[175,191]	<183,8>	[144,160]	<152,8>
94	[181,185]	<183,2>	[148,160]	<154,6>
95	[183,191]	<187,4>	[158,170]	<164,6>

3 實證

3.1 懸浮微粒PM₁₀濃度預測

近年來，中國大陸霾害因地球暖化關係，有越來越嚴重情況，霾害不只對大陸產生影響，台灣也深受其害；台灣因為冬天吹東北季風，所以每年入冬之後，來自蒙古與中國北方的亞洲沙塵暴經過中國沿海工業區，使得存在自然環境中的大小顆粒與工業污染物也跟著吹過來，導致空氣品質受到影響，整個台灣上空都是灰濛濛的一片[5]。許多學者也紛紛探討懸浮微粒對於健康的危害，以及對於懸浮微粒PM₁₀濃度的預測。Ul-Saufie A.Z & Yahya A.S & Ramli N.A (2010) 等人，藉由氣象因子（濕度、溫度、風速）和化學因子（CO, NO₂, SO₂, PM₁₀）建立 PCR 模型，預測PM₁₀濃度[31]。溫志忠、蔡立宏、楊尚威 (2009) 利用統計分析發現CO, NO₂, SO₂, PM₁₀彼此間具高相關性，與能見度則呈現負相關性[6]。Branis and Vetvicka (2010); Barmpadimos et al. (2011)，文中指出溫度和濕度對於PM₁₀皆呈現負相關。風速（風向）扮演著運輸與稀釋PM₁₀的角色；同時也扮演著土壤風塵的製造者。蔡春進等 (1999) 指出揚塵所需之風速起始值約為 5m/s，當超過 6m/s 之後揚塵量與風速呈現正相關，但是當風速超過某一上限，則呈負相關[7]。

本研究文中探討的氣象因子為溫度、濕度、風速；化學因子為CO, NO₂, SO₂, PM₁₀, O₃。我們蒐集士林地區 104 年 10 月~105 年 4 月的因子變化資料，建構模糊迴歸架構，其架構如下：

$$Y_{PM_{10},t} = A_0 + A_1X_{PM_{10},t-1} + A_2X_{濕度,t-1} + A_3X_{溫度,t-1} + A_4X_{NO_x,t-1} + A_5X_{SO_2,t-1} + A_6X_{CO,t-1} + A_7X_{O_3,t-1} + A_8X_{風速,t-1} \quad (3.1)$$

$Y_{PM_{10},t}$ 為今日 24 小時 PM_{10} 濃度的變化，以區間模糊數表示

$X_{PM_{10},t-1}$ 為昨日 24 小時 PM_{10} 濃度的變化，以區間模糊數表示

$X_{濕度,t-1}$ 為昨日 24 小時濕度的變化，以區間模糊數表示

$X_{溫度,t-1}$ 為昨日 24 小時溫度的變化，以區間模糊數表示

$X_{NO_x,t-1}$ 為昨日 24 小時 NO_2 濃度的變化，以區間模糊數表示

$X_{SO_2,t-1}$ 為昨日 24 小時 SO_2 濃度的變化，以區間模糊數表示

$X_{CO,t-1}$ 為昨日 24 小時 CO 濃度的變化，以區間模糊數表示

$X_{O_3,t-1}$ 為昨日 24 小時 O_3 濃度的變化，以區間模糊數表示

$X_{風速,t-1}$ 為昨日 24 小時風速的變化，以區間模糊數表示

採用二維模糊迴歸方法，估算出中心迴歸模式與半徑迴歸模式如下：

(1) 中心迴歸模式：

$$Y_{PM_{10},t,C} = 46.9 + 0.22C_{PM_{10},t-1} - 0.28C_{濕度,t-1} - 0.21C_{溫度,t-1} - 0.43C_{NO_x,t-1} + 0.14C_{SO_2,t-1} + 34.4C_{CO,t-1} + 0.27C_{O_3,t-1} - 4.79C_{風速,t-1} \quad (3.2)$$

(2) 半徑迴歸模式：

$$Y_{PM_{10},t,S} = 11.32 + 0.12S_{PM_{10},t-1} + 0.41S_{濕度,t-1} - 0.97S_{溫度,t-1} - 0.28S_{NO_x,t-1} + 0.55S_{SO_2,t-1} + 24.9S_{CO,t-1} + 0.27S_{O_3,t-1} - 0.76S_{風速,t-1} \quad (3.3)$$

藉由中心迴歸模式與半徑迴歸模式，我們針對 105 年 1 月份 PM_{10} 的濃度進行估計，其結果如下：

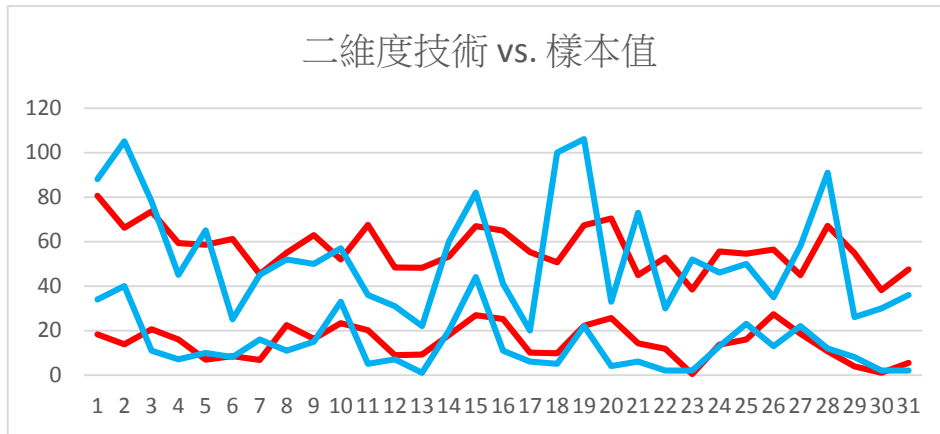


圖 3.1 二維度技術之估計值 vs. 樣本值 (105 年 1 月)

圖 3.1 中，紅色曲線範圍為模糊迴歸模式下， PM_{10} 估計濃度區間，藍色曲線為 PM_{10} 實際樣本濃度區間。

藉由定義 2.1 (區間模糊數區間距離) 和定義 2.2 (區間模糊數平均距離)，計算出 31 天估計值與樣本值的區間差距如下：

$$\begin{aligned} \sum_{i=1}^{31} D(Y_i, \hat{Y}_i) &= \sum_{i=1}^{31} f(< L_{Y_i \hat{Y}_i}, M_{Y_i \hat{Y}_i}; C_{Y_i \hat{Y}_i} >) = f(< 0, 1671.69; 863.64 >) \\ &= 404.03 \end{aligned} \quad (3.4)$$

由結果顯示，估計值與實際值 31 天距離加總，最近距離總和為 0；最遠距離總和為 1671.69；重疊距離總和為 863.64；平均距離總和為 404.03。

接著我們將比較 2.6 節 (其它模糊迴歸方法)，陳孝煒、吳柏林 (2007)，區間迴歸與模糊樣本分析中，所採用的方法。

首先，將中心迴歸參數與半徑迴歸參數代入 PM_{10} 模糊迴歸架構(3.1)中，並改寫如下：

(1) 中心與半徑迴歸模式：

$$Y_{PM_{10},t} = < 46.9, 11.32 > + < 0.22, 0.12 > X_{PM_{10},t-1} + < -0.28, 0.41 >$$

$$X_{溫度,t-1} + < -0.21, -0.97 > X_{溫度,t-1} + < -0.43, -0.28 > X_{NO_x,t-1} + <$$

$$0.14, 0.55 > X_{SO_2,t-1} + < 34.4, 24.9 > X_{CO,t-1} + < 0.27, 0.27 >$$

$$X_{O_3,t-1} + < -4.79, -0.76 > X_{\text{風速},t-1}$$

(3.5)

(2) 區間迴歸模式：

$$Y_{PM_{10},t} = [35.58, 58.22] + [0.1, 0.34]X_{PM_{10},t-1} + [-0.69, 0.13]X_{\text{濕度},t-1} + [-1.18, 0.77]X_{\text{溫度},t-1} + [-0.71, -0.14]X_{NO_x,t-1} + [-0.41, 0.69]X_{SO_2,t-1} + [9.5, 59.3]X_{CO,t-1} + [-0.002, 0.54]X_{O_3,t-1} + [-5.55, -4.03]X_{\text{風速},t-1} \quad (3.6)$$

藉由區間數乘法定義 2.3 至 2.5 求得三種估計值 $\widehat{Y(X_1)}_{2.3}$ 、 $\widehat{Y(X_1)}_{2.4}$ 、 $\widehat{Y(X_1)}_{2.5}$ ，其結果如下：

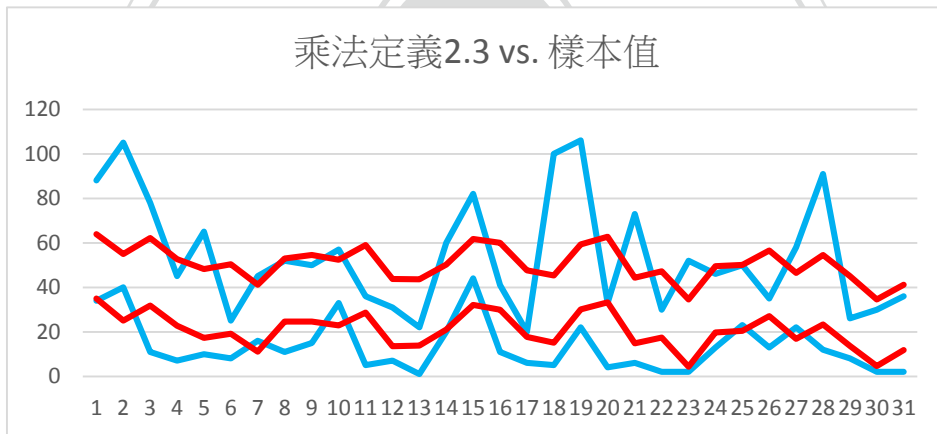


圖 3.2 乘法定義 2.3 之估計值 vs. 樣本值 (105 年 1 月)

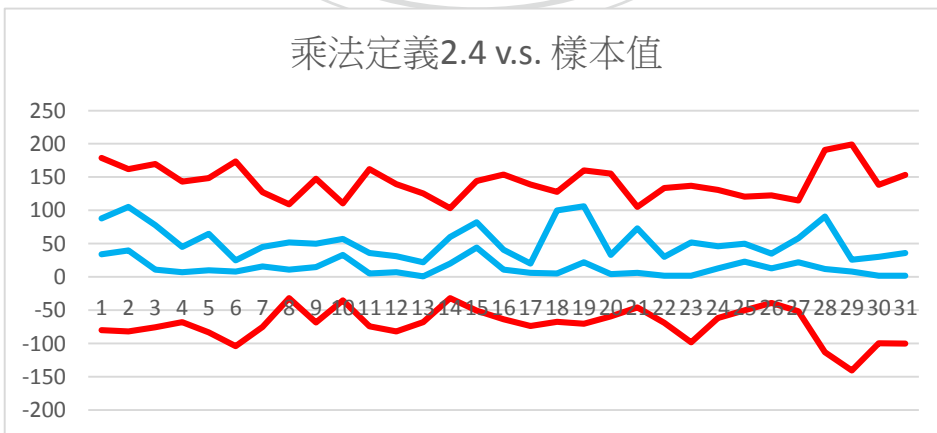


圖 3.3 乘法定義 2.4 之估計值 vs. 樣本值 (105 年 1 月)

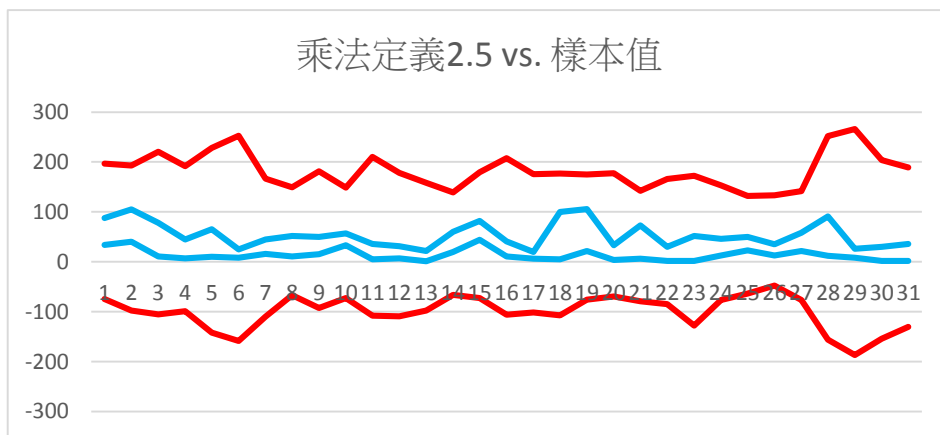


圖 3.4 乘法定義 2.5 之估計值 vs. 樣本值 (105 年 1 月)

以下計算 31 天，估計值與實際值的總區間距離與總平均距離：

- (1) 乘法定義 2.3 之估計值與實際值距離加總

$$\sum_{i=1}^{31} D(Y_i, \hat{Y}_i) = f(< 0.211, 1480.33; 634.44 >) = 423.05$$

區間距離為：< 0.211, 1480.33; 634.44 >

平均距離為：423.05

- (2) 乘法定義 2.4 之估計值與實際值距離加總

$$\sum_{i=1}^{31} D(Y_i, \hat{Y}_i) = f(< 0, 4494.46; 1253 >) = 1620.73$$

區間距離為：< 0, 4494.46; 1253 >

平均距離為：1620.73

- (3) 乘法定義 2.5 之估計值與實際值距離加總

$$\sum_{i=1}^{31} D(Y_i, \hat{Y}_i) = f(< 0, 5452.56; 1253 >) = 2099.78$$

區間距離為：< 0, 5452.56; 1253 >

平均距離為：2099.78

由上述結果可知，定義 2.3 之估計值與實際值區間平均距離加總最小，因此，樣本的最佳區間相乘定義為乘法定義 2.3。

由圖 3.3、圖 3.4 的結果中，可以明顯地觀察出區間乘法定義 2.4 與 2.5 所估計出來的 $Y(X_i)_{2.4}$ 、 $Y(X_i)_{2.5}$ 與實際樣本值差異非常大。(陳孝煒、吳柏林(2007)，區間迴歸與模糊樣本分析中表示)關於模糊參數的運算定義方式，可以依據不同領域的問題；不同的樣本特性、群性，來做更貼近母體的定義，不是狹義的只能採用幾個特別定義，只要能夠符合樣本特性，選取適當乘法定義的最終目的是讓迴歸估計誤差最小。因此，圖 3.3、圖 3.4 的結果可能是因為區間數乘法與區間數加法定義與此樣本特性的配適程度所造成的。

雖然，乘法定義 2.3 為此樣本最佳的區間乘法定義，但由圖 3.2 中，我們可以觀察出，估計值區間的半徑有過於保守的趨勢（半徑偏小），這也造成估計區間與實際區間的重疊距離，相較其它結果來的小，且在 1 月 20 日的估計區間完全沒有覆蓋住實際區間。

二維度技術與乘法定義 2.3，在區間距離的結果中，二維度技術估計值與實際值在 31 筆資料中，皆與實際值資料有交集，因此最近距離總和為零。雖然，最遠距離總和的結果中，二維度技術的誤差相較於乘法定義 2.3 大。但若以覆蓋程度（重疊距離）的結果比較，二維度技術比起乘法定義 2.3 佳。

藉由區間數平均距離定義 2.2，同時考量三者因素（最近距離、重疊距離、重疊距離），二維度技術在整體的表現中，較貼近樣本（平均距離最小）。

3.2 台灣加權股價指數價量分析

探討報酬波動的行為一直是經濟學家和計量經濟學家們非常感興趣的研究。本研究欲從模糊理論的觀點切入，並以區間迴歸分析架構，藉由價量分析預測價格的波動區間。本研究蒐集 2015 年 6 月 1 日~ 2016 年 4 月 30 日台灣證券交易所發佈的台灣加權股價指數價量資料，指數的價為自迴歸行為的變數，前一期的價與一周的交易量為獨立變數，資料採取最大值與最小值的區間資料作為我們的模糊樣本。

考慮區間迴歸模型如下：

$$Y_{\text{價},t} = A_0 + A_1 X_{\text{量}} + A_2 X_{\text{價},t-1} \quad (3.7)$$

$Y_{\text{價},t}$ 為今日價格變化，以區間模糊數表示

$X_{\text{量}}$ 為交易量(前五天交易日)變化，以區間模糊數表示

$X_{\text{價},t-1}$ 為昨日價格變化，以區間模糊數表示

採用二維模糊迴歸方法，估算出中心迴歸模式與半徑迴歸模式如下：

(1) 中心迴歸模式：

$$Y_{\text{價},t,C} = 205 + 0.97C_{\text{價},t-1} + 0.14C_{\text{量}} \quad (3.8)$$

(2) 半徑迴歸模式：

$$Y_{\text{價},t,S} = 34.11 + 0.27S_{\text{價},t-1} + 0.04S_{\text{量}} \quad (3.9)$$

藉由中心迴歸模式與半徑迴歸模式，我們針對 105 年 5 月 3 日~5 月 20 日的台灣加權股價指數價格資料進行估計，其結果如下：

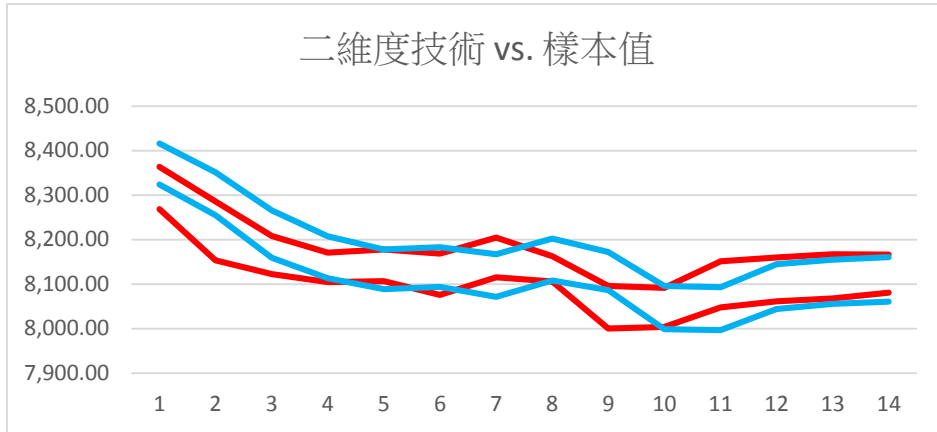


圖 3.5 二維度技術之估計值 vs. 樣本值 (5/3~5/20)

圖 3.5 中，藍色曲線範圍為模糊迴歸模式下，台灣加權股價指數價格估計區間，紅色曲線為實際樣本價格區間。

藉由定義 2.1 (區間模糊數區間距離) 和定義 2.2 (區間模糊數平均距離)，計算出 14 天估計值與樣本值的區間差距如下：

$$\begin{aligned} \sum_{i=1}^{14} D(Y_i, \hat{Y}_i) &= \sum_{i=1}^{14} f(\langle L_{Y_i \hat{Y}_i}, M_{Y_i \hat{Y}_i}; C_{Y_i \hat{Y}_i} \rangle) = f(\langle 0, 1768.51; 823.02 \rangle) \\ &= 472.75 \end{aligned} \quad (3.10)$$

由結果顯示，估計值與實際值 14 天距離加總，最近距離總和為 0；最遠距離總和為 1768.51；重疊距離總和為 823.02；平均距離總和為 472.75。

接著我們將比較 2.6 節 (其它模糊迴歸方法)，陳孝煒、吳柏林 (2007)，區間迴歸與模糊樣本分析中，所採用的方法。

首先，將中心迴歸參數與半徑迴歸參數代入台灣加權股價指數模糊迴歸架構 (3.7) 中，並改寫如下：

(1) 中心與半徑迴歸模式：

$$Y_{\text{價},t} = \langle 205, 34.11 \rangle + \langle 0.97, 0.27 \rangle X_{\text{價},t-1} + \langle 0.14, 0.04 \rangle X_{\text{量}} \quad (3.11)$$

(2) 區間迴歸模式：

$$Y_{PM_{10,t}} = [170.89, 239.11] + [0.7, 1.24]X_{價,t-1} + [0.1, 0.18]X_{量} \quad (3.12)$$

藉由區間數乘法定義 2.3~2.5 求得三種估計值 $\widehat{Y(X_1)_{2.2}}$ 、 $\widehat{Y(X_1)_{2.3}}$ 、 $\widehat{Y(X_1)_{2.4}}$ ，

其結果如下：

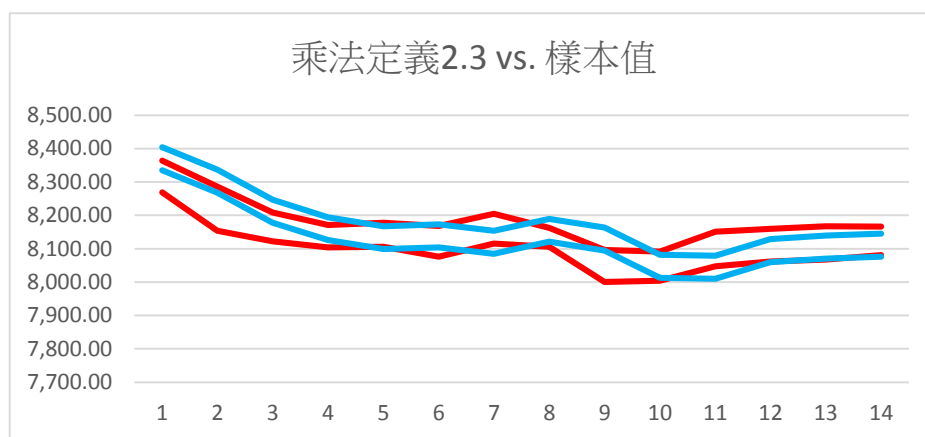


圖 3.6 乘法定義 2.3 之估計值 vs. 樣本值 (5/3~5/20)

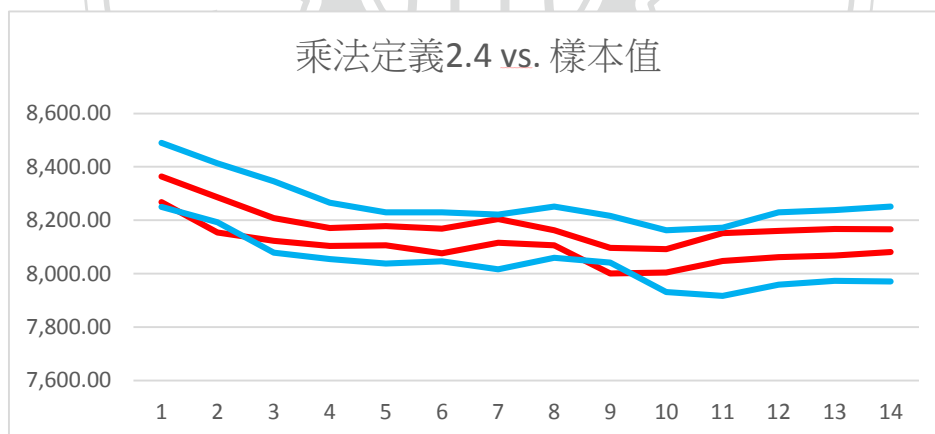


圖 3.7 乘法定義 2.4 之估計值 vs. 樣本值 (5/3~5/20)

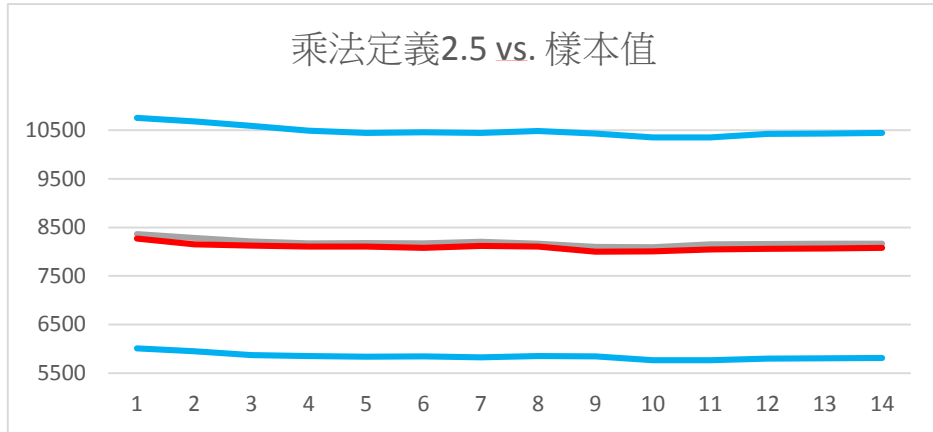


圖 3.8 乘法定義 2.5 之估計值 vs. 樣本值 (5/3~5/20)

以下計算 14 天，估計值與實際值的總區間距離與總平均距離：

(1) 乘法定義 2.3 之估計值與實際值距離加總

$$\sum_{i=1}^{14} D(Y_i, \hat{Y}_i) = f(< 0, 1582.79; 629.41 >) = 476.69$$

區間距離為：< 0, 1582.79; 629.41 >

平均距離為：476.69

(2) 乘法定義 2.4 之估計值與實際值距離加總

$$\sum_{i=1}^{14} D(Y_i, \hat{Y}_i) = f(< 0, 2693.88; 1179.95 >) = 756.97$$

區間距離為：< 0, 2693.88; 1179.95 >

平均距離為：756.97

(3) 乘法定義 2.5 之估計值與實際值距離加總

$$\sum_{i=1}^{14} D(Y_i, \hat{Y}_i) = f(< 0, 33634.08; 1260.3 >) = 16186.89$$

區間距離為：< 0, 33634.08; 1260.3 >

平均距離為：16186.89

由上述結果可知，定義 2.3 估計值與實際值區間平均距離加總最小，因此樣本的最佳區間相乘定義為乘法定義 2.3。

圖 3.7、圖 3.8 的結果中，明顯的觀察出區間乘法定義 2.4 與 2.5 所估計出來的 $Y(X_i)_{2.4}$ 、 $Y(X_i)_{2.5}$ 與實際樣本值差異非常大。結果與 3.1 節懸浮微粒 PM_{10} 濃度預測的結果相似，因此乘法定義 2.4 與 2.5 與此樣本的配適度不佳。

由圖 3.6 與圖 3.7，我們可看出乘法定義 2.3 與 2.4 兩著之間的差異，對於區間半徑的預測。乘法定義 2.4 預測過大，也造成較大的誤差，與乘法定義 2.3 相反。

比較二維度技術與乘法定義 2.3 的結果，兩者與 14 筆實際值資料皆有交集，最小距離總和皆為零。雖然，最遠距離總和的結果中，二維度技術的誤差相較於乘法定義 2.3 來的差。但以覆蓋率（重疊距離）的結果，乘法定義 2.3 較不佳。

藉由區間數平均距離定義 2.2，同時考量三者因素（最近距離、重疊距離、重疊距離），二維度技術在整體的表現中，較貼近樣本（平均距離最小）。

4 結論

傳統迴歸和模糊迴歸建立在不同的理論基礎上。傳統迴歸是以機率為基礎；而模糊迴歸則是以隸屬度為基礎。傳統迴歸應變數的估計值與觀測值之間的差異為應變數的隨機性；而模糊迴歸應變數的估計值則產生於觀測值的模糊性。兩者最大的不同處在於樣本資料分析的方式：模糊迴歸模式獨立變數與應變數皆為模糊數（區間值）；傳統迴歸係數的預測值均為精確值。然而，生活中許多事物，若採用假性的「精確值」，可能誤導模型的建構，也可能擴大預測結果和實際狀態之間的誤差。

文中我們考慮以二維度技術，建構出中心迴歸模式與半徑迴歸模式，並分別估計出應變數的中心與半徑。如此一來，可避免區間數乘法與區間數加法與樣本的配適性的問題。也不用依據不同領域的問題、樣本特性、群性，來重新定義更貼近母體的模糊參數運算法則。

此外，為了衡量兩個區間之間的差異程度，我們首度引進區間數區間距離的概念，並結合克卜勒第三運動定律中，行星平均距離的概念，將區間數轉換為實數值，以方便做比較，並且在 2.5 節推導賦距測度等性質，此距離定義與相關性質推導，為本文最主要的貢獻。

藉由上述距離定義，在實例探討中，我們發現二維度技術的估計區間相較於其它三種乘法定義的估計區間，與實際樣本值區間差異最小。且二維度技術下，估計的半徑也不會有過於保守或是過大等問題。

由文中所舉懸浮微粒濃度預測與大盤指數價格預測實例中，我們可以瞭解到現實生活中許多事物隨時都處於波動的狀態。利用傳統迴歸求得的估計值為一個實數值，不但不符合樣本特性，也可能具有很大的誤差。

對於隨時都在變動的觀測值，我們無法以單一數值精準預估。但藉由長期觀察，可發現變動樣本週期性的規律。就像是週期波的形狀，具有一定的波峰跟波谷，亦即數據變動範圍有上下界，超出上界或低於下界的機會很低。因此，如何估計出上下界是很重要的一門技術，這是傳統迴歸較少考慮的問題。

傳統迴歸與時間序列都是預測下一刻的數值(橫向預測)，然而在統計學中，這樣的預測往往會因隨機性造成誤差。而模糊迴歸除了橫向預測外，同時進行縱向預測(上下界預測)。藉由縱向的預測，我們可以做出相對應的決策。因此，我們認為以模糊迴歸作為工具，配合我們提出之距離定義來選取適切之模型參數，來表達與解釋許多生活中的樣本特性，可彌補傳統迴歸不足之處。



5 參考文獻

- [1] 吳柏林 (2015), 模糊統計導論—方法與應用, 五南出版社
- [2] 陳孝煒、吳柏林 (2007), 區間迴歸與模糊樣本分析, 管理科學與統計決策, 4, 54-65
- [3] 楊敏生 (1994), 模糊理論簡介, 數學傳播季刊, 18
- [4] 邱鼎泰 (2012), 學習技術分析, 理財周刊, 614
- [5] 李聯雄、賴嘉祥 (2011), 大陸沙塵暴對我國室外作業勞工健康影響之研究, 行政院勞工委員會勞工安全衛生研究所
- [6] 溫志中、蔡立宏、楊尚威 (2009) 台灣港區大氣能見度特性探討—基隆、台中、高雄, 港灣報導, 82, 31-39
- [7] 蔡春進、繆敦耀、簡聰智、陳泰任 (1999), 裸露地逸散性粒狀空氣污染物的控制技術研究, 第十六屆空氣污染控制技術研討會, 台中朝陽科技大學台灣
- [8] 細懸浮微粒(PM₂)對人體健康危害之預防策略研究公聽會 (2013), 國家衛生研究院國家環境毒物研究中心
- [9] 江世民 (1999), 台北地區臭氧及懸浮微粒型態與其他汙染物及氣象條件之分析比較, 第十六屆空氣污染控制技術研討會, 台中朝陽科技大學
- [10] 潘南飛、李思賢 (2005), 模糊迴歸用於分析作業之工期, 工程科技與教育學刊, 2, 152-160
- [11] 台灣十大死因 都是空氣惹的禍, 遠見雜誌
- [12] Wu, B. & Tseng, N. (2002). A new approach to fuzzy regression models with application to business cycle analysis. *Fuzzy Sets and System*. 130, 33-42.

- [13] Zadeh, L.A. (1965). Fuzzy sets, *Information and Control* 8, 338-353.
- [14] Tanaka, H. & Uejima, S. & Asai, K. (1980). Fuzzy Linear Regression Model. International Congress on Applied Systems Research and Cybernetics. Aculpoco, Mexico.
- [15] Tanaka, H. & Uejima, S. & Asai, K. (1982). Linear Regression Analysis with Fuzzy model. *IEEE Trans. SystemsMan Cybernet*, vol SMC 12, 903-907.
- [16] Diamond, P. (1988). Fuzzy least squares, *Information Sciences* 46(3), 141-157.
- [17] Curtis, R. K, (1990) .Complexity and predictability: The application of chaos theory to economic forecasting. *Futures Research Quarterly*, 6(4), 57-70.
- [18] Kacprzyk J. & Fedrizzi M. (1992). Fuzzy Regression Analysis.
- [19] Tanaka, H. & Ishibuchi, H. (1993). Anarchitecture of neural networks with interval weights and its application to fuzzy regression analysis. *Fuzzy Sets and Systems*, 57, 27-39.
- [20] Chang P. T. & Lee E. S., (1994), Fuzzy Linear Regression with spreads Unrestricted in Sign, *Computers Math. Applic*, 281994, 61-70.
- [21] Chang P. T. et al, (1996). Applying fuzzy linear regression to VDT legibility, *Fuzzy Sets and Systems*, 80, 197-204.
- [22] Chang Yun-His O. & Ayyub, B. M (1996). Hybrid Fuzzy Regression Analysis and Its applications. *Uncertainty Modeling and Analysis in Civil Engineering*, CRC Press Inc. 33-41.
- [23] Chang P. T. (1997). Fuzzy Seasonality Forecasting. *Fuzzy Sets and Systems*, 90, 1-10.
- [24] Yen K. K. & Ghoshray S. & Roig G. (1999). A Linear Regression Model Using

- Triangular Fuzzy Number Coefficient, *Fuzzy Sets and Systems*, 106, 167-177.
- [25] Wang H. F. & Tsaur R. C. (2000). Insight of a fuzzy regression model, *Fuzzy Sets and Systems*, 112, 355-369.
- [26] Lee, H. T. & Chen, S. H. (2001). Fuzzy regression model with fuzzy input and output data for manpower forecasting, *Fuzzy Sets and Systems*, 119, 205-213.
- [27] Tseng Yu-Heng & Paul Durbin & Tzeng Gwo-Hshiong (2001). Using a Fuzzy Piecewise Regression Analysis to Predict the Nonlinear Time-Series of Turbulent Flows with Automatic Change-Point Detection, *Flow, Turbulence and Combustion*, 67, 81-106.
- [28] Kao Chiang & Chyu Chin - Lu (2002). A fuzzy linear regression model with better explanatory power. *Fuzzy Sets and Systems*, 126(3), 401-409
- [29] Kutner & Nachtsheim & Neter (2004). *Applied Linear Regression Model*, Fourth Edition.
- [30] Reza Ghodsi & MohammadSaleh Zakerinia & Mahdi Joka (2010). Neural network and Fuzzy Regression Model for Forecasting Short Term Price in Ontario Electricity Market. *Proceedings of the 41st International Conference on Computers & Industrial Engineering*.954-959.
- [31] Ul-Saufie A.Z & Yahya A.S & Ramli N.A (2010). Improving multiple linear regression model using principal component analysis for predicting PM_{10} concentration in Seberang Prai, Pulau Pinang, *International Journal Of Environmental Sciences* Vol2, No 2, 2011.
- [32] Ghodratollah Emamverdi & Ebrahim Siami Araghi & Fatemeh Fahimifar (2013). Presenting a Fuzzy ARIMA Model for Forecasting Stock Market Price Index of

Iran. J. Basic. Appl. Sci. Res, 3(6), 37-43, 2013.

[33] Arwa S. Sayegh & Said Munir & Turki M. Habeebullah (2014). Comparing the Performance of Statistical Models for Predicting PM_{10} Concentrations, Aerosol and Air Quality Research, 14, 653–665.

[34] Liem Tran & Lucien Duckstein (2002). Comparison of fuzzy numbers using a fuzzy distance measure, Fuzzy Sets and Systems, 130, 331-341.



6 附錄

Lemma 4 三角不等式： $D(\tilde{A}, \tilde{C}) \leq D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \quad \forall \tilde{A}, \tilde{B}, \tilde{C} \in \mathbb{R}$

<pf> Suppose that $\frac{A_1+A_2}{2} \leq \frac{B_1+B_2}{2} \leq \frac{C_1+C_2}{2}$.

This implies that $M_{AB} = B_2 - A_1$, $M_{BC} = C_2 - B_1$ and $M_{AC} = C_2 - A_1$.

Case 1 $C_2 \leq B_2 \leq A_2$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(B_1 - B_2) \quad C_{BC} = C_{AC} = -(C_1 - C_2)$$

$$\Rightarrow \frac{(B_1 - B_2) + (B_2 - A_1)}{2} + \frac{(C_1 - C_2) + (C_2 - B_1)}{2} = \frac{(C_1 - C_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) = D(\tilde{A}, \tilde{C})$$

Case 2 $B_2 \leq C_2 \leq A_2$ 且 $A_1 \leq B_1 \leq C_1$ 且 $B_2 \geq C_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(B_1 - B_2) \quad C_{BC} = -(C_1 - B_2) \quad C_{AC} = -(C_1 - C_2)$$

$$\Rightarrow \frac{(B_1 - B_2) + (B_2 - A_1)}{2} + \frac{(C_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - C_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 3 $B_2 \leq C_2 \leq A_2$ 且 $A_1 \leq B_1 \leq C_1$ 且 $B_2 \leq C_1$

$$L_{AB} = L_{AC} = 0 \quad L_{BC} = C_1 - B_2$$

$$C_{AB} = -(B_1 - B_2) \quad C_{BC} = 0 \quad C_{AC} = -(C_1 - C_2)$$

$$\Rightarrow \frac{(B_1 - B_2) + (B_2 - A_1)}{2} + \frac{(C_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - C_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 4 $B_2 \leq C_2 \leq A_2$ 且 $A_1 \leq C_1 \leq B_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(B_1 - B_2) \quad C_{BC} = -(B_1 - B_2) \quad C_{AC} = -(C_1 - C_2)$$

$$\Rightarrow \frac{(B_1 - B_2) + (B_2 - A_1)}{2} + \frac{(B_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - C_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 5 $C_2 \leq A_2 \leq B_2$ 且 $A_1 \leq B_1 \leq C_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(B_1 - A_2) \quad C_{BC} = -(C_1 - C_2) \quad C_{AC} = -(C_1 - C_2)$$

$$\Rightarrow \frac{(B_1 - A_2) + (B_2 - A_1)}{2} + \frac{(C_1 - C_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - C_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 6 $C_2 \leq A_2 \leq B_2$ 且 $B_1 \leq A_1 \leq C_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(A_1 - A_2) \quad C_{BC} = -(C_1 - C_2) \quad C_{AC} = -(C_1 - C_2)$$

$$\Rightarrow \frac{(A_1 - A_2) + (B_2 - A_1)}{2} + \frac{(C_1 - C_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - C_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 7 $A_2 \leq C_2 \leq B_2$ 且 $A_1 \leq B_1 \leq C_1$ 且 $A_2 \geq C_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(B_1 - A_2) \quad C_{BC} = -(C_1 - C_2) \quad C_{AC} = -(C_1 - A_2)$$

$$\Rightarrow \frac{(B_1 - A_2) + (B_2 - A_1)}{2} + \frac{(C_1 - C_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 8 $A_2 \leq C_2 \leq B_2$ 且 $A_1 \leq B_1 \leq C_1$ 且 $C_1 \geq A_2 \geq B_1$

$$L_{AB} = L_{BC} = 0 \quad L_{AC} = (C_1 - A_2)$$

$$C_{AB} = -(B_1 - A_2) \quad C_{BC} = -(C_1 - C_2) \quad C_{AC} = 0$$

$$\Rightarrow \frac{(B_1-A_2)+(B_2-A_1)}{2} + \frac{(C_1-C_2)+(C_2-B_1)}{2} \geq \frac{(C_1-A_2)+C_2-A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 9 $A_2 \leq C_2 \leq B_2$ 且 $A_1 \leq B_1 \leq C_1$ 且 $B_1 \geq A_2$

$$L_{BC} = 0 \quad L_{AB} = (B_1 - A_2) \quad L_{AC} = (C_1 - A_2)$$

$$C_{AB} = 0 \quad C_{BC} = -(C_1 - C_2) \quad C_{AC} = 0$$

$$\Rightarrow \frac{(B_1-A_2)+(B_2-A_1)}{2} + \frac{(C_1-C_2)+(C_2-B_1)}{2} \geq \frac{(C_1-A_2)+C_2-A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 10 $A_2 \leq C_2 \leq B_2$ 且 $B_1 \leq A_1 \leq C_1$ 且 $A_2 \geq C_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(A_1 - A_2) \quad C_{BC} = -(C_1 - C_2) \quad C_{AC} = -(C_1 - A_2)$$

$$\Rightarrow \frac{(A_1-A_2)+(B_2-A_1)}{2} + \frac{(C_1-C_2)+(C_2-B_1)}{2} \geq \frac{(C_1-A_2)+C_2-A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 11 $A_2 \leq C_2 \leq B_2$ 且 $B_1 \leq A_1 \leq C_1$ 且 $A_2 \leq C_1$

$$L_{AB} = L_{BC} = 0 \quad L_{AC} = (C_1 - A_2)$$

$$C_{AB} = -(A_1 - A_2) \quad C_{BC} = -(C_1 - C_2) \quad C_{AC} = 0$$

$$\Rightarrow \frac{(A_1-A_2)+(B_2-A_1)}{2} + \frac{(C_1-C_2)+(C_2-B_1)}{2} \geq \frac{(C_1-A_2)+C_2-A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 12 $A_2 \leq C_2 \leq B_2$ 且 $B_1 \leq C_1 \leq A_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(A_1 - A_2) \quad C_{BC} = -(C_1 - C_2) \quad C_{AC} = -(A_1 - A_2)$$

$$\Rightarrow \frac{(A_1-A_2)+(B_2-A_1)}{2} + \frac{(C_1-C_2)+(C_2-B_1)}{2} \geq \frac{(A_1-A_2)+C_2-A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 13 $B_2 \leq A_2 \leq C_2$ 且 $C_1 \leq A_1 \leq B_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(B_1 - B_2) \quad C_{BC} = -(B_1 - B_2) \quad C_{AC} = -(A_1 - A_2)$$

$$\Rightarrow \frac{(B_1 - B_2) + (B_2 - A_1)}{2} + \frac{(B_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(A_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 14 $B_2 \leq A_2 \leq C_2$ 且 $A_1 \leq C_1 \leq B_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(B_1 - B_2) \quad C_{BC} = -(B_1 - B_2) \quad C_{AC} = -(C_1 - A_2)$$

$$\Rightarrow \frac{(B_1 - B_2) + (B_2 - A_1)}{2} + \frac{(B_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 15 $B_2 \leq A_2 \leq C_2$ 且 $A_1 \leq B_1 \leq C_1$ 且 $B_2 \geq C_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(B_1 - B_2) \quad C_{BC} = -(C_1 - B_2) \quad C_{AC} = -(C_1 - A_2)$$

$$\Rightarrow \frac{(B_1 - B_2) + (B_2 - A_1)}{2} + \frac{(C_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 16 $B_2 \leq A_2 \leq C_2$ 且 $A_1 \leq B_1 \leq C_1$ 且 $B_2 \leq C_1 \leq A_2$

$$L_{AB} = L_{AC} = 0 \quad L_{BC} = (C_1 - B_2)$$

$$C_{AB} = -(B_1 - B_2) \quad C_{BC} = 0 \quad C_{AC} = -(C_1 - A_2)$$

$$\Rightarrow \frac{(B_1 - B_2) + (B_2 - A_1)}{2} + \frac{(C_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 17 $B_2 \leq A_2 \leq C_2$ 且 $A_1 \leq B_1 \leq C_1$ 且 $A_2 \leq C_1$

$$L_{AB} = 0 \quad L_{BC} = (C_1 - B_2) \quad L_{AC} = (C_1 - A_2)$$

$$C_{AB} = -(B_1 - B_2) \quad C_{BC} = 0 \quad C_{AC} = 0$$

$$\Rightarrow \frac{(B_1 - B_2) + (B_2 - A_1)}{2} + \frac{(C_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 18 $A_2 \leq B_2 \leq C_2$ 且 $A_1 \leq B_1 \leq C_1$ 且 $A_2 \geq C_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(B_1 - B_2) \quad C_{BC} = -(C_1 - B_2) \quad C_{AC} = -(C_1 - A_2)$$

$$\Rightarrow \frac{(B_1 - B_2) + (B_2 - A_1)}{2} + \frac{(C_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 19 $A_2 \leq B_2 \leq C_2$ 且 $A_1 \leq B_1 \leq C_1$ 且 $C_1 \geq A_2 \geq B_1$ 且 $B_2 \geq C_1$

$$L_{AB} = L_{BC} = 0 \quad L_{AC} = (C_1 - A_2)$$

$$C_{AB} = -(B_1 - B_2) \quad C_{BC} = -(C_1 - B_2) \quad C_{AC} = 0$$

$$\Rightarrow \frac{(B_1 - B_2) + (B_2 - A_1)}{2} + \frac{(C_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 20 $A_2 \leq B_2 \leq C_2$ 且 $A_1 \leq B_1 \leq C_1$ 且 $C_1 \geq A_2 \geq B_1$ 且 $B_2 \leq C_1$

$$L_{AB} = 0 \quad L_{AC} = (C_1 - A_2) \quad L_{BC} = (C_1 - B_2)$$

$$C_{AB} = -(B_1 - B_2) \quad C_{BC} = 0 \quad C_{AC} = 0$$

$$\Rightarrow \frac{(B_1 - B_2) + (B_2 - A_1)}{2} + \frac{(C_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 21 $A_2 \leq B_2 \leq C_2$ 且 $A_1 \leq B_1 \leq C_1$ 且 $A_2 \leq B_1$ 且 $B_2 \geq C_1$

$$L_{BC} = 0 \quad L_{AC} = (C_1 - A_2) \quad L_{AB} = (B_1 - B_2)$$

$$C_{AB} = 0 \quad C_{BC} = -(C_1 - B_2) \quad C_{AC} = 0$$

$$\Rightarrow \frac{(B_1-B_2)+(B_2-A_1)}{2} + \frac{(C_1-B_2)+(C_2-B_1)}{2} \geq \frac{(C_1-A_2)+C_2-A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 22 $A_2 \leq B_2 \leq C_2$ 且 $A_1 \leq B_1 \leq C_1$ 且 $A_2 \leq B_1$ 且 $B_2 \leq C_1$

$$L_{BC} = (C_1 - B_2) \quad L_{AC} = (C_1 - A_2) \quad L_{AB} = (B_1 - B_2)$$

$$C_{AB} = 0 \quad C_{BC} = 0 \quad C_{AC} = 0$$

$$\Rightarrow \frac{(B_1-B_2)+(B_2-A_1)}{2} + \frac{(C_1-B_2)+(C_2-B_1)}{2} \geq \frac{(C_1-A_2)+C_2-A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 23 $A_2 \leq B_2 \leq C_2$ 且 $B_1 \leq A_1 \leq C_1$ 且 $A_2 \geq C_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(A_1 - A_2) \quad C_{BC} = -(C_1 - B_2) \quad C_{AC} = -(C_1 - A_2)$$

$$\Rightarrow \frac{(A_1-A_2)+(B_2-A_1)}{2} + \frac{(C_1-B_2)+(C_2-B_1)}{2} \geq \frac{(C_1-A_2)+C_2-A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 24 $A_2 \leq B_2 \leq C_2$ 且 $B_1 \leq A_1 \leq C_1$ 且 $A_2 \leq C_1 \leq B_2$

$$L_{AB} = L_{BC} = 0 \quad L_{AC} = (C_1 - A_2)$$

$$C_{AB} = -(A_1 - A_2) \quad C_{BC} = -(C_1 - B_2) \quad C_{AC} = 0$$

$$\Rightarrow \frac{(A_1-A_2)+(B_2-A_1)}{2} + \frac{(C_1-B_2)+(C_2-B_1)}{2} \geq \frac{(C_1-A_2)+C_2-A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 25 $A_2 \leq B_2 \leq C_2$ 且 $B_1 \leq A_1 \leq C_1$ 且 $C_1 \geq B_2$

$$L_{AB} = 0 \quad L_{AC} = (C_1 - A_2) \quad L_{BC} = (C_1 - B_2)$$

$$C_{AB} = -(A_1 - A_2) \quad C_{BC} = 0 \quad C_{AC} = 0$$

$$\Rightarrow \frac{(A_1-A_2)+(B_2-A_1)}{2} + \frac{(C_1-B_2)+(C_2-B_1)}{2} \geq \frac{(C_1-A_2)+C_2-A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 26 $A_2 \leq B_2 \leq C_2$ 且 $A_1 \leq C_1 \leq B_1$ 且 $A_2 \geq B_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(B_1 - A_2) \quad C_{BC} = -(B_1 - B_2) \quad C_{AC} = -(C_1 - A_2)$$

$$\Rightarrow \frac{(B_1 - A_2) + (B_2 - A_1)}{2} + \frac{(B_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 27 $A_2 \leq B_2 \leq C_2$ 且 $A_1 \leq C_1 \leq B_1$ 且 $C_1 \leq A_2 \leq B_1$

$$L_{BC} = L_{AC} = 0 \quad L_{AB} = (B_1 - A_2)$$

$$C_{AB} = 0 \quad C_{BC} = -(B_1 - B_2) \quad C_{AC} = -(C_1 - A_2)$$

$$\Rightarrow \frac{(B_1 - A_2) + (B_2 - A_1)}{2} + \frac{(B_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 28 $A_2 \leq B_2 \leq C_2$ 且 $A_1 \leq C_1 \leq B_1$ 且 $A_2 \leq C_1$

$$L_{BC} = 0 \quad L_{AB} = (B_1 - A_2) \quad L_{AC} = (C_1 - A_2)$$

$$C_{AB} = 0 \quad C_{BC} = -(B_1 - B_2) \quad C_{AC} = 0$$

$$\Rightarrow \frac{(B_1 - A_2) + (B_2 - A_1)}{2} + \frac{(B_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(C_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 29 $A_2 \leq B_2 \leq C_2$ 且 $C_1 \leq A_1 \leq B_1$ 且 $A_2 \geq B_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(B_1 - A_2) \quad C_{BC} = -(B_1 - B_2) \quad C_{AC} = -(A_1 - A_2)$$

$$\Rightarrow \frac{(B_1 - A_2) + (B_2 - A_1)}{2} + \frac{(B_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(A_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 30 $A_2 \leq B_2 \leq C_2$ 且 $B_1 \leq C_1 \leq A_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(A_1 - A_2) \quad C_{BC} = -(C_1 - B_2) \quad C_{AC} = -(A_1 - A_2)$$

$$\Rightarrow \frac{(A_1 - A_2) + (B_2 - A_1)}{2} + \frac{(C_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(A_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

Case 31 $A_2 \leq B_2 \leq C_2$ 且 $C_1 \leq B_1 \leq A_1$

$$L_{AB} = L_{BC} = L_{AC} = 0$$

$$C_{AB} = -(A_1 - A_2) \quad C_{BC} = -(B_1 - B_2) \quad C_{AC} = -(A_1 - A_2)$$

$$\Rightarrow \frac{(A_1 - A_2) + (B_2 - A_1)}{2} + \frac{(B_1 - B_2) + (C_2 - B_1)}{2} \geq \frac{(A_1 - A_2) + C_2 - A_1}{2}$$

$$\Rightarrow D(\tilde{A}, \tilde{B}) + D(\tilde{B}, \tilde{C}) \geq D(\tilde{A}, \tilde{C})$$

