

學號 0411509 姓名:許家維

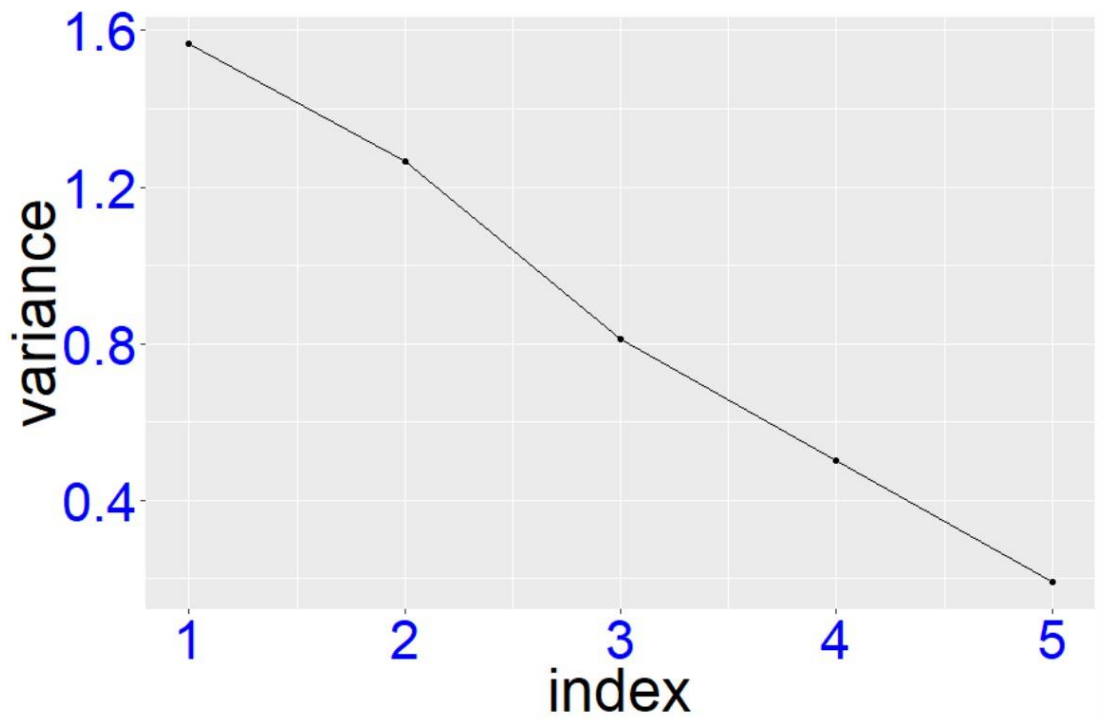
此次作業我沿用上次作業的 data 他的參數主要是透過 Wavelet Transformed 得到鈔票圖像的

- A. Variance
- B. Skewness
- C. Kurtosis
- D. 圖像亂度(entropy)

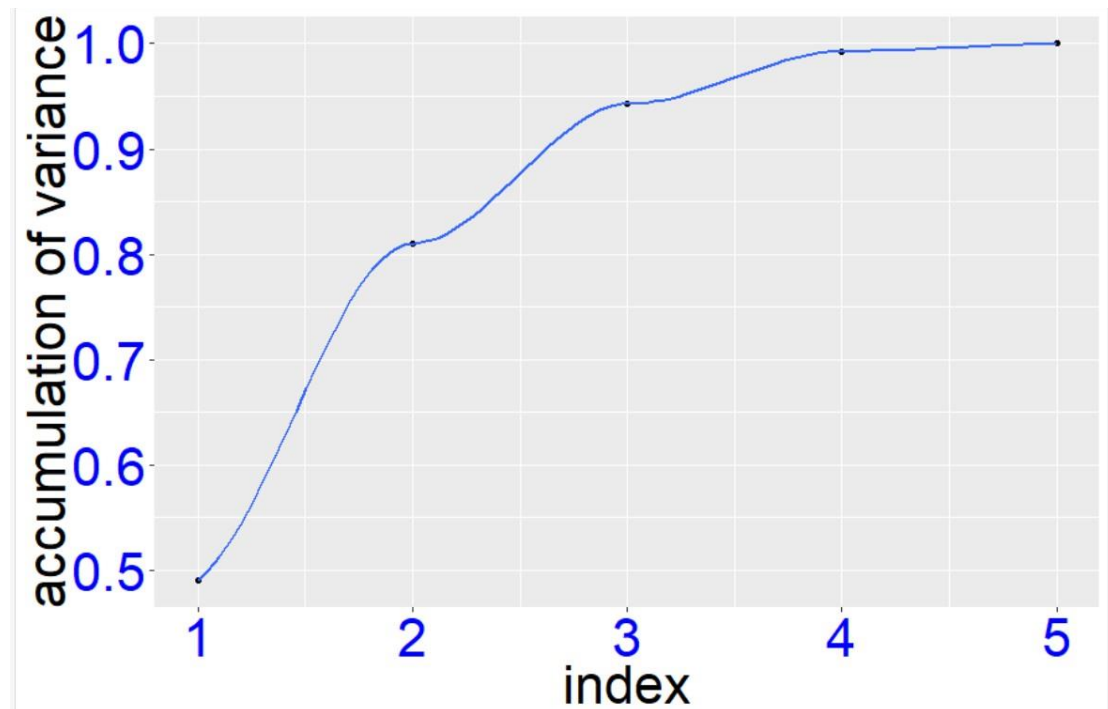
目標設定為能夠辨識出真假鈔票

我使用 R 語言並實作 PCA(Principal Component Analysis)分析, 先將複雜數據降維以利觀察趨勢

1. 首先我觀察 INDEX 的主成分變異數, 發現前 2 個數比較大(有超過 1)



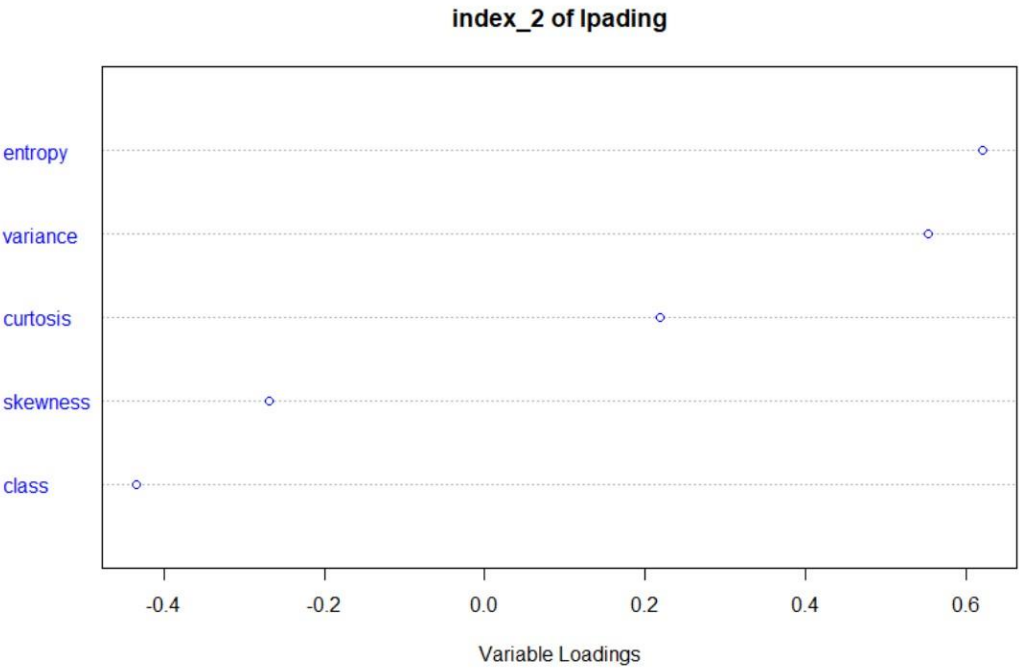
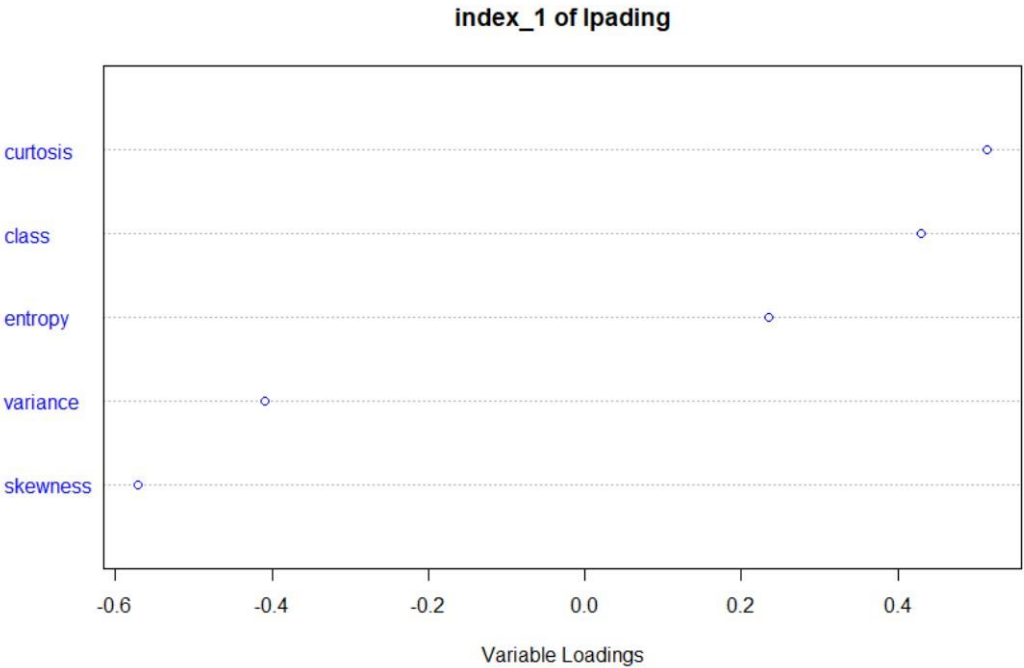
然後對 INDEX 做每個主成分特徵/總特質的累進圖,發現前 2 個參數已囊括約 8 成的變異,接下來將會使用 2 維平面圖做數據呈現



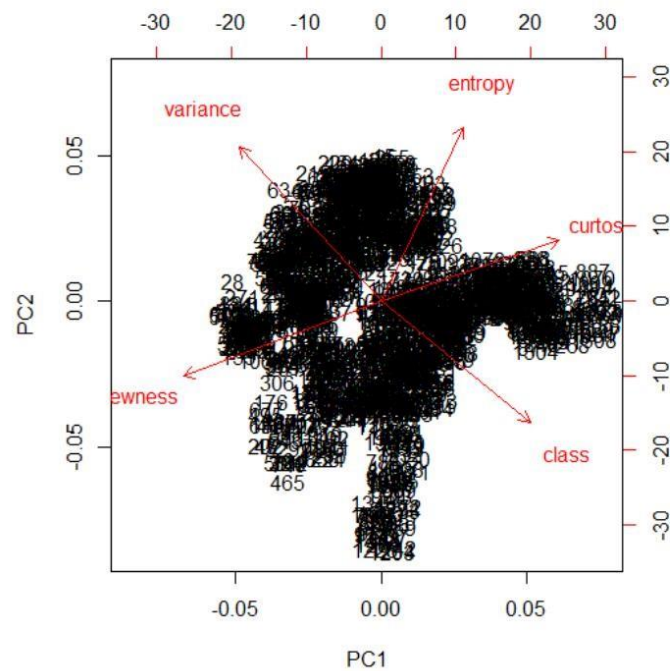
並且我叫出主成分的特徵值,觀察數據間互相的影響,發現大部分數據間都有著高度正相關和負相關

	PC1	PC2	PC3	PC4	PC5
variance	-0.4089107	0.5523651	0.12767053	-0.58042680	0.41771977
skewness	-0.5712405	-0.2683287	-0.01317329	0.54283647	0.55393058
curtosis	0.5151075	0.2191741	-0.62782137	0.09121597	0.53305437
entropy	0.2359805	0.6215523	0.50572528	0.54945741	0.01801391
class	0.4306435	-0.4341981	0.57758893	-0.24126644	0.48394178

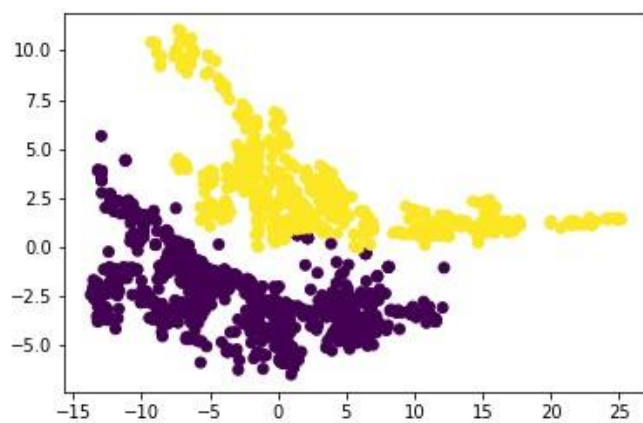
藉由此數據繪製主成分負荷圖，並列出較大關係的第一第二主成分



最後結合上面 2 張圖繪製平面負荷圖並結合 CLASS(辨識鈔票的分類)



可以看到藉由 pc1 與 pc2 作圖,數據被大致分為了 2 類



結合上週使用 ISOMAP 進行降維的結果,也是很明顯的分為 2 類,圖中不同顏色的資料點代表鈔票的真假,不過這周發現 variance 跟鈔票真偽有著相反的趨勢