

國立臺北大學 多變量分析報告

阿拉斯加氣候之多變量分析

王姿文¹

摘要

本報告為多變量分析報告，將分析由1981-2003年共23年的阿拉斯加資料。在此報告中，會藉由Principal Component Analysis以及Factor Analysis兩種維度縮減方法萃取資料資訊，並使用Clustering將維度縮減後的資料分群，最後以MANOVA作檢定來探討分群結果是否合適。

¹ 臺北大學經濟學系

目錄

壹、研究動機.....	2
貳、資料介紹.....	2
參、分析方法.....	3
一、階層式分群法 (Hierarchical Clustering).....	3
二、維度縮減法.....	4
(一) 主成分分析法 (Principal Component Analysis, PCA)	4
(二) 因子分析法 (Factor Analysis, FA).....	4
三、MANOVA.....	4
肆、實證分析.....	5
一、維度縮減前之多變量分析.....	5
二、運用維度縮減法於多變量分析.....	6
(一) 主成分分析法 (Principal Component Analysis, PCA)	6
(二) 因子分析法 (Factor analysis, FA).....	11
伍、結論.....	15
陸、致謝.....	15
參考文獻.....	15

壹、研究動機

氣候資料對於生態研究十分重要，因為生態環境指標的組成包含著氣候，因此研究阿拉斯加氣候能夠幫助解釋阿拉斯加的生物變遷以及生態變化。

貳、資料介紹

此報告的氣候資料共有 17 個變數，為阿拉斯加的年度氣候資料，由 1981-2003 年共 23 筆數據，圖 1 為資料原始型態。本資料是在一個名叫 Fundacion BBVA²的多變量數據網站取得，並在 Fundacion BBVA 找到關於本資料的文獻³。根據內容得知，其中變數 AO(Artic Oscillation)、NPI(North Pacific Index)為 Climate indices；Temperature、Rain 為 Meteorological variables；Ice coverage、Ice-free days 為 Various measures，也能從表 1 清楚明白各個變數。表 2 則為 17 個變數的 Measure scales。

	AO	AO_wint	AO_summ	NPI	NPI_spring	NPI_winter	Temp	Temp_summ	Temp_wint	Rain	Rain_summ	Rain_wint	Ice	Ice_JanJul	Ice_OctDec	IceCover	IceFreeDays
1981	-0.4346	-0.1683	-0.2410	-2.09	-0.15	-4.46	-3.9	8.8	-16.66	23.62	4.25	1.46	12.309	14.647	11.367	-0.64	140
1982	0.2977	-0.3750	0.3083	0.75	0.13	1.70	-4.7	10.5	-17.17	27.03	3.71	1.92	12.673	14.983	11.907	-1.65	144
1983	0.0319	0.1733	0.4653	-2.54	0.30	-6.44	-4.4	8.2	-17.75	28.75	5.65	0.68	12.493	14.735	11.573	-0.34	116
1984	-0.1917	0.2627	0.0240	-1.20	-0.23	-2.62	-7.0	8.7	-20.03	26.04	5.75	0.74	12.089	14.428	11.103	0.15	134
1985	-0.5192	-1.2667	0.2678	0.52	-0.43	1.11	-5.9	9.6	-15.15	27.28	3.62	1.24	12.208	14.758	11.153	-0.21	120
1986	0.0848	-1.8067	-1.8067	-1.84	-0.38	-4.11	-5.7	9.0	-14.65	25.32	6.01	0.91	12.404	14.627	11.690	-0.32	154
1987	-0.5442	-0.8537	-0.8537	-1.25	-0.30	-2.93	-5.2	9.6	-16.41	20.45	3.22	1.35	12.266	14.828	10.405	-1.43	152
1988	0.0402	-0.4450	-0.4450	-0.24	0.35	-0.77	-5.1	9.5	-16.63	22.45	3.60	1.70	12.094	14.532	11.647	0.19	142
1989	0.9500	2.6880	2.6880	1.35	-0.18	2.84	-4.8	10.7	-17.47	34.75	6.25	1.91	12.147	14.322	11.497	-1.16	159
1990	1.0241	1.2530	1.2530	0.91	0.33	2.00	-6.4	11.0	-22.81	37.49	5.27	1.40	11.888	14.342	11.310	0.16	149
1991	0.1970	0.3747	0.3747	0.99	0.20	2.34	-5.2	10.4	-20.46	22.00	2.66	2.31	11.914	14.315	11.150	0.53	154
1992	0.4366	1.0950	1.0950	-1.10	1.33	-2.45	-7.3	8.0	-20.02	24.94	3.35	1.42	12.226	14.295	11.640	0.65	137
1993	0.0792	1.7687	1.7687	-0.16	1.53	-0.15	-4.2	9.4	-19.06	34.06	4.17	1.52	12.108	14.567	11.480	-0.63	173
1994	0.5324	-0.4180	-0.4180	0.67	0.23	1.46	-6.3	9.3	-15.46	34.52	7.31	1.38	12.176	14.465	11.437	0.00	148
1995	-0.2746	0.7230	0.7230	-0.02	0.53	-0.34	-4.9	10.8	-17.96	23.80	2.28	1.06	11.601	14.060	10.963	0.74	159
1996	-0.4565	-1.0547	-1.0547	-0.57	0.30	-1.13	-5.6	8.4	-19.15	26.09	3.89	1.58	11.948	13.980	11.030	-1.89	157
1997	-0.0398	-0.0963	-0.0963	-0.22	1.05	-0.23	-4.5	10.9	-18.06	34.87	6.07	1.20	11.831	14.230	10.987	-0.38	183
1998	-0.2709	-0.7783	-0.7783	-1.57	-1.73	-3.42	-3.8	9.6	-19.66	36.04	6.83	0.91	11.943	14.463	10.953	0.89	162
1999	0.1126	0.6483	0.6483	0.26	-0.73	0.55	-6.2	10.0	-19.84	21.34	4.07	1.33	11.876	14.388	10.990	0.43	133
2000	-0.0465	1.1297	1.1297	-0.29	-0.40	-0.45	-5.0	7.4	-19.70	30.28	5.93	1.56	11.661	14.063	10.757	0.67	156
2001	-0.1615	-1.3117	-1.3117	-0.95	-0.63	-1.82	-5.3	9.7	-11.32	25.73	4.40	1.59	11.771	14.243	10.783	-2.24	137
2002	0.0717	0.4543	0.0187	0.13	-0.18	0.30	-3.3	9.4	-20.00	26.31	3.68	1.40	11.568	14.072	10.803	0.78	203
2003	0.1521	-0.6453	0.0399	-1.67	-0.40	-3.84	-3.8	8.9	-16.89	31.98	4.52	1.38	11.563	14.090	10.587	-1.60	179

圖 1

表 1

VARIABLES

CLIMATE INDICES	AO(Artic Oscillation)、AO_wint、AO_summ、NPI(North Pacific Index)、NPI_spring、NPI_winter
METEOROLOGICAL VARIABLES	Temp(Temperature)、Temp_summ、Temp_wint、Rain、Rain_summ、Rain_wint
VARIOUS MEASURES	Ice、Ice_JanJul、Ice_OctDec、Ice coverage、Ice-free days

² <https://www.fbbva.es/microsite/multivariate-statistics/data.html>

³ <https://www.fbbva.es/microsite/multivariate-statistics/publications.html>

表 2

Variables	Measurement Scales
AO	Interval Scale
AO_wint	Interval Scale
AO_summ	Interval Scale
NPI	Interval Scale
NPI_spring	Interval Scale
NPI_winter	Interval Scale
Temp	Interval Scale
Temp_summ	Interval Scale
Temp_wint	Interval Scale
Rain	Ordinal Scale
Rain_summ	Ordinal Scale
Rain_wint	Ordinal Scale
Ice	Ordinal Scale
Ice_JanJul	Ordinal Scale
Ice_OctDec	Ordinal Scale
IceCover	Interval Scale
IceFreeDays	Ordinal Scale

參、分析方法

一、階層式分群法 (Hierarchical Clustering)

本研究採用聚合式階層分群法，是由樹狀結構的底層開始，將相似的資料點層層向上群聚而成。群聚的過程，以平均連結法 (Average) 來定義群間的距離。對於分群法來說，如何決定群的數目是相當重要的問題，R語言的套件

「pheatmap」能夠幫助我們客製化熱圖，透過視覺化，可以更清楚地判斷群間的關係；此外也能運用SAS幫助我們得知不同群間的摘要統計值，並以MANOVA驗證分群數量。平均連結法的公式如下：

$$d(C_i, C_j) = \sum_{a \in C_i, b \in C_j} \frac{d(a, b)}{|C_i||C_j|} \quad \text{式 1}$$

式 1 中， C_i 、 C_j 個別代表兩群， $|C_i|$ 、 $|C_j|$ 為群內資料點的數量。

二、 維度縮減法

(一) 主成分分析法 (Principal Component Analysis, PCA)

PCA是一個非監督式的維度縮減方法，藉由預測變數的線性組合，也就是主成分，解釋資料矩陣的變異結構，最後選擇可以解釋最大變異的主成分，達到維度縮減的目的。主成分分析法的公式如下：

$$PC_i = W_i^T X = w_{i1}X_1 + w_{i2}X_2 + \cdots + w_{ip}X_p, i = 1, 2, \cdots k \quad \text{式 2}$$

$$Var(PC_i) = W_i^T \Sigma_X W_i = \lambda_i, i = 1, 2, \cdots k \quad \text{式 3}$$

式 3 中 PC_i 為第 i 個主成分， X_1, X_2, \cdots, X_p 為預測變數， Σ_X 為預測變數的共變異矩陣，而 $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 0$ 為其特徵值， W_1, W_2, \cdots, W_k 為特徵向量，而預測變數 X_i 一般都會予以標準化，避免單位不同影響主成分分析的結果。

(二) 因子分析法 (Factor Analysis, FA)

FA是一個非監督式的維度縮減方法，藉由尋找潛在的因素，也就是因子，解釋資料矩陣的變異結構，最後選擇可以解釋最大變異的因子，達到維度縮減的目的。因子分析法的公式如下：

$$X - \mu = LF + \epsilon \quad \text{式 4}$$

$$Cov(X) = LL' + \Psi \quad \text{式 5}$$

$$Var(x_i) = l_{i1}^2 + \cdots + l_{im}^2 + \Psi_i \quad \text{式 6}$$

式 4 中 F 為因子(common factor)， X 為預測變數， ϵ 為誤差(specific factor)，而 l 為factor loading， L 為factor loading的矩陣， Ψ 為 ϵ 的共變異矩陣。而預測變數 X_i 一般都會予以標準化，避免單位不同影響因子分析的結果。

三、 MANOVA

MANOVA 是多變量變異數分析方法，藉由指定兩個或兩個以上反應變數的變異數分析，來處理多個母群體平均數間的比較，為一種檢定方法。MANOVA 的假設如下：

$$\begin{cases} H_0: \mu_1 = \cdots = \mu_g \\ H_a: \text{at least one } \mu_i \neq \mu_j, i = 1, 2, \dots, g, j = 1, 2, \dots, g, i \neq j \end{cases} \quad \text{式 7}$$

若是式 7 的虛無假設被拒絕，則至少有兩個母群體之間的平均不同，亦即這兩群體間有顯著差異。

肆、實證分析

一、維度縮減前之多變量分析

將資料中的 17 個變數設為 x1-x17 並跑出分析結果。圖 2(a)為維度縮減前之階層式分群熱圖，集群分析方法為平均法，欄為變數，列為觀測值。而圖 2(b)為圖 2(a)做完標準化後所產生的階層式分群熱圖，集群分析方法一樣為平均法。

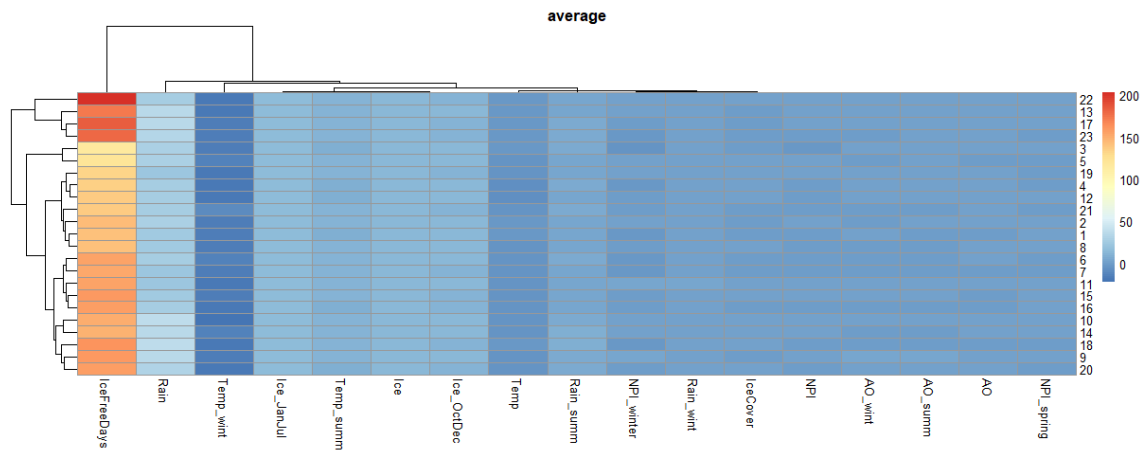


圖 2(a)

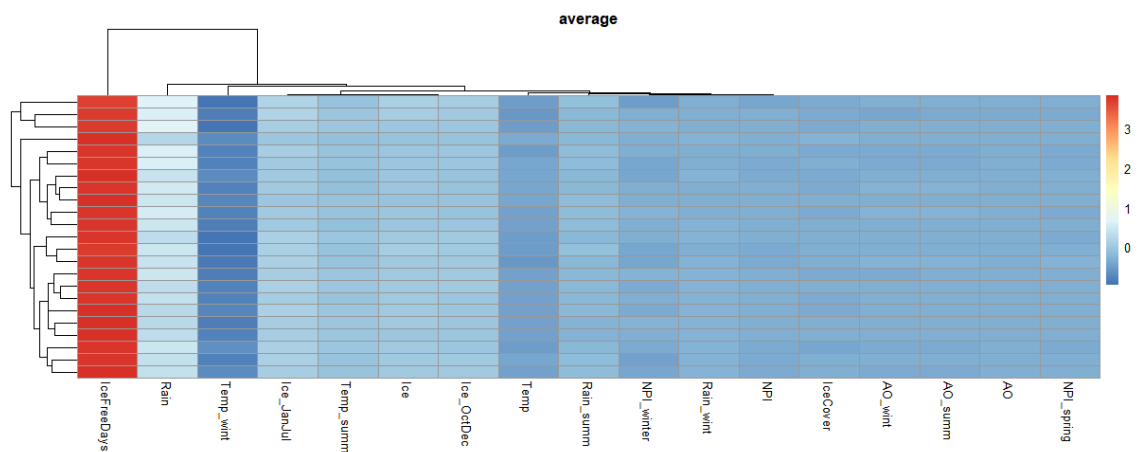


圖 2(b)

由圖 2(a)能看出 Ice free days 的值變化較大，而其他欄位在不同觀測值的變化不明顯，能些微看出大概能分三到四群。而圖 2(b)為標準化後再做熱圖，若是先做標準化則 23 個觀測值用眼睛看皆沒有明顯的分群。可能是因為這是一分在不同年度間的氣候資料，氣候變遷於短期內(23 年)變化並不大因此分群不明顯。

接著做 MANOVA 分析，表 3 為 MANOVA 分析的報表，根據報表做虛無假設：

$$\begin{cases} H_0: \text{無整體cluster效果} \\ H_a: \text{有整體cluster效果} \end{cases}$$

表 3

無整體 CLUSTER 效果假設的 MANOVA 檢定準則和 F 近似					
H = 類型 III SSCP 矩陣 - 針對 CLUSTER					
E = 誤差 SSCP 矩陣					
S=3 M=6.5 N=0.5					
統計值	值	F 值	分子自由度	分母自由度	Pr > F
Wilks' Lambda	0.00104680	1.72	51	9.7373	0.1816
Pillai's Trace	2.39826705	1.17	51	15	0.3830
Hotelling-Lawley Trace	68.32881405	6.06	51	2.5676	0.1051
Roy's Greatest Root	61.35597729	18.05	17	5	0.0023
附註: Roy Greatest Root 的 F 統計值是上限。					

根據表 3 的 Wilk' s Lamda 可以得知 P-value=0.1816，由於 SAS 預設的 α 值為 0.05，因此接受虛無假設，維度縮減前的資料作集群分析無整體 cluster 效果。

二、運用維度縮減法於多變量分析

由於本報告的氣候資料為一多變量資料，因此決定採用維度縮減法縮減資料維度來萃取資料的資訊，再做集群分析，最後以 MANOVA 作檢定來判斷分析效果好壞。

(一) 主成分分析法 (Principal Component Analysis, PCA)

將資料中的 17 個變數設為 x1-x17 並跑出分析結果。圖 3 為 PCA 的陡坡圖，橫軸為維度(主成分)，縱軸為解釋變異百分比，表 4 則為 PCA 的累計解釋變異百分比。

表 4

	Prin1	Prin2	Prin3	Prin4	Prin5	Prin6	Prin7	Prin8
累計解釋變異百分比	27.9%	45.6%	58.8%	69.6%	77.3%	83.1%	87.7%	91.4%

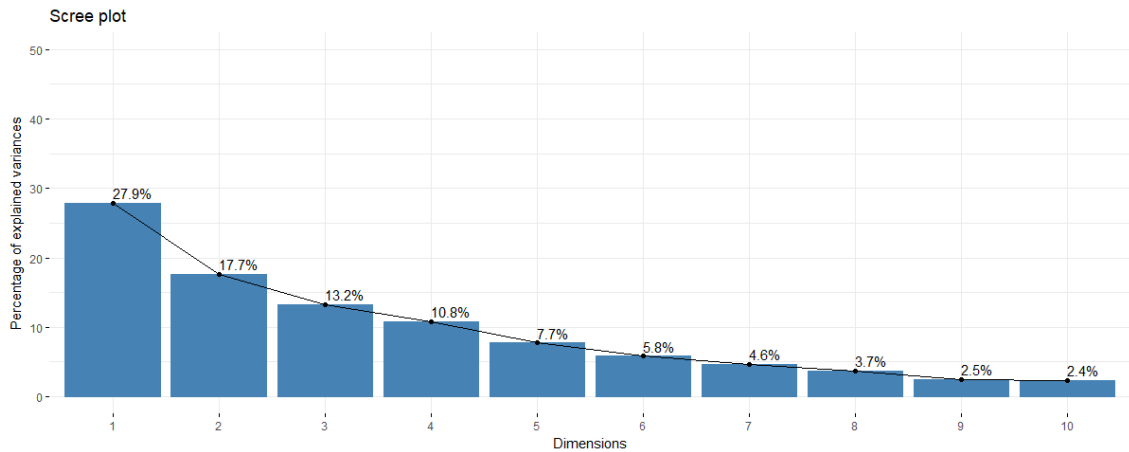


圖 3

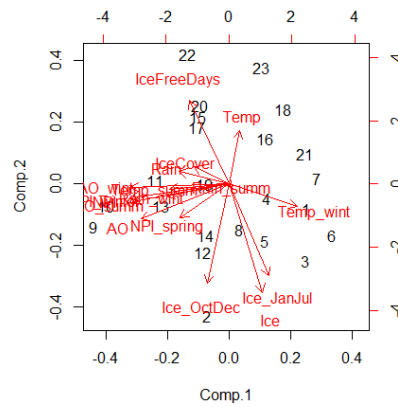


圖 4

由表 4 可以得知若希望解釋變異程度超過 90%則需要用到第 8 個主成分，也就是八個維度，然而在分群上，若是維度過多則分群特性難以解釋，因此後續集群分析上僅會採用主成分 1 及主成分 2。而這兩個主成分一共可以解釋 45.6%的變異數，主成分 1 可以解釋 27.9%而主成分 2 可以解釋 17.7%，圖 4 為 23 個觀測值投射在二維空間的散佈圖，橫軸為主成分 1，縱軸為主成分 2，紅色向量為 17 個變數分別在兩個主成分的貢獻度(重要程度)。

藉由 17 個變數分別在主成分 1 以及主成分 2 的貢獻度不同，我們能夠為主成分作命名。圖 5(a)為不同變數在主成分 1 的貢獻度百分比，圖 5(b)為不同變數在主成分 2 的貢獻度百分比橫軸為變數，縱軸為貢獻百分比，當變數的貢獻百分比高於水平虛線，則代表該變數有一定重要性。圖 6 則為不同變數在二維空間貢獻度的向量圖，橫軸為主成分 1，縱軸為主成分 2，附圖色條為貢獻程度，色條由藍到紅，顏色越紅代表貢獻度越高，顏色越藍則代表貢獻度越低。

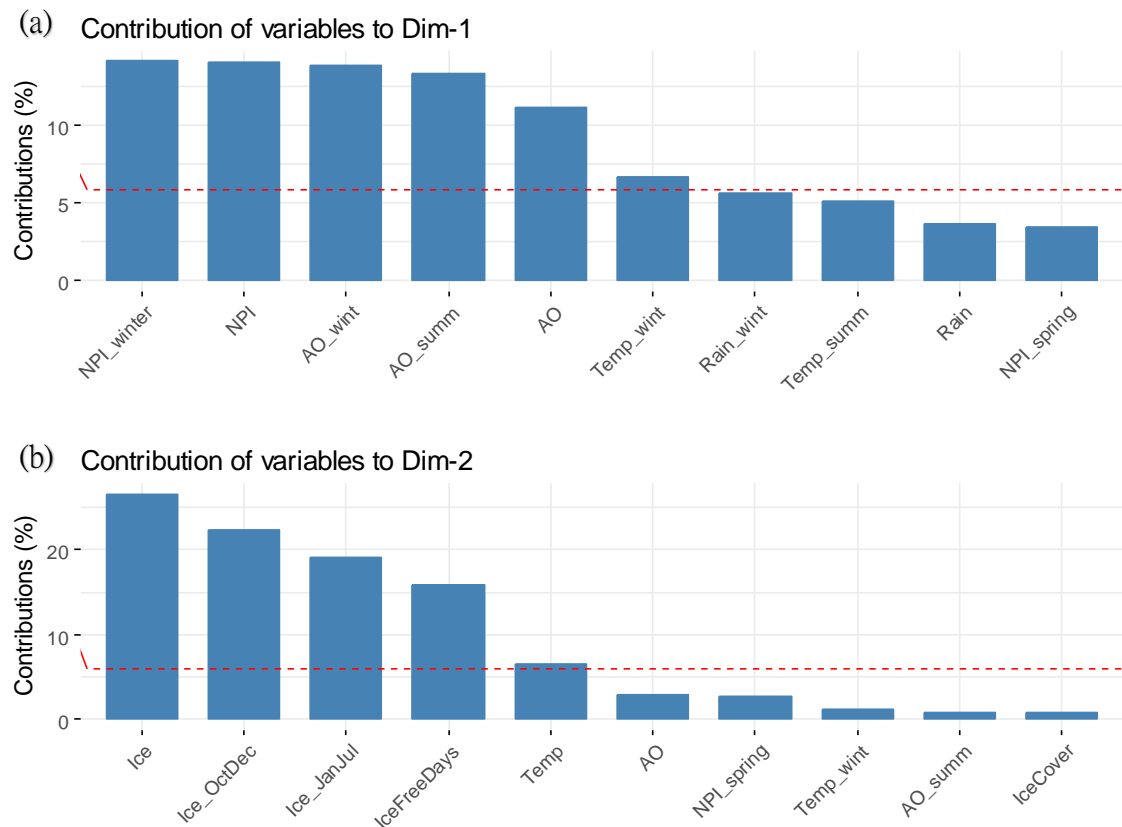


圖 5

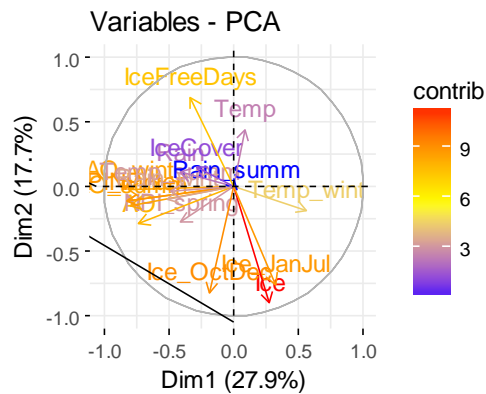


圖 6

由圖 5、圖 6 能將主成分 1 命名為 Climate indices Component，主成分 2 命名為 Various measures Component，命名規則能參照表 1 的變數介紹得知。

接下來將投射至 Climate indices Component 及 Various measures Component 的觀測值作集群分析。圖 7 為集群分析的階層式分群樹狀圖，分群方法為平均法，橫軸為觀測值，縱軸為群集間的平均距離，而我根據樹狀圖決定將資料分成四群。而表 5 為四個群集中 Climate indices Component 及 Various measures Component 的平均值。

表 5

	Climate indices Component	Various measures Component
<i>Cluster1</i>	0.054	-2.14
<i>Cluster2</i>	-0.07	1.618
<i>Cluster3</i>	2.814	0.438
<i>Cluster4</i>	-2.8	0.616

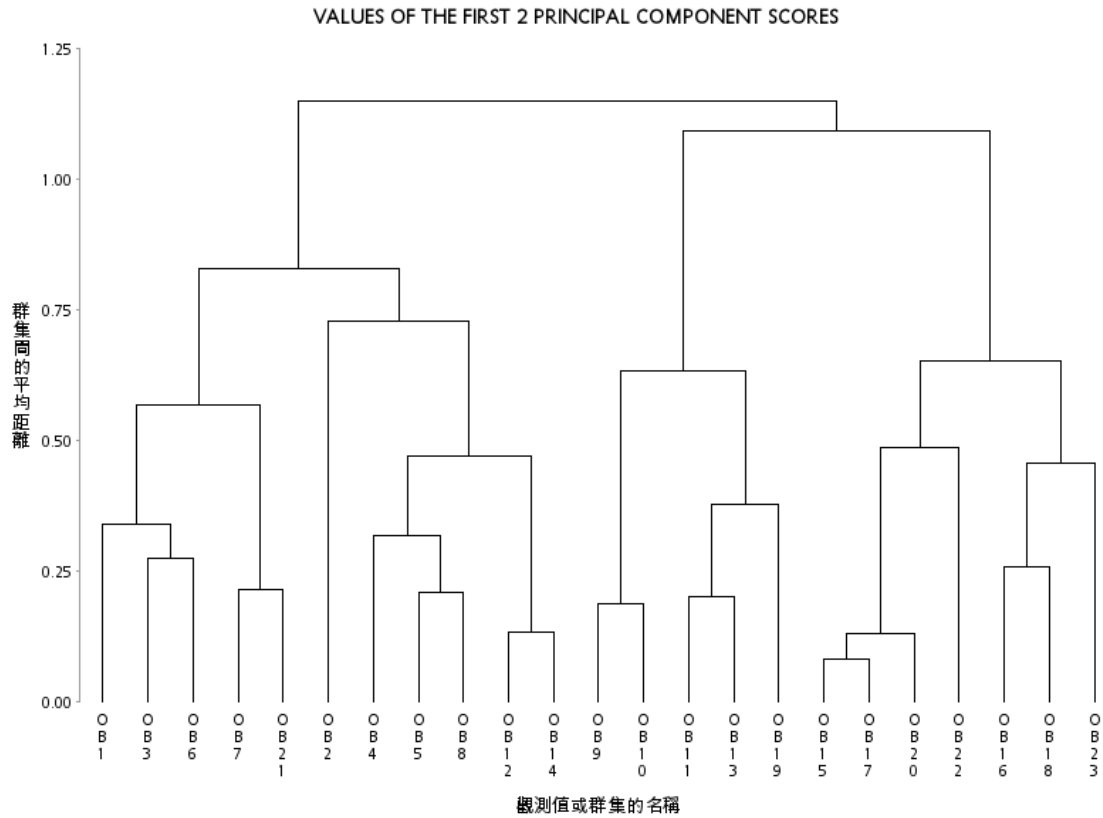


圖 7

接著分別對 Climate indices Component、Various measures Component 以及整體 cluster 做 MANOVA 分析，表 6、表 7、表 8 為 MANOVA 分析的報表，根據報表做虛無假設：

$$\text{Climate indices Component} \begin{cases} H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 \\ H_a: \text{at least one } \mu_i \neq \mu_j, i = 1, 2, 3, 4, j = 1, 2, 3, 4, i \neq j \end{cases}$$

$$\text{Various measures Component} \begin{cases} H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 \\ H_a: \text{at least one } \mu_i \neq \mu_j, i = 1, 2, 3, 4, j = 1, 2, 3, 4, i \neq j \end{cases}$$

$$\text{整體 cluster} \begin{cases} H_0: \text{無整體cluster效果} \\ H_a: \text{有整體cluster效果} \end{cases}$$

表 6

來源	自由度	平方和	均方	F 值	Pr > F
模型	3	83.3393896	27.7797965	25.21	<.0001
誤差	19	20.9344901	1.1018153		
已校正的總計	22	104.2738797			

表 7

來源	自由度	平方和	均方	F 值	Pr > F
模型	3	49.79034803	16.59678268	19.31	<.0001
誤差	19	16.33452923	0.85971206		
已校正的總計	22	66.12487726			

表 8

無整體 CLUSTER 效果假設的 MANOVA 檢定準則和 F 近似 H = 類型 III SSCP 矩陣 - 針對 CLUSTER E = 誤差 SSCP 矩陣 S=2 M=0 N=8					
統計值	值	F 值	分子自由度	分母自由度	Pr > F
Wilks' Lambda	0.04958890	20.94	6	36	<.0001
Pillai's Trace	1.55221006	21.95	6	38	<.0001
Hotelling-Lawley Trace	7.03004378	20.59	6	22.308	<.0001
Roy's Greatest Root	3.98365788	25.23	3	19	<.0001
附註: Roy Greatest Root 的 F 統計值是上限。					
附註: Wilks Lambda 的 F 統計值是精確值。					

表 6 為 Climate indices Component 的 MANOVA 檢定，表 7 為 Various measures Component 的 MANOVA 檢定，兩者的 p-value 均小於預設 α 值 0.05，故兩者均拒絕虛無假設，此資料在兩個主成分皆有明顯分群。表 8 為整體 cluster 的整體檢定，因為 p-value 小於預設 α 值 0.05，故拒絕虛無假設，也就是整體有 cluster 效果，因此可以得知用 PCA 作維度縮減後再分群的效果更勝未作維度縮減即分群的效果。圖 8 為兩個主成分的盒狀圖，橫軸為群集，左圖為 Climate indices Component，右圖為 Various measures Component，能夠藉由圖 8 明顯看出分群。

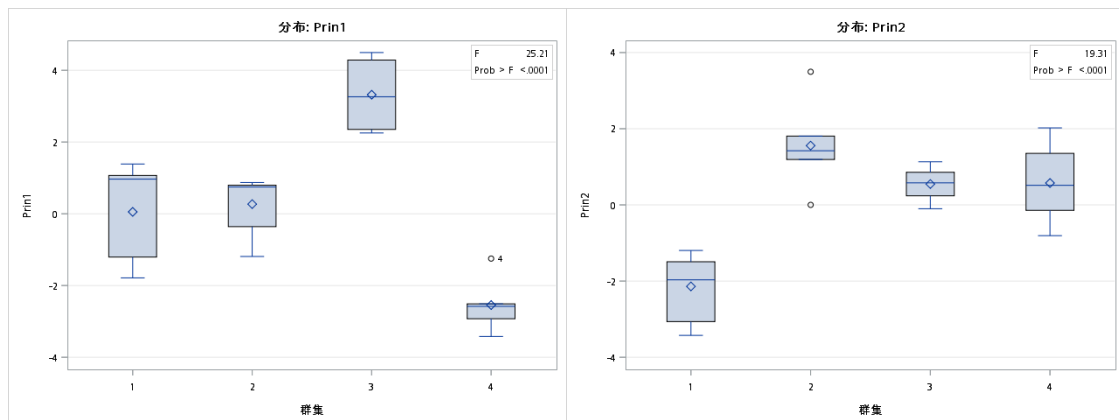


圖 8

(二) 因子分析法 (Factor analysis, FA)

將資料中的 17 個變數設為 x1-x17 並跑出分析結果，首先檢定是否需要因子：

$$\begin{cases} H_0: \text{沒有公因子} \\ H_a: \text{至少一個公因子} \end{cases}$$

表 9 為檢定結果。根據表 9 得知 p-value < 預設 α 值 0.05，故拒絕虛無假設，也就是至少有一個公因子。

本報告的因子分析分別採用主成分以及最大概度兩種方法來分析，藉由主成分方法所產生的因子，若希望累計解釋變異程度超過 90% 則需要用到第八個因子；而藉由最大概度方法產生的因子僅需要兩個因子就能使解釋變異程度超過 90%，由於維度縮減後將要作集群分析，因此採用維度較小的最大概度方法較優。

表 9

根據 23 觀測值的顯著性檢定			
檢定	自由度	卡方	Pr > 卡方
H0: 沒有公因子	136	322.1379	<.0001
HA: 至少一個公因子			

沒有 Bartlett 校正的卡方	457.22798
Akaike 訊息準則	185.22798
Schwarz 貝氏準則	30.80077
Tucker 和 Lewis 的可靠度係數	0.00000

表 10

	Factor1	Factor2
NPI_winter	0.82640	-0.05059
NPI	0.82549	-0.04910
AO_summ	0.80895	-0.01371
AO_wint	0.80086	-0.13193
AO	0.77226	0.13868
Rain_wint	0.52322	-0.00764
Temp_summ	0.48874	-0.05431
NPI_spring	0.45182	0.19269
Rain	0.38482	-0.18625
IceCover	0.25337	-0.19850
Temp_wint	-0.51480	0.29272
Ice	-0.09454	0.93321
Ice_JanJul	-0.17196	0.80793
Ice_OctDec	0.34153	0.76705
Rain_summ	-0.01645	0.03120
Temp	-0.17466	-0.41349
IceFreeDays	0.18512	-0.74092

表 10 為最大概度法以 Varimax 方法產生的旋轉因子模型，為了之後的集群分析，故僅採用兩個因子(兩個維度)。根據表 10 能將因子 1 命名為 Climate indices Factor，因子 2 命名為 Various measures Factor，命名規則能參照表 1 的變數介紹得知。

接下來將投射至 Climate indices Factor 及 Various measures Factor 的觀測值作集群分析。圖 9 為集群分析的階層式分群樹狀圖，分群方法為平均法，橫軸為觀測值，縱軸為群集間的平均距離，而我根據樹狀圖決定將資料分成四群。而表 11 為四個群集中 Climate indices Factor 及 Various measures Factor 的平均值。

表 11

	Climate indices Component	Various measures Component
Cluster1	-0.22	-1.21
Cluster2	0.154	0.921
Cluster3	1.317	-0.01
Cluster4	-1.19	0.605

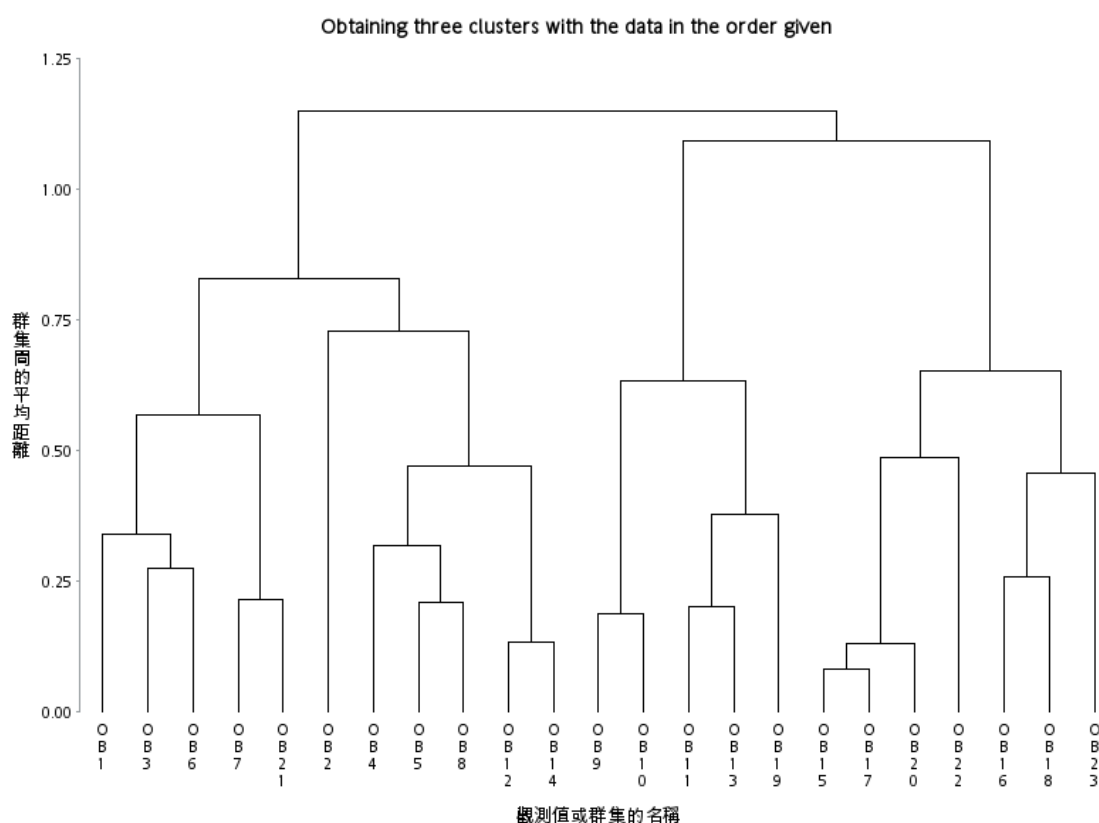


圖 9

接著分別對 Climate indices Factor、Various measures Factor 以及整體 cluster 做 MANOVA 分析，表 12、表 13、表 14 為 MANOVA 分析的報表，根據報表做虛無假設：

$$\text{Climate indices Component} \begin{cases} H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 \\ H_a: \text{at least one } \mu_i \neq \mu_j, i = 1, 2, 3, 4, j = 1, 2, 3, 4, i \neq j \end{cases}$$

$$\text{Various measures Component} \begin{cases} H_0: \mu_1 = \mu_2 = \mu_3 = \mu_4 \\ H_a: \text{at least one } \mu_i \neq \mu_j, i = 1, 2, 3, 4, j = 1, 2, 3, 4, i \neq j \end{cases}$$

$$\text{整體 cluster} \begin{cases} H_0: \text{無整體cluster效果} \\ H_a: \text{有整體cluster效果} \end{cases}$$

表 12 為 Climate indices Factor 的 MANOVA 檢定，表 13 為 Various measures Factor 的 MANOVA 檢定，兩者的 p-value 均小於預設 α 值 0.05，故兩者均拒絕虛無假設，此資料在兩個主成分皆有明顯分群。表 14 為整體 cluster 的整體檢定，因為 p-value 小於預設 α 值 0.05，故拒絕虛無假設，也就是整體有 cluster 效果，因此可以得知用 PCA 作維度縮減後再分群的效果更勝未作維度縮減即分群的效果。

表 12

來源	自由度	平方和	均方	F 值	Pr > F
模型	3	16.42024210	5.47341403	18.64	<.0001
誤差	19	5.57975790	0.29367147		
已校正的總計	22	22.00000000			

表 13

來源	自由度	平方和	均方	F 值	Pr > F
模型	3	17.61396408	5.87132136	25.43	<.0001
誤差	19	4.38603592	0.23084400		
已校正的總計	22	22.00000000			

表 14

無整體 CLUSTER 效果假設的 MANOVA 檢定準則和 F 近似 H = 類型 III SSCP 矩陣 - 針對 CLUSTER E = 誤差 SSCP 矩陣 S=2 M=0 N=8					
統計值	值	F 值	分子自由度	分母自由度	Pr > F
Wilks' Lambda	0.05038561	20.73	6	36	<.0001
Pillai's Trace	1.54700937	21.63	6	38	<.0001
Hotelling-Lawley Trace	6.99047593	20.48	6	22.308	<.0001
Roy's Greatest Root	4.09542838	25.94	3	19	<.0001
附註: Roy Greatest Root 的 F 統計值是上限。					
附註: Wilks Lambda 的 F 統計值是精確值。					

圖 10 為兩個主成分的盒狀圖，橫軸為群集，左圖為 Climate indices Factor，右圖為 Various measures Factor，能夠藉由圖 10 看出明顯的分群

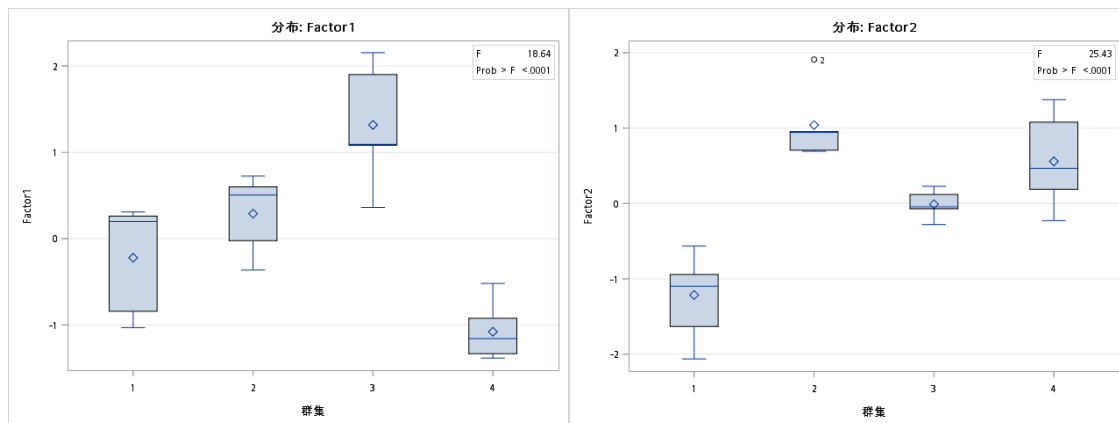


圖 10

伍、 結論

本報告中維度縮減後的分群效果相較於維度縮減前變好許多，無論是PCA或是FA的表現都很好。因為原本有17個維度，而觀測值僅有23個，故無法做好集群分析，而使用維度縮減方法將維度縮減為二維後則能清楚辨別不同的群集，然而缺點是在兩個維度的解釋變異程度並沒有想像中來的好。若之後要改進應該要採用 $n \gg p$ 的資料來作多變量分析會更好，或是本報告的氣候資料累積更多觀測值後再試試看作分析，因為這個資料的 n 與 p 實在太過接近，因此在一些多變量分析上的表現並不會這麼良好。

陸、 致謝

感謝許玉雪教授半學期來的傾囊相授，教會了我關於多變量分析的概念以及應用方法，我才能夠以課堂上習得的知識完成本次報告。

參考文獻

- Subhash Sharma. “Applied Multivariate techniques” .
 Johnson Wichern. “Applied Multivariate statistical analysis” .
 Michael Greenacre, Raul Primicerio. 2013. “Multivariate analysis of ecological data” .

全文完