Introduction to Chaotic Dynamics Final Project

State Space Reconstruction for Stock Movements of Taiwan

許智翔 B06502023 羅恩至 B06502027 郭俊廷 B06502047

1 Introduction

由於先前觀察每日各支股票的變動情形均不同，而單支股票於不同時間下的漲跌情況亦不一致，並且難以掌握其變化規律，無法預測未來之股價變化，因此想透過本次混沌力學期末報告，將股票視為一種動態系統，將各支股票的股價與交易量進行State Space Reconstruction，來觀察股票的變化情形是否具有混沌現象。

2 Analysis methods

由於我們想對台灣不同產業的股票作分析，比較不同產業對於股票趨勢的影響，因此決定選擇台塑化(油電燃氣)、中華電(通信網路)、大立光(光電)、台積電(其他電子)、鴻海(半導體)和聯發科(半導體)等六支股票來進行分析。而對於每一支股票，使用2011年1月至2021年1月為期十年的每日交易資料，並針對開盤價(Open)、當日最高點(High)、當日最低點(Low)、收盤價(Close)和交易量(Volumn)等來作一維時間序列的狀態空間重建。

重建的方法與流程為，先分別對每支股票的開盤價(Open)、當日最高點(High)、當日最低點(Low)、收盤價(Close)和交易量(Volumn)，透過Matlab分析得到各自於不同time delay下的mutual information(I(m)值)趨勢圖，並從趨勢圖中找出第一次出現最低I(m)值的time delay，之後再利用此time delay值進行後續的embedding dimension分析，算出於不同維度下的False nearest neighbors所占百分比，並得到每種資料的維度。

而每支股票的分析結果詳見下方點3所示。

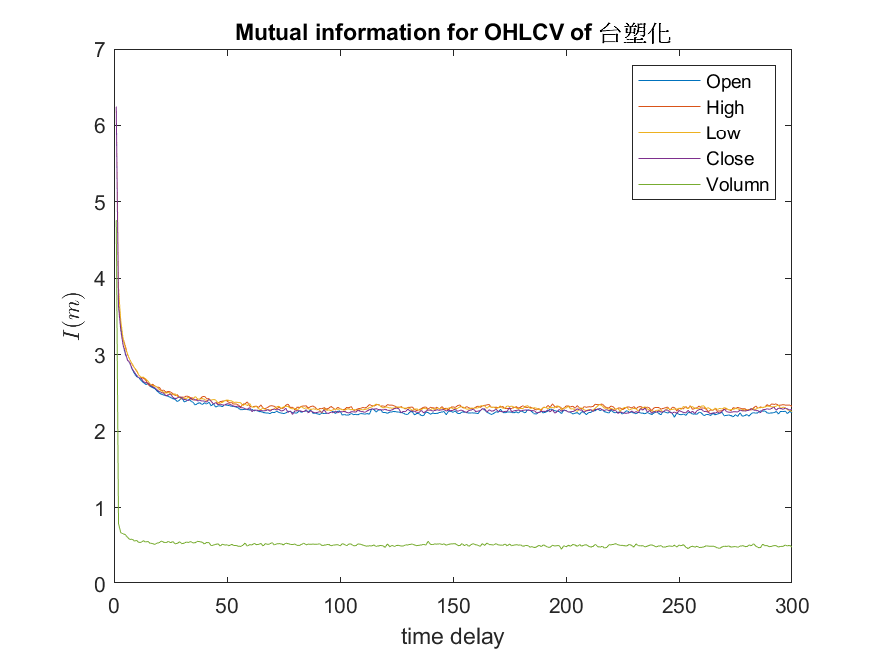
3 Results and conclusions

**Mutual information**

1. 個股分析

　　我們發現在六支股票當中，其中有三支股票的mutual information很快地趨近於平緩，較難看出第一次出現最小值的點。以下進行分析的順序按照平緩程度排序。

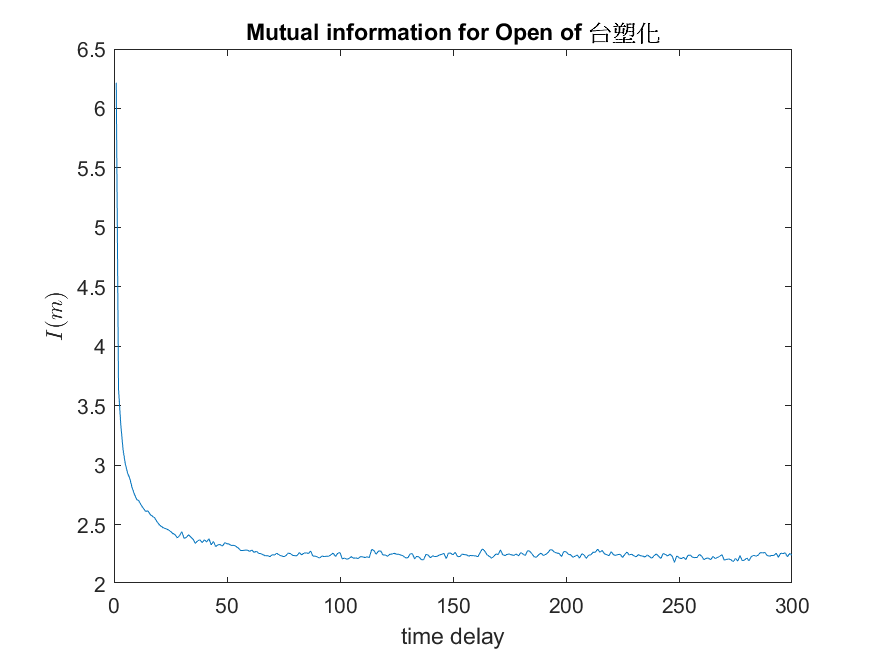
1. 台塑化



圖一：台塑化OHLCV資料的mutual information

由圖中可以發現，開盤價、最高點、最低點和收盤價（後簡稱這四個資訊為OHLC）之間的差異不大，都很快地趨於平緩，並維持在相近的值。而交易量的值則明顯比其他資訊都來得小，但同樣也都會很快地趨於平緩。

由此可以發現，進行state-space reconstruction時，OHLC所需要的time delay都相近，可以由此推測，OHLC之間的關聯十分密切，很OHLC受到同一個系統所影響。而交易量則與OHLC相差較遠，可能受到不同系統的影響，或者是僅受到系統部分參數所控制。



圖二：台塑化開盤價資料的mutual information

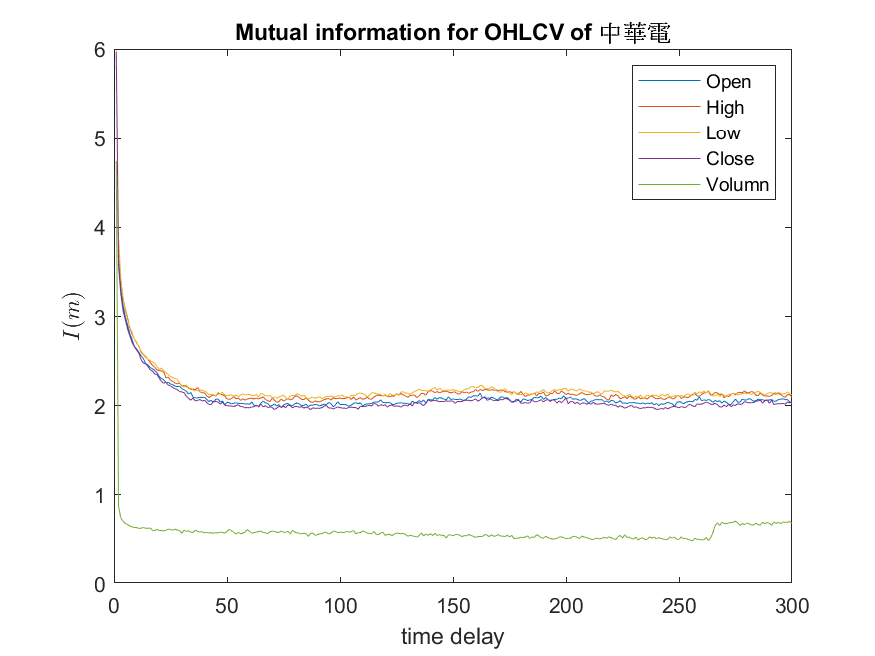
觀察圖二，由於曲線很快地趨近平緩，無法明顯看出第一次最小值的位置，因此我們除了用人工進行判斷外，也使用了matlab內建的函式進行分析，得到以下結果：

由於針對台塑化的資料難以由人工進行準確判斷，後續我們將採用內建函式的數據進行embedding dimension的分析。這邊可以注意到，雖然在圖一中OHLC的曲線類似，但最終計算time delay數據時，可以發現開盤價與其他三者差異甚大，這可能是因為曲線太過平緩所導致的誤差。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 開盤價 | 最高點 | 最低點 | 收盤價 | 交易量 |
| 內建函式 | 59 | 34 | 34 | 35 | 5 |
| 人工判斷 | 80 | 80 | 68 | 65 | 10 |

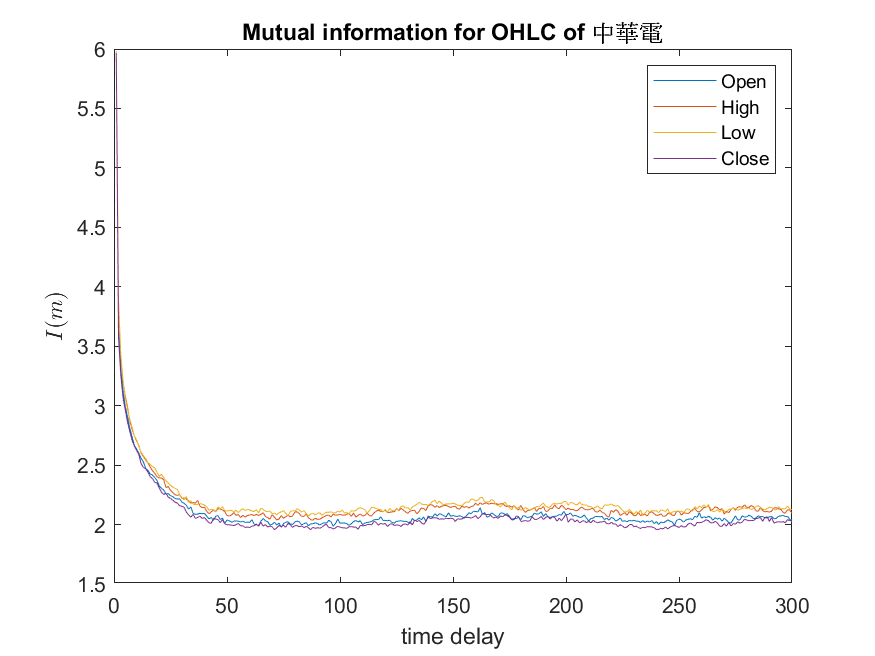
表一：台塑化各項資料的time delay選取值

1. 中華電



圖三：中華電OHLCV資料的mutual information

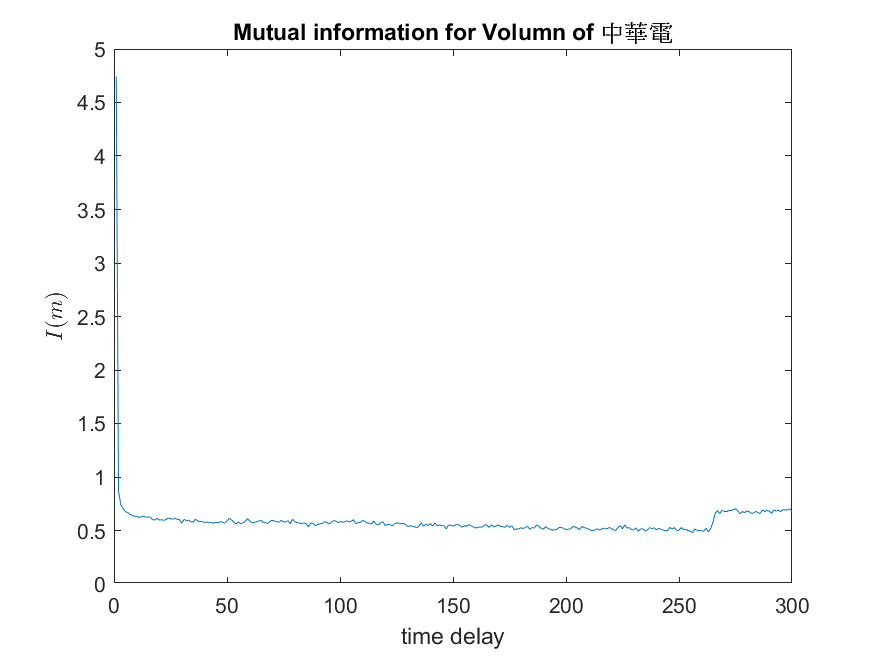
圖三中依然可以發現到，交易量和OHLC之間的差距甚大。以下用另外用兩張圖表觀察OHLC資料和交易量資料的特徵。



圖四：中華電OHLC資料的mutual information

由圖四可以發現，最高點和最低點的資料所得到的曲線相近，而開盤價和收盤價的資料也相近，但最高點和最低點的資料稍微比另外兩者高。由此可以推測，影響最高點和最低點的系統參數，可能又與影響開盤價和收盤價的系統參數有所區別。

另外，可以發現，OHLC資料在下降後，有微幅地上升，不過大致上仍然很平緩，較難觀察出詳細的第一次最小值位置。



圖五：中華電交易量資料的mutual information

由圖五可以觀察到，在time delay取到260左右時，突然出現了一個向上的斷層，這代表距離260天左右以後的資料，與當天的交易量資料的關聯性會再次下降。這個現象雖然沒有在台塑化中看到，但在中華電以及後續分析的部份股票中都可以觀察到。

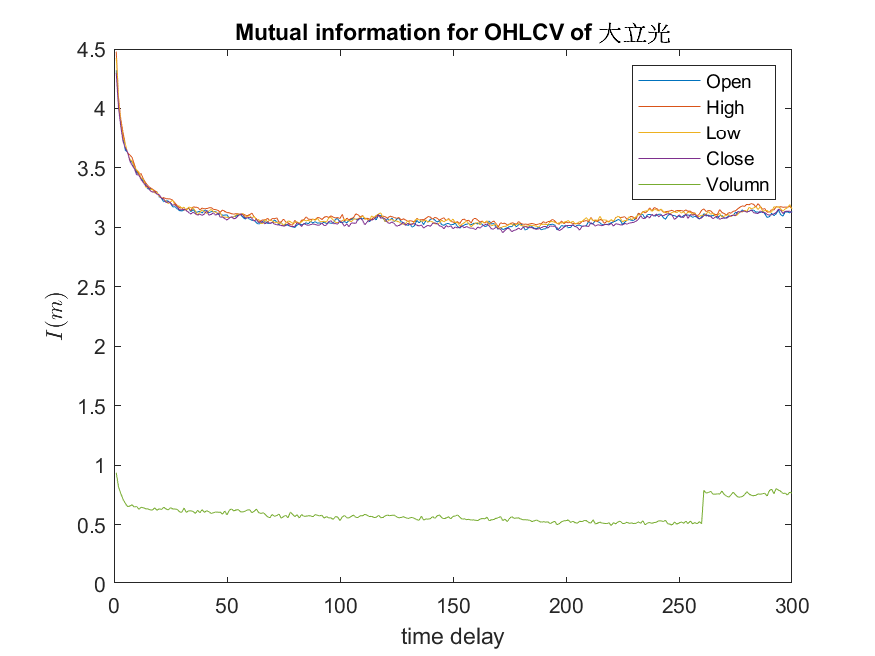
另外，交易量的曲線很快的趨近平緩，無法準確以人工觀察第一次最小值的位置。因此以下我們依然使用內建函式與人工判斷兩方法分析，並最終使用內建函式的數據進行後續embedding deimension的分析。

從表中可以發現，使用內建函式判斷的結果，OHLC的資料得到的time delay都十分相近，因此可以說明OHLC四項資料很可能受到同一個系統所影響。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 開盤價 | 最高點 | 最低點 | 收盤價 | 交易量 |
| 內建函式 | 72 | 71 | 73 | 67 | 4 |
| 人工判斷 | 70 | 70 | 70 | 70 | 10 |

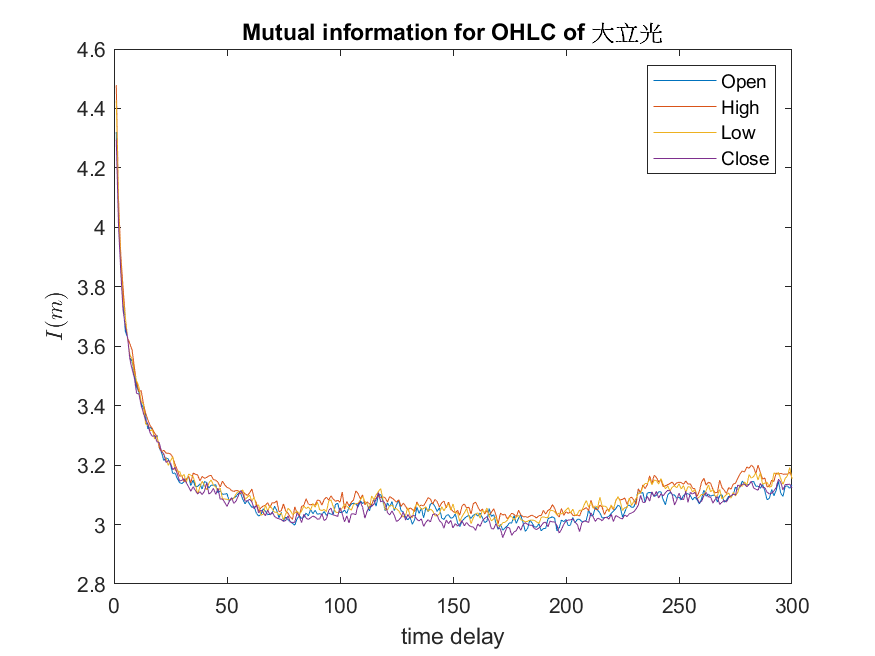
表二：中華電各項資料的time delay選取值

1. 大立光



圖六：大立光OHLCV資料的mutual information

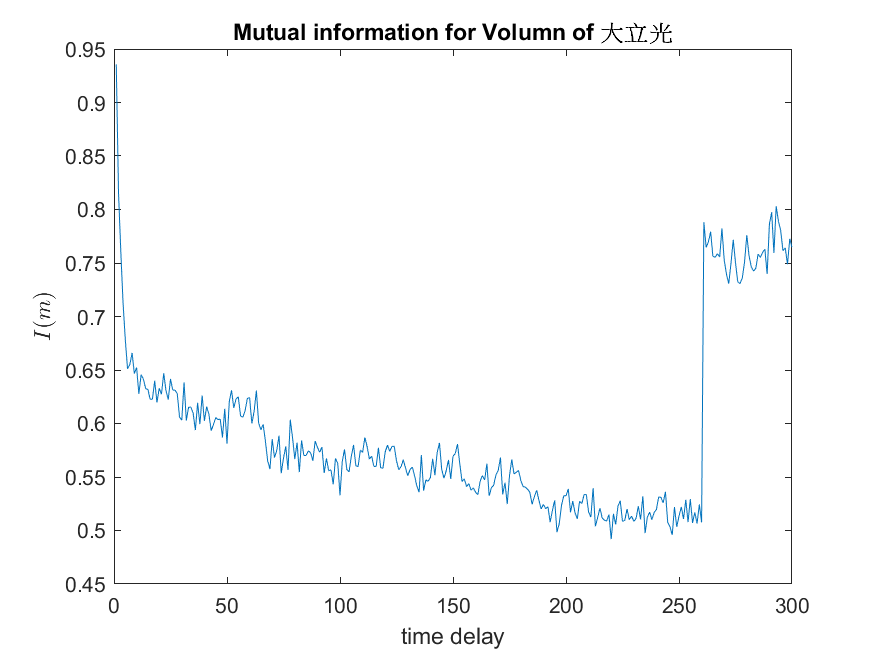
圖六中依然可以發現到，交易量和OHLC之間的差距甚大。由於圖六中y軸的尺度較小，不易觀察OHLC資料的細節，因此以下用另外用兩張圖表觀察OHLC資料和交易量資料的特徵。



圖七：大立光OHLC資料的mutual information

可以發現，OHLC資料在下降後，稍微明顯地上升，因此較容易觀察到第一次的最小值，故我們採用人工判斷的數據。

但觀察交易量資料的mutual information時，我們發現了較為困惑的現象。



圖八：大立光交易量資料的mutual information

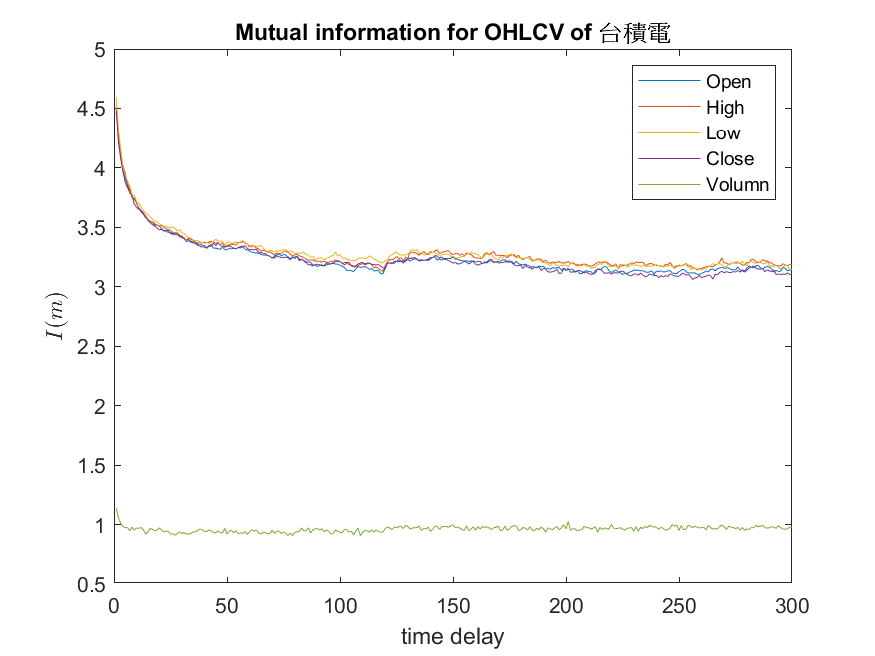
由圖八中可以發現，資料的mutual information的尺度很小，因此看起來會是快速波動的樣子。可以發現在time delay為260左右，同樣地出現了向上的斷層。值得注意的是，資料的mutual information不斷地下降，難以找到適當的第一次最小值，最終依據time delay為50左右時有微幅上升，而選擇time delay為50。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 開盤價 | 最高點 | 最低點 | 收盤價 | 交易量 |
| time delay | 80 | 79 | 75 | 78 | 50 |

表三：大立光各項資料的time delay選取值

從表中依然可以發現，OHLC資料得到的time delay都類似。

1. 台積電

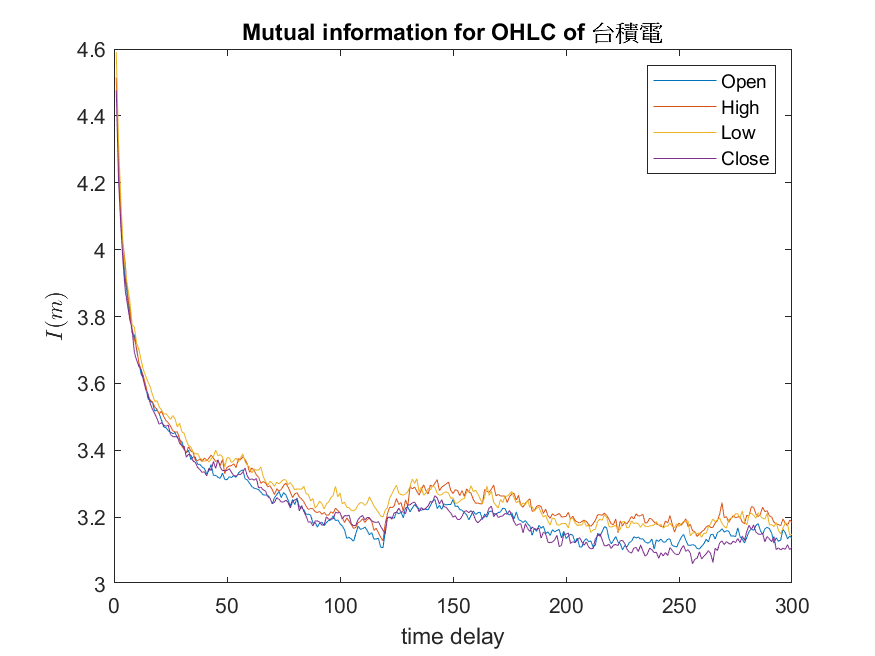


圖九：台積電OHLCV資料的mutual information

圖九中依然可以發現到，交易量和OHLC之間的差距甚大。

由於圖九中，y軸的尺度較小，不易觀察OHLC資料的細節，因此以下用另外用一張圖表觀察OHLC資料和交易量資料的特徵。

另外，值得注意台積電的交易量資料中，並沒有在time delay為260時出現向上的斷層。



圖十：台積電OHLC資料的mutual information

可以發現，OHLC資料在下降後，有明顯地上升，因此很容易觀察到第一次的最小值，故我們採用人工判斷的數據。

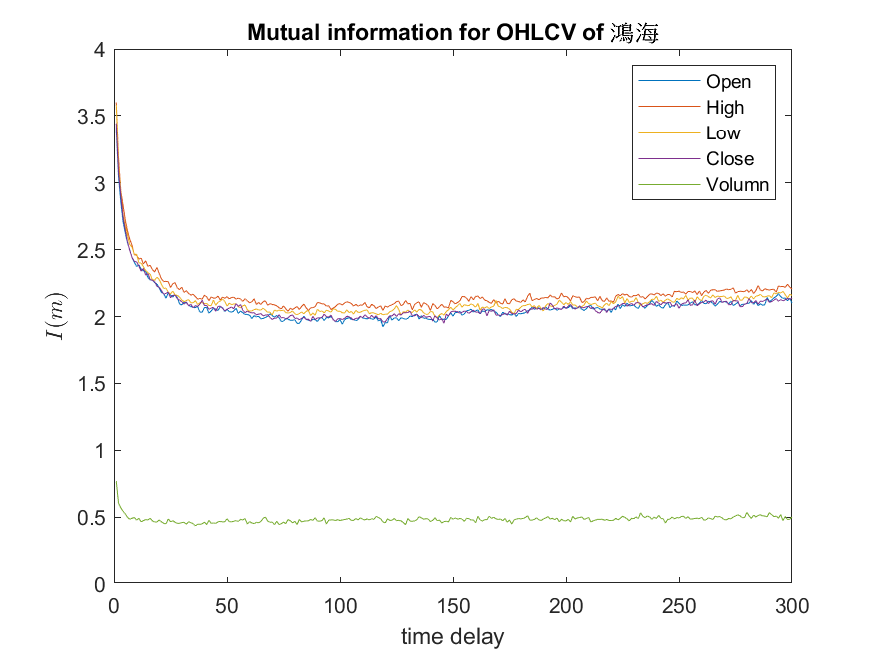
關於交易量資料的mutual information，則與前述股票類似。

從表中依然可以發現，OHLC資料得到的time delay都類似。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 開盤價 | 最高點 | 最低點 | 收盤價 | 交易量 |
| time delay | 119 | 119 | 119 | 121 | 79 |

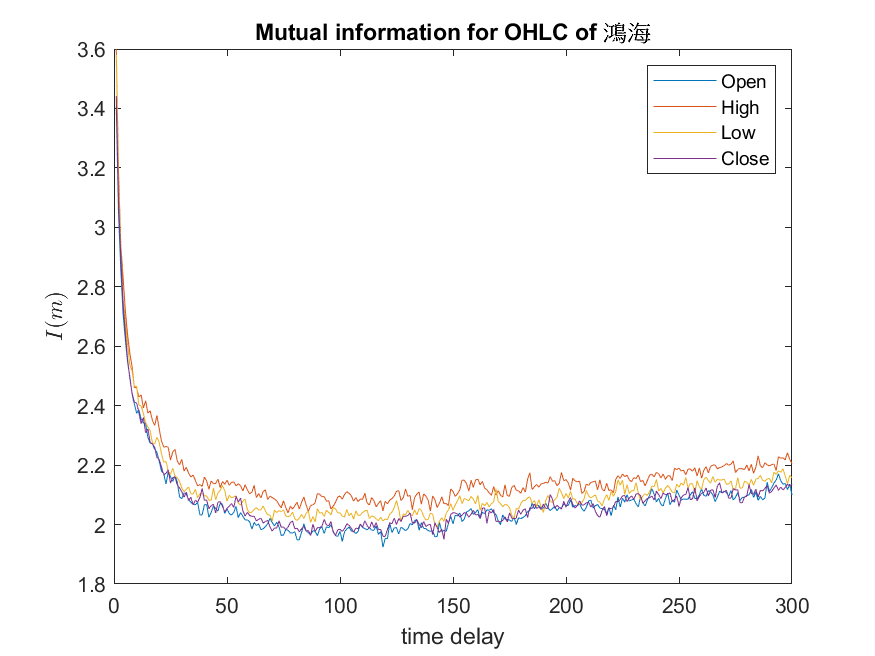
表四：台積電各項資料的time delay選取值

1. 鴻海



圖十一：鴻海OHLCV資料的mutual information

圖十一中依然可以發現到，交易量和OHLC之間的差距甚大。由於圖十一中，y軸的尺度較小，不易觀察OHLC資料的細節，因此以下用另外用一張圖表觀察OHLC資料和交易量資料的特徵。另外，值得注意鴻海的交易量資料中，仍然沒有在time delay為260時出現向上的斷層。



圖十二：鴻海OHLC資料的mutual information

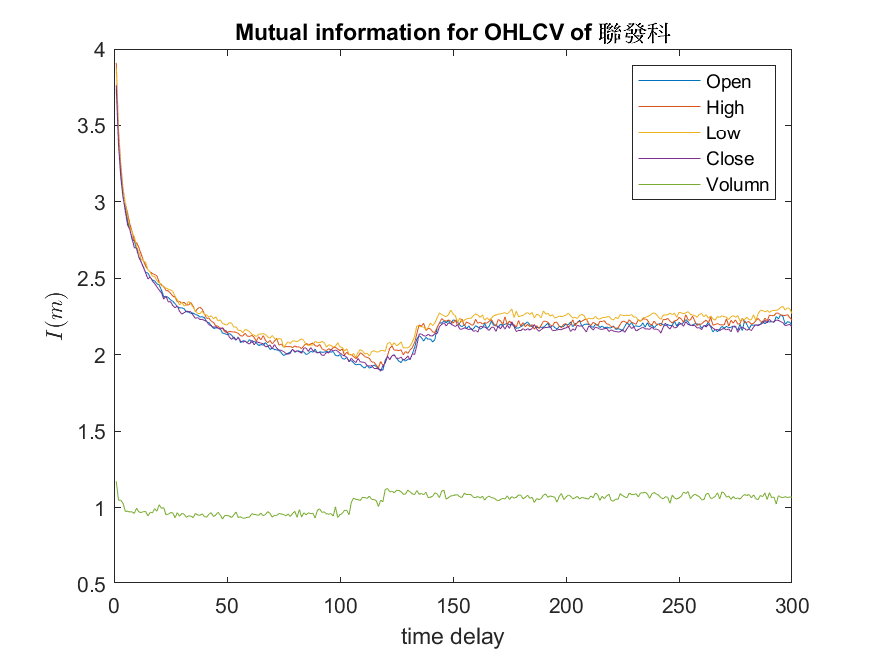
可以發現，OHLC資料在下降後，於time delay為80附近又開始上升，因此很容易觀察到第一次的最小值，故我們採用人工判斷的數據。而關於交易量資料的mutual information，則與前述股票類似。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 開盤價 | 最高點 | 最低點 | 收盤價 | 交易量 |
| time delay | 81 | 86 | 70 | 87 | 36 |

表五：鴻海各項資料的time delay選取值

從表中依然可以發現，OHLC資料得到的time delay都類似。

1. 聯發科



圖十三：聯發科OHLCV資料的mutual information

圖十三中依然可以發現到，交易量和OHLC之間的差距甚大。就可以明顯看出，OHLC的資料在下降後就開始上升，並趨於穩定，因此第一次的最小值可以清楚被觀察到。

　　另外，雖然在time delay為260時出現向上的斷層，但可以發現在time delay為100左右時，出現了突然突然向上的趨勢。

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 開盤價 | 最高點 | 最低點 | 收盤價 | 交易量 |
| time delay | 119 | 117 | 107 | 118 | 48 |

表五：聯發科各項資料的time delay選取值

1. 綜合分析

　　整理所有股票資料的time delay，標註上產業類別後可以得到下表：

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | 產業 | 開盤價 | 最高點 | 最低點 | 收盤價 | 交易量 |
| 台塑化 | 油電燃氣 | 59 | 34 | 34 | 35 | 5 |
| 中華電 | 通信網路 | 72 | 71 | 73 | 67 | 4 |
| 大立光 | 光電 | 80 | 79 | 70 | 87 | 36 |
| 台積電 | 其他電子 | 119 | 119 | 119 | 121 | 79 |
| 鴻海 | 半導體 | 81 | 86 | 70 | 87 | 36 |
| 聯發科 | 半導體 | 119 | 117 | 107 | 118 | 48 |

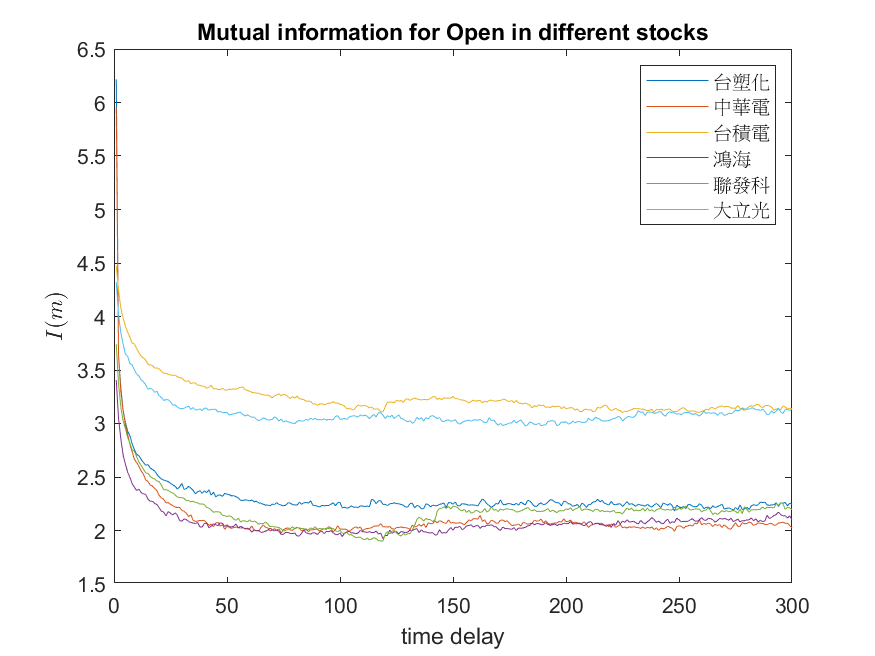
表六：股票各項資料的time delay選取值

其中台塑化及中華電的time delay是由matlab內建函式計算得到的。

　　個股來看，OHLC的資料所得到的time delay都差距不大，因此可以推斷OHLC的資料可能是由相同的系統所控制的。而交易量的資料，則明顯與OHLC資料相差較大，可能是由不同的系統參數所控制。

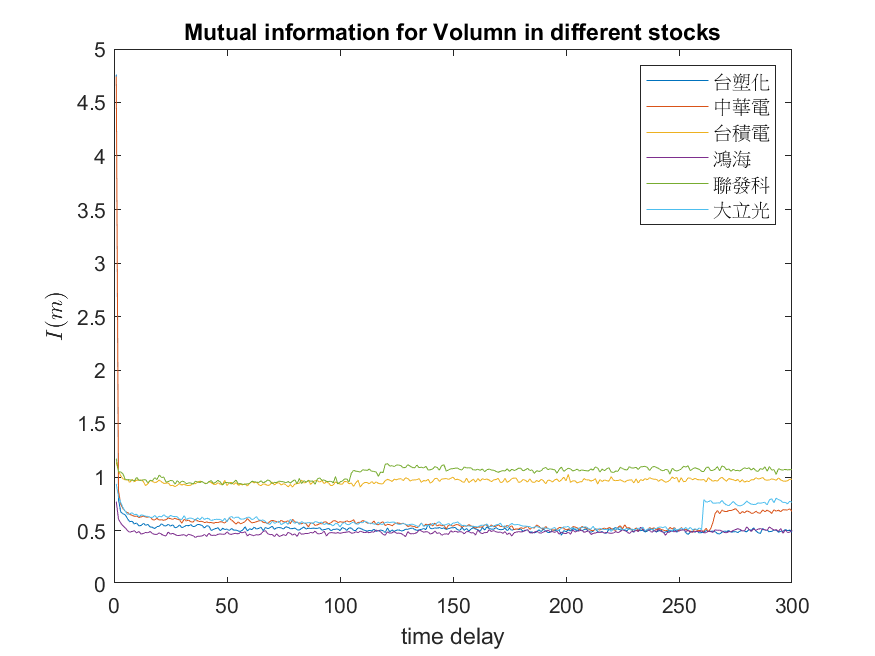
　　整體來看，可以發現產業類別和time delay的關聯較小。例如鴻海和聯發科皆是半導體產業，但兩者的time delay卻十分不同，反而是大立光和鴻海或是台積電和聯發科，反而time delay較為相近。

　　接下來我們對不同股票的同一種資料進行分析，由於分別對開盤價、最高點、最低點和收盤價資料進行分析時，圖形長相都十分類似，以下只以不同股票的開盤價資料進行比較。



圖十四：不同股票間開盤價資料的mutual information

　　由圖十四可以發現，台積電和大立光的資料自相關度較為不好，而另外四支股票較好。



圖十五：不同股票間交易量資料的mutual information

　　由圖十五可以發現，大立光和中華電在time delay為260的地方，出現了向上的斷層，而聯發科則是在time delay為100的地方出現。

綜觀以上分析，可得以下結論：

1. time delay值與股票的產業性質未有太大的關聯性

2. 在六種股票中，交易量的time delay均較OHLC來得低，由此推測交易量受到外在環境影響的反映較其他參數快。

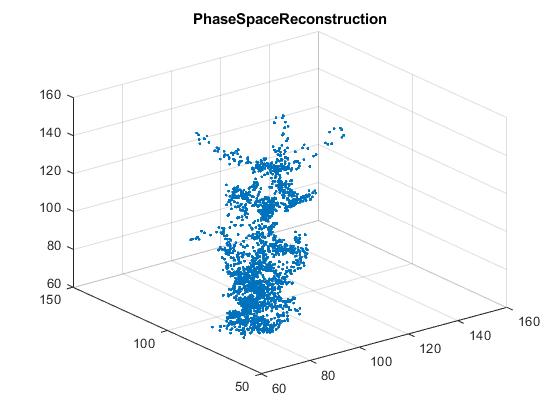
**Embeddig Dimension**

**Phase Space Reconstruction**

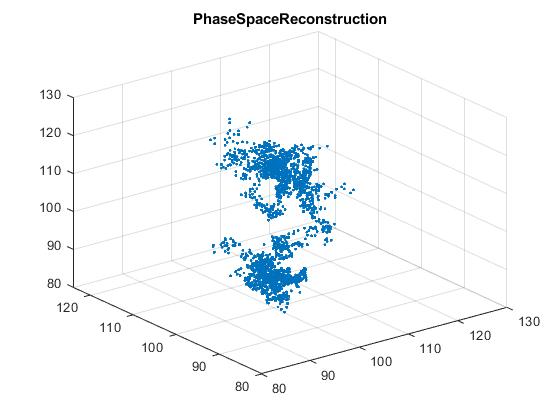
透過Mutual information與Embedding dimension得到每支股票的time delay以及dimension後，套用Matlab的內建phaseSpaceRecostruction函式，計算每支股票OHLCV重建後的序列XR，再作圖繪出每個資料的三維Phase plot，來觀察是否具有混沌現象。

由於六支股票各自的OHLC所得的圖形均相當類似，因此以下僅針對六股開盤價與交易量的圖形進行分析探討。

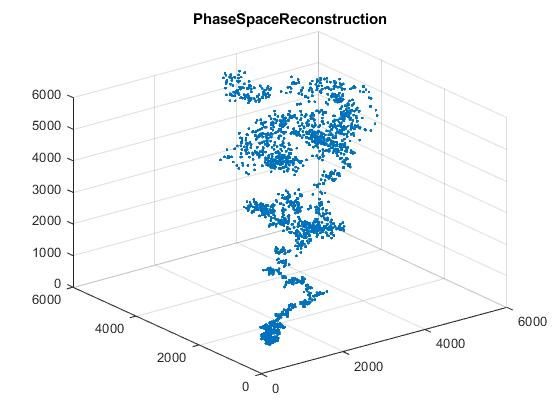
一、開盤價



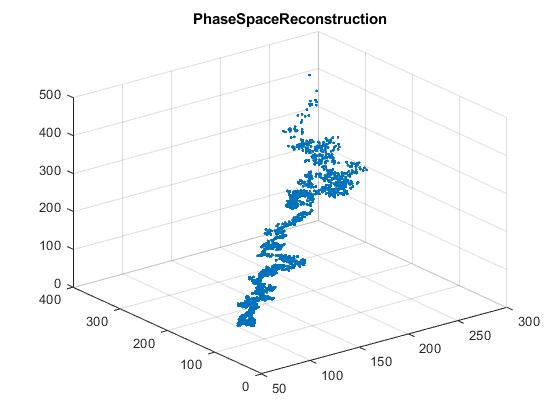
台塑化開盤價的phase space recostruction



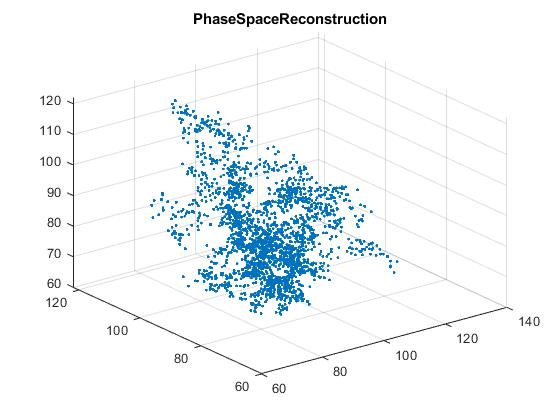
中華電開盤價的phase space recostruction



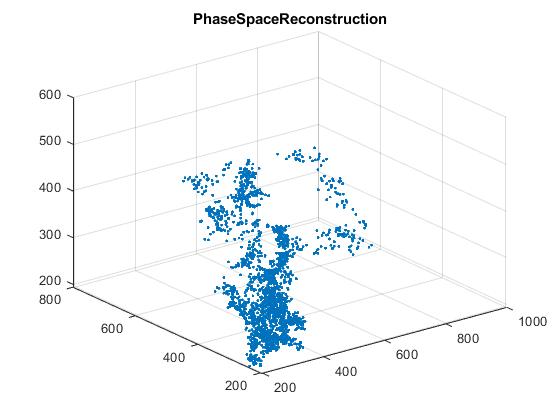
大立光開盤價的phase space recostruction



台積電開盤價的phase space recostruction



鴻海開盤價的phase space recostruction

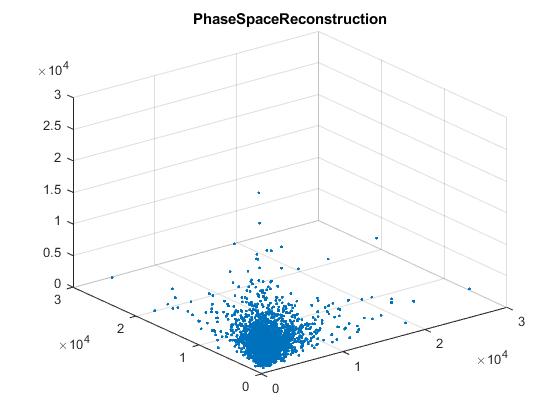


聯發科開盤價的phase space reconstruction

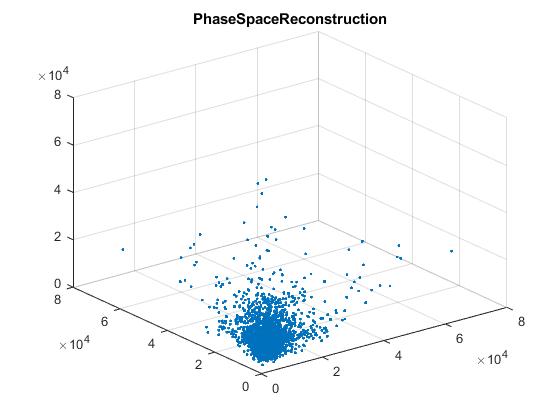
從上方六張圖中可以看出，每支股票reconstruction後每點所圍出來的圖形均不太一樣，相當獨特，像是台塑化的圖形類似鳳梨形狀、大立光與台積電類似龍捲風形狀、鴻海類似帽子狀等等，也由此可以看出，各支股票的OHLC，均有混沌現象發生。

二、交易量

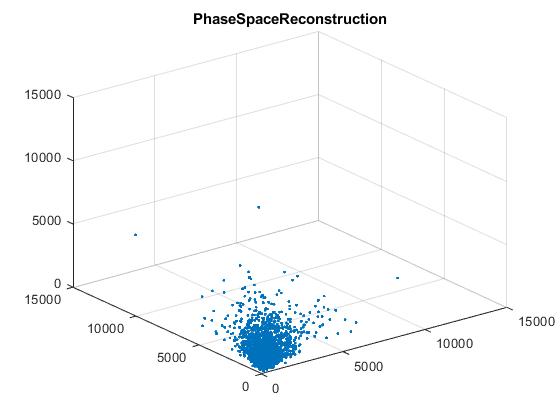
六支股票交易量的phase space recostruction結果如下：



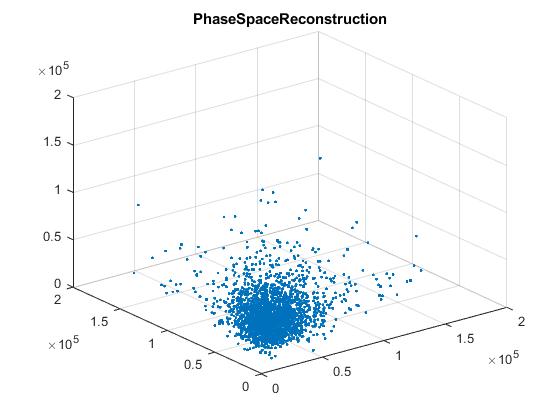
台塑化交易量的phase space recostruction



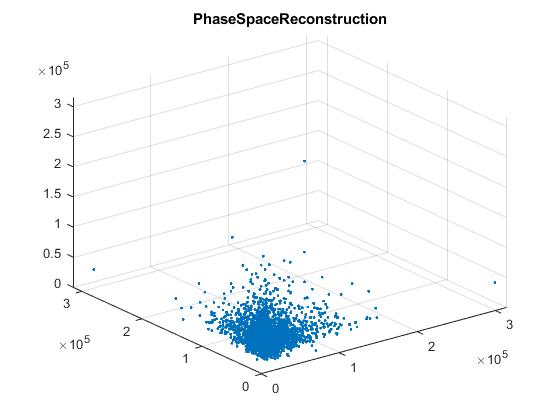
中華電交易量的phase space recostruction



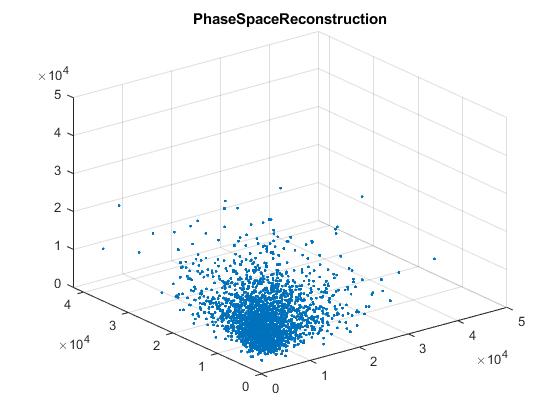
大立光交易量的phase space recostruction



台積電交易量的phase space recostruction



鴻海交易量的phase space recostruction



聯發科交易量的phase space reconstruction

觀察六支股票交易量的reconstruction圖形，均呈現橢圓形密集分布，類似於隨機分布的狀態，未有明顯的混沌現象。

三、小結

如先前提到的，由於六支股票的OHLC的維度均相同，在動態系統裡可視為相同變數，因此phase space reconstruction出來的結構也類似；而不同的股票OHLC數值大小不同，指數每日的變動幅度亦不同，可以視為不同的動態系統，因此會得到不同的混沌圖形。

至於在成交量的部分，六支股票的圖形均呈現隨機分布，這可能是由於各股的成交量較不受股票本身的結構影響；此外，交易量也會因當日市場的活躍度、投資人的購買能力等因素有關，總體受變動的因素較多且較複雜，導致phase space reconstruction後的結果呈現隨機分布的狀態。

4 References

<https://www.cnyes.com/twstock/ps_historyprice.aspx>

Kunsch, P., Theys, M., Chevalier, A. and Lacopetta, J. P. (2000). A System Dynamics Model

of Stock Price Movements. Decision making: Recent development and world-wide applications Applied Optimization Volume 45: 157-171. doi: [10.1007/978-1-4757-4919-9\_11](https://www.researchgate.net/deref/http%3A%2F%2Fdx.doi.org%2F10.1007%2F978-1-4757-4919-9_11)