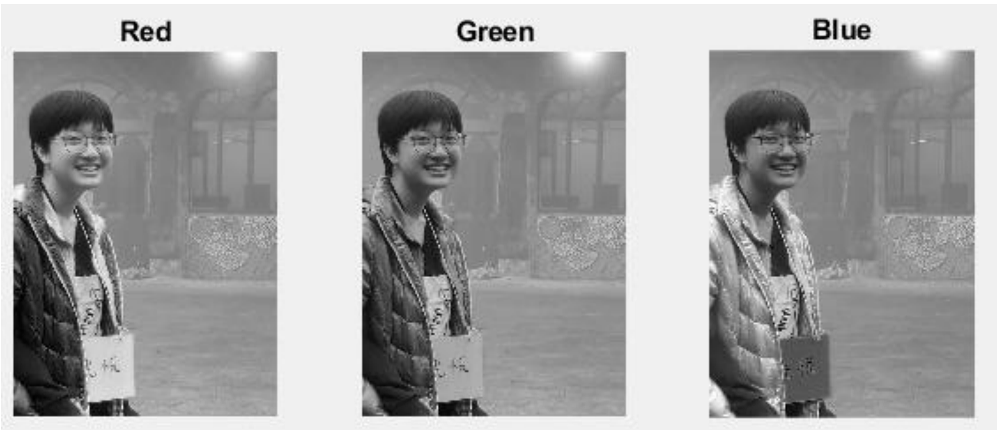
線性代數 期末報告

陳光悅

10811020

Part 1

Part 2



Part 3

Part 4

Part 5

Part 6

Part 4 的特徵分解和 Part 5 的 SVD 分解，兩者的結果不盡相同，尤其是在保留前 10 大特徵值的圖片。造成此差異的原因是因為特徵分解和 SVD 分解對於矩陣的分解方式不同。

特徵分解時，矩陣 $A_{n\times n}$ 被分解成 PDP^{-1}。其中，D 是 diagonal matrix，對角線上的 entries 是 eigenvalue；P 是對應的 eigenvector 的矩陣。若將這個看作是 transformation matrix，則 eigenvalue 對應到的 eigenvector 就是在描述這個矩陣變化方向。eigenvector 組成 transformation matrix 的 basis，每一個 eigenvector 皆為 linear independent。因此將矩陣 A 做特徵分解，便是將 A 投影到 n 個 eigenvector（orthogonal basis）上。eigenvalue 越大，表示他在對應的 eigenvector 上的信息量越多，若將它保留下來，可以還原更多的圖片信息。

而 SVD 分解時，矩陣 $A_{m\times n}$ 被分解成 $U\Sigma V^T$。其中，U 和 V 是 orthogonal matrix；Σ 是 diagonal matrix，對角線上的 entries 是 singular value。進行 SVD 分解時，其實就是進行 A^TA 的特徵分解。其中，A^TA 的 eigenvalue 對應到的 eigenvector 形成了矩陣 V；AA^T 的 eigenvalue 對應到的 eigenvector 形成了矩陣 U。而因為部分的特徵或信息可以通過其他特徵或信息推導得到，或是矩陣中的變異性主要集中在少數幾個主要的 singular value 上。因此在多數的情況下，前 10%甚至 1%的 singular value 的和就佔了全部的 singular value 之和的 99%以上。而此時僅需以這少數幾%的 singular value 即可還原原始的圖片信息。

故若觀察 Part 4 和 Part 5 的結果，便可以發現 SVD 分解時，保留前 100 個

和保留 200 個最大特徵值的結果差異不大；然特徵值分解時，保留前 100 個和保留 200 個最大特徵值的結果仍頗多。

因此，特徵分解在解釋能力上更強，因為他能夠提供 eigenvalue 和 eigenvector，可以用來描述矩陣如何在新的 basis 上做 transformation。在機器學習的特徵提取中，經常使用特徵分解來實現，如：主成分分析（Principal Component Analysis；PCA）。

而 SVD 分解則更適用於矩陣的數據壓縮。他能夠將矩陣中較為重要的部分拆解出來，在數據、圖片、影片的壓縮應用中，被廣泛使用。

參考資料：

https://geek.digiasset.org/pages/ml/basic/eigenvalue-singular-value-decomposition-pca/