

國立政治大學統計學系
多變量分析期中作業

交通流量之 PCA 與 FA

統計三 109304008 賴冠維

組員：歐西四 107508006 陳葳芃

此份資料為交通部公路總局所提供的 111 年度公路交通量調查統計表，收集了新北市、桃園市、臺中市等地 5 月 12 日到 9 月 18 日之交通資料。

接下來會對以下變數進行分析：

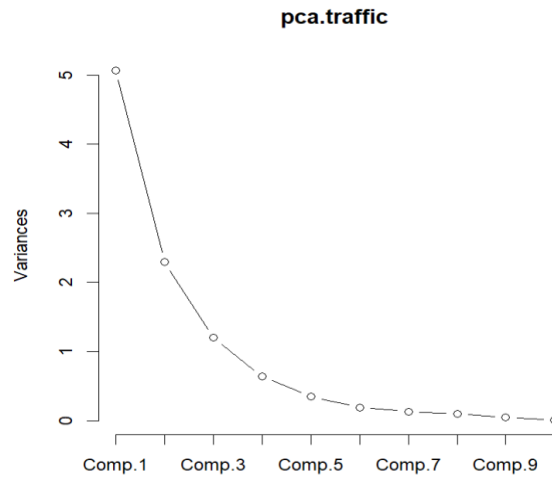
1. 路面寬度(公尺)
2. 各車種車輛數(輛/日)_小型車
3. 各車種車輛數(輛/日)_大客車
4. 各車種車輛數(輛/日)_大貨車
5. 各車種車輛數(輛/日)_全聯結車
6. 各車種車輛數(輛/日)_半聯結車
7. 各車種車輛數(輛/日)_機車
8. 總計_流量(PCU)
9. 尖峰小時_交通量(PCU)
10. 車道數量

最後一點為新增變數以取代原資料中之快車道及機慢車道寬度。因為每個路段之快車道及慢車數量不一，且車道寬度的變異性並不大，故新增變數取代之以便分析。

一、主成分分析

	Comp.1	Comp.2	Comp.3	Comp.4
標準差	2.2512386	1.5135926	1.0948776	0.79772307
解釋變異	0.5068075	0.2290963	0.1198757	0.06363621
累積解釋	0.5068075	0.7359038	0.8557795	0.91941569

表<一> PCA 解釋變異表



圖<一> Scree plot for PCA

由表<一>、圖<一>可以看出當取三個線性組合時，可以解釋約 86%的變數，並且其解釋變異 > 0.1 ，因此選取至 Comp.3。

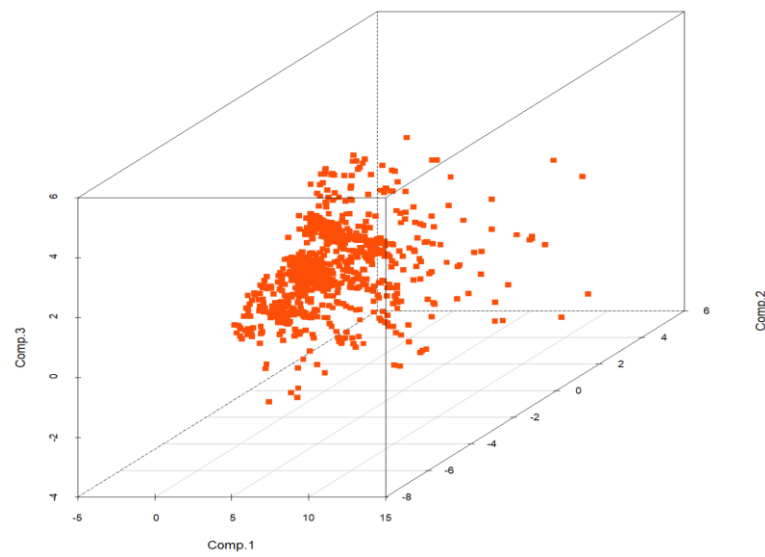
	Comp.1	Comp.2	Comp.3
路面寬度	0.284	0.132	0.635
小型車數	0.400		-0.223
大客車數	0.303	0.276	-0.254
大貨車數	0.352	-0.323	
全聯結車數	0.166	-0.528	0.124
半聯結車數	0.169	-0.538	0.116
機車數	0.231	0.417	-0.141
總流量	0.429		-0.200
尖峰流量	0.421		-0.153
車道數量	0.271	0.238	0.601

表<二> PCA 配適值

從表<二>中我們可以看到在 1st PC 中所有變數都有正的載荷，並且車輛數及流量有較大的載荷，這可能代表 1st PC 與整體道路的交通流量有較大的關係。

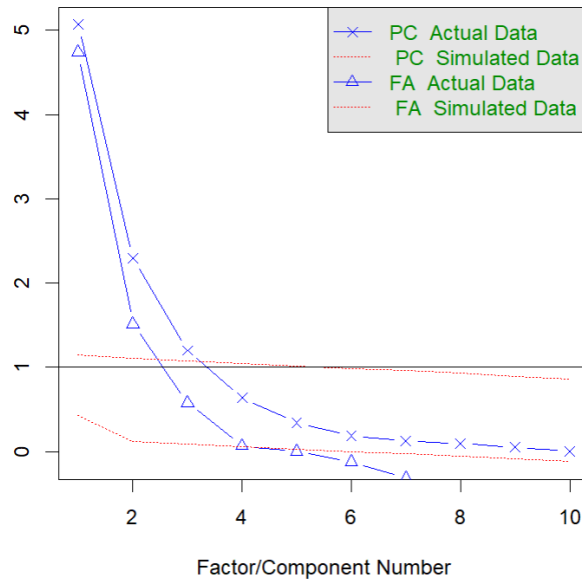
2nd PC 中大客車數和機車數有較大的正載荷，其餘與車輛有關的大多有負載荷，這可能顯示了 2nd PC 與平均每人之車道使用量或是車輛行駛模式有關。例如同樣的體積下機車和公車相較於其他車種可以運輸較多的人，又或者在台灣，機車和公車因為制度的關係更偏向於行駛右側的車道。

3rd PC 中路面寬度和車道數量有非常大的正載荷，這表示 3rd PC 與道路的寬敞度有關。



圖<二> PCA 三維圖

二、因素分析



圖<三> Parallel Analysis Scree Plots

根據圖<三>我們可以得知程式建議我們取至 Factor3 和 PC3

	Factor1	Factor2	Factor3
SS Loadings	3.292	2.154	2.039
Proportion Var	0.329	0.215	0.204
Cumulative Var	0.329	0.545	0.748

表<三> 變異解釋

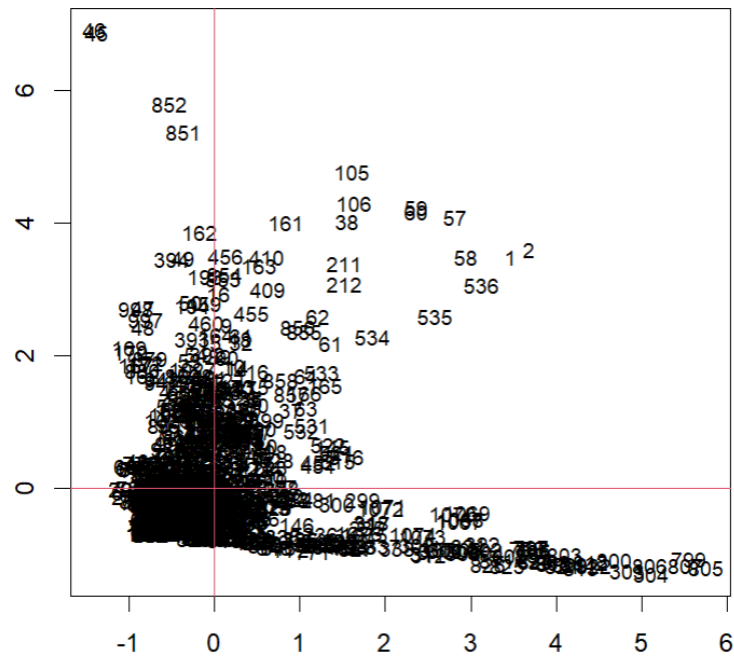
	Factor1	Factor2	Factor3
路面寬度	0.340	0.308	0.123
小型車數	0.958	0.250	0.126
大客車數	0.473	0.587	
大貨車數	0.770		0.508
全聯結車數	0.171		0.798
半聯結車數	0.138		0.987
機車數	0.109	0.975	-0.179
總流量	0.844	0.481	0.231
尖峰流量	0.756	0.554	0.214
車道數量	0.312	0.394	

表<四> 因素分析

根據表<三>及表<四>，我們可以將所有的變數組成三個新的因素：

1. 尖峰時段交通順暢度：由路面寬度、小型車數、大貨車數、總流量及尖峰流量組成
2. 道路使用效率：由大客車數、機車數及車道數量組成
3. 大型車數量：全聯結車數及半聯結車數組成

第二點與 2nd PC 的分析結果相似因此取其名，而第一點則是因為小客車及大貨車的道路使用效率較低，加上尖峰流量及總流量都是影響尖峰時段交通順暢度的關鍵因素。



圖<四>中橫坐標為 Factor1，縱座標為 Factor2。在圖中越靠右下角的點代表其尖峰時間的車流量較高且道路使用效率較低，交通狀況偏向壅塞。左上角則代表尖峰時間車流量較低且道路使用效率較高，交通狀況相對於右下角的點較順暢。

附錄

附錄一 程式碼

```
traffic_data <- read.csv("C:/Users/david/OneDrive/桌面/多變量/111 年 0512 至 09
18 交通資料.csv")
library(dplyr)
library(ggplot2)
library(stringr)
library(stats)
str(traffic_data)

##將 27:36 欄轉成數值型資料
num = c(27:36)
for(i in 1:length(num)){
  traffic_data[,num[i]] = as.numeric(traffic_data[,num[i]])
}

##計算各車道數量
fast_lane = NULL
slow_lane = NULL
for(i in 1:nrow(traffic_data)){
  fast_lane = c(fast_lane, 5 - sum(traffic_data[i,19:23]==0))
  slow_lane = c(slow_lane, 2 - sum(traffic_data[i,24:25]==0))
}

##新增車道總數量
traffic_data = mutate(traffic_data, 車道數量 = fast_lane+slow_lane)
traffic = traffic_data |>
  select(路面寬度.公尺., 各車種車輛數.輛.日._小型車:各車種車輛數.輛.
日._機車, 總計_流量.PCU., 尖峰小時_交通量.PCU., 車道數量)
##PCA
pca.traffic = princomp(traffic, cor=T,)
summary(pca.traffic)

##Scree plot
screeplot(pca.traffic, type = "l")
```

```
loadings(pca.traffic)
```

```
##3D 圖
```

```
pcs.traffic = predict(pca.traffic)
library(scatterplot3d)
s3d <- scatterplot3d(pcs.traffic[,1:3],
                     pch = 15,
                     color = c('#FC4E07'),
                     lty.hide = 2
)
```

```
##FA
```

```
library(psych)
traffic.cor = cor(traffic)
traffic.per = fa.parallel(traffic.cor, n.obs = 1076, fa="both")
traffic.fa = factanal(traffic, factors = 3, rotation = 'varimax', socrs = "regression")
traffic.fa
```

```
##FA 二維圖
```

```
fa = factanal(traffic,factors=3,scores="regression")
plot(fa$scores[,1],fa$scores[,2],type="n", ylab='Factor 2',xlab= 'Factor 1')
text(fa$scores[,1:2],row.names(traffic),cex=0.8)
abline(h=0, v=0,col=2)
```

資料來源：

https://www.thb.gov.tw/News_Download.aspx?n=273&sms=12823