主成分分析

Step 1 : 利用Eigenvalues決定取前幾組PC作分析

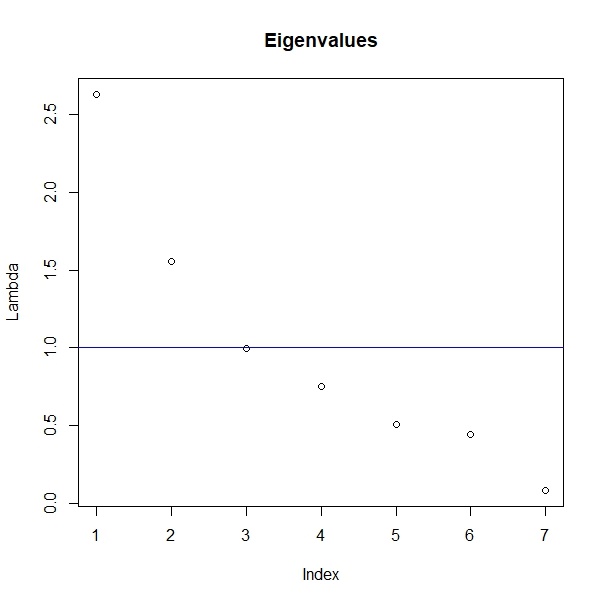
為了瞭解什麼樣的因素跟我們的目標有最大的關係，我們使用標準化後的資料進行主成分分析，首先計算相關係數矩陣以求得特徵值、特徵向量：

　　接著要決定取前幾組特徵值，可以先算出各特徵值可以變異的百分比，整理的表格如下：

表一：特徵值及其解釋變異的百分比

|  |  |  |  |
| --- | --- | --- | --- |
| Component | Eigenvalue | Proportion | Cumulative |
| 1 | 2.630 | 0.377 | 0.377 |
| 2 | 1.555 | 0.223 | 0.600 |
| 3 | 0.996 | 0.143 | 0.743 |
| 4 | 0.754 | 0.108 | 0.851 |
| 5 | 0.508 | 0.073 | 0.924 |
| 6 | 0.446 | 0.064 | 0.988 |
| 7 | 0.085 | 0.012 | 1 |

圖一、特徵值陡坡圖



根據表一顯示第一組主成分可以解釋大約37.7%的總變異，前兩組主成分可以解釋60%的變異，前三組則有74.3%的變異解釋力。另外，也可透過特徵值的陡坡圖(圖一)看出，在第三組主成分之後坡度才逐漸趨緩，因此我們選擇取前三組來進行本次主成分分析。

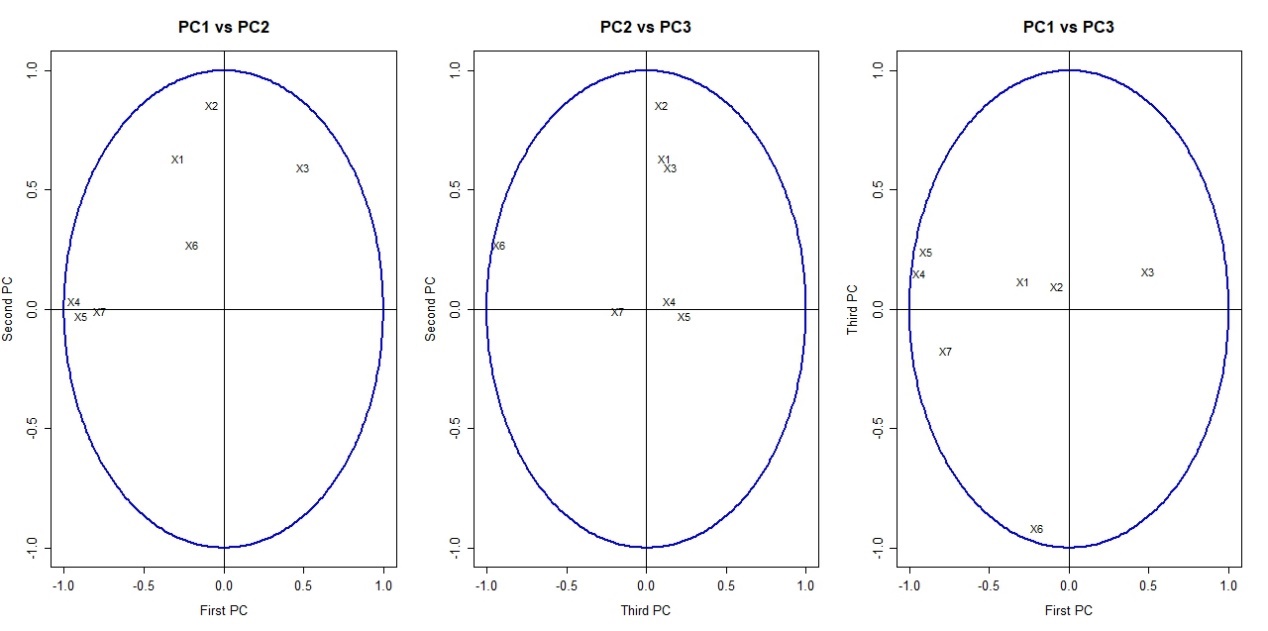
Step 2 : 利用Eigenvectors計算PC Scores

前三組主成分分數：

表二、前三組主成分與變數關係

|  |  |  |  |
| --- | --- | --- | --- |
|  | PC1 | PC2 | PC3 |
| 溫度 | -0.284 | **0.632** | 0.115 |
| 降雨量 | -0.071 | **0.857** | 0.095 |
| 相對溼度 | 0.499 | **0.594** | 0.157 |
| 住宅 | **-0.935** | 0.035 | 0.149 |
| 服務業 | **-0.894** | -0.028 | 0.240 |
| 農林漁牧 | -0.197 | 0.269 | **-0.918** |
| 工業 | **-0.770** | -0.006 | -0.174 |

圖二、前三組主成分與變數相關係數圖



Step 3 : 主成分分析

　　PC可以解釋為每個主成分與變數之間的關係，當PC離0越遠(正的越大；負的越小)表示他跟該變數間存在強烈的關係。以下由表二、圖二來分別說明三組主成分分析。

第一組主成分分析(PCA1)：第一組主成分與住宅、服務業、工業有高度負相關，表示這這三個變數會有同向性的變動，並且三者上升時回帶動PC1 Score下降。從表二可以看出確實非常貼近半徑為1的相關係數圓圈邊界，證實它們在PC1有高度負相關。因此，PC1 可以看作解釋住宅用電、服務業與工業用電的主成分。

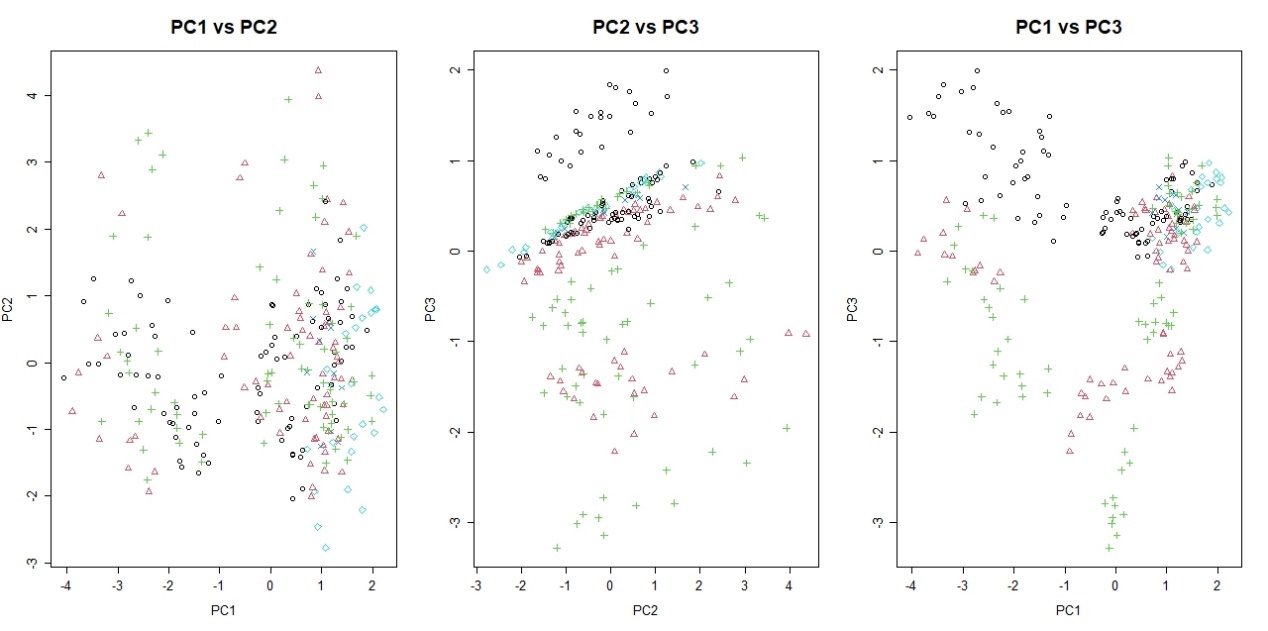
第二組主成分分析(PCA2)：第二組主成分與溫度、降雨量、相對濕度有有中到高度的正相關，表示當溫度、降雨量、相對濕度上升時，PC2 Score才會隨著上升。從圖二的左邊及中間兩張相關係數圖可以驗證第二組主成分最有關聯的是三個變數。因此，第二組主成分可以被定位為跟天氣有關連的主成分。

第三組主成分分析(PCA3)：第三組主成分僅與農林漁牧有高度負相關，PC3可以被看做測量農林漁牧用電量有多低的指標。根據圖二的中間及右邊的圖能夠得到一樣的結果。

Step 4 : 比較各地區分布差異

　　由於本資料為各地區的天氣與用電情況，我們考慮以台灣北、中、南、東及離島為區分，觀察天氣與用電是否存在地域性的差異。

圖三、PC散佈圖



○：北部 △：中部 ＋：南部　×：東部　◇：離島（金門、連江）

　　由圖三左圖無法看出地區上有明顯的差異。中間的圖可以看到各地區在PC2與PC3之間存在正向關係，另外，對PC3而言，南部與北部存在較明顯的差異，這個部分由右圖也可以觀察到，表示南部與北部在農林漁牧的用電量上存在較明顯的差距。右圖也可以看到離島與東部在PC1分佈較集中在右側，表示離島與東部在住宅、服務業、工業用電上較為類似。

典型相關分析

Step 1 : 用相關係數分析兩類變數關聯

　　為了研究台灣天氣與用電量是否有關聯，我們首先將7個變數拆分為天氣與用電量兩大類：, ，接著生成7個變數的相關係數矩陣，並直觀觀察各自類別與類別之間的相關性。

　　計算出相關係數矩陣後，可以先觀察左上角的部分，天氣之間皆呈現正相關係，但彼此之間都沒有特別明顯的高度正相關；而右下角部分，用電量之間可以發現住宅用電與服務業用電有高度正相關、與工業用電也有中等程度以上的正相關；服務業用電與工業用電有中度正相關，農林漁牧用電則沒有特別跟其他產業的用電量有太明顯的關聯。另外，比較兩類別之間的關聯以右上或左下的矩陣為主，大致來看天氣與用電都沒有看到中度以上的關聯性。溫度與四種用電量的比較下，它與住宅用電是最有相關；降雨量則跟四種用電量沒有特別明顯的關聯性；相對濕度與四種用電皆有低度負相關，其中與服務業用電有稍高的負相關。

Step 2 : 典型相關分析

表一、典型相關變數與典型相關

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |
|  | -0.32 | -0.47 | 0.97 | 0.46 |  | -0.42 | -0.40 | -0.20 | -0.24 |
|  | 0.17 | 0.77 | 0.32 | 0.17 |  | 1.72 | -1.38 | 0.63 | -0.77 |
|  | 1.01 | -0.72 | 0.37 | 0.03 |  | 1.86 | -1.68 | -0.77 | -0.02 |

第一組典型相關變數是：

其對應的典型相關。

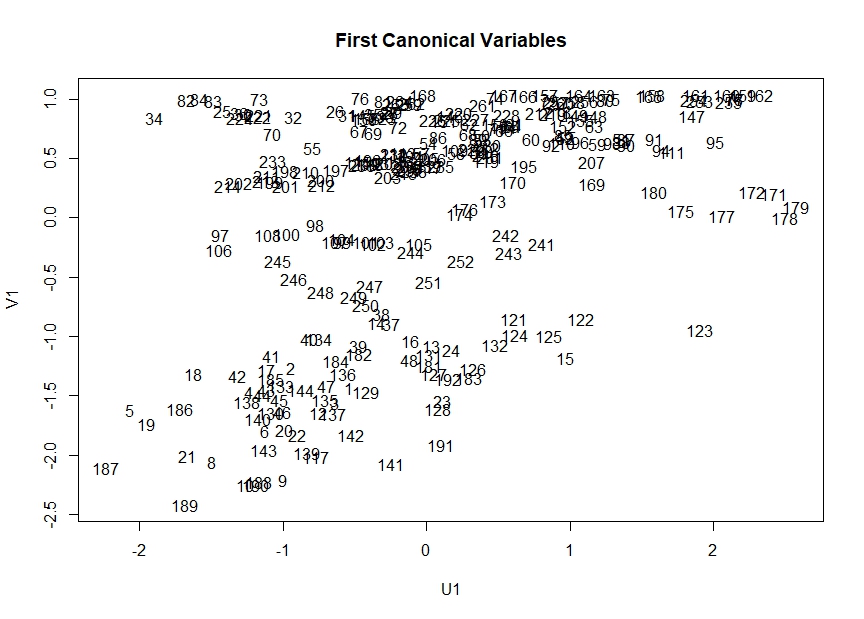
　　由表一可以看出第一組的canonical correlation為0.46，表示天氣與用電量有中度的正相關。第二組，表示透過第二組典型相關僅能發現天氣與用電量的低度相關。因此，以下我們只考慮第一組的典型相關變數。

　　由第一組典型相關分析可得知相對濕度在天氣類別的權數相當高，表示對天氣的變動而言，相對溫度佔有重要的解釋力；住宅用電與服務業用電在用電量類別的權數較高，表示在用電量上住宅用電與服務業用電為較強的變動要素。兩類別比對可推測：當相對濕度較高時，住宅用電與服務業用電會較低。而顯示天氣變數與用電量變數之間存在部分的重疊。

表二、原變數與典型相關變數的相關性

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | 樣本典型相關變數 | |  | 樣本典型相關變數 | |
|  |  |  |  |  |  |
| 溫度 |  |  | 住宅 |  |  |
| 降雨量 |  |  | 服務業 |  |  |
| 相對溼度 |  |  | 農林漁牧 |  |  |
|  |  |  | 工業 |  |  |

圖一、第一組典型相關變數散佈圖



Step 3 : 檢定典型相關係數

　　由於本資料為大樣本，因此需要檢定典型相關係數是否顯著，根據表三可以得到當假設全部時，結果為拒絕虛無假設，表示至少有一個。所以需要接著假設，其餘，結果為不拒絕，表示若本資料服從多變量常態，則第一組的Canonical correlation不為零。

表三、檢定結果

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | 虛無假設 | 檢定統計量(Bartlett correction) | 卡方分配 | 結果 |
| 1 |  |  |  | Reject |
| 2 |  | 7.829 |  | Do not reject |

研究背景

　　世界氣象組織(WMO)指出在最糟的情況下，全球年均溫將比工業化前高出1.6度；我國中央氣象局也從國內各個氣象測站資料表示國內的年均溫呈現一路上升的趨勢。2022年夏天全台用電持續創新高，其原因除了是氣溫高居不下，還加上疫情的影響，使民眾多待在家裡使用空調，拉高住宅用電量，且產業用電並未下降，形成用電量飆高的現象。另外，台灣在2022年初也由國發會宣布台灣2050淨零碳排放的目標，並規劃四大轉型策略及兩大基礎，能源的轉型是其中的一大方向，預期未來各項技術成熟能源的需求會趨緩，但可能更多產業朝電氣化發展進而演變為電力需求提高。如何在全球暖化、氣候變遷這些不可逆的情況下使台灣用電量維持穩定是未來台灣能源政策需要面臨的課題。

研究動機與目的

　　面對氣候的變遷，國家必須了解國人真實的用電概況才能制定有效率的減碳計畫，民眾需要了解用電量對環境的影響才會開始付諸減碳行動。而每年夏天總是有許多「今日氣溫飆破歷年最高」或是「今日尖峰負載創歷史新高」等等類似的新聞，究竟台灣氣候與用電量之間使否存在關聯性是本次研究想探究的問題。

　　本研究目的可整理為以下三點，

1. 了解台灣天氣資料與用電資料間的關聯性

2. 各地區天氣情況是存在差異

3. 各地區用電情況是存在差異

研究方法

1. 典型相關分析(Canonical Correlation Analysis)

2. 主成分分析(Principle Component Analysis)

3. 分群分析(Cluster Analysis)