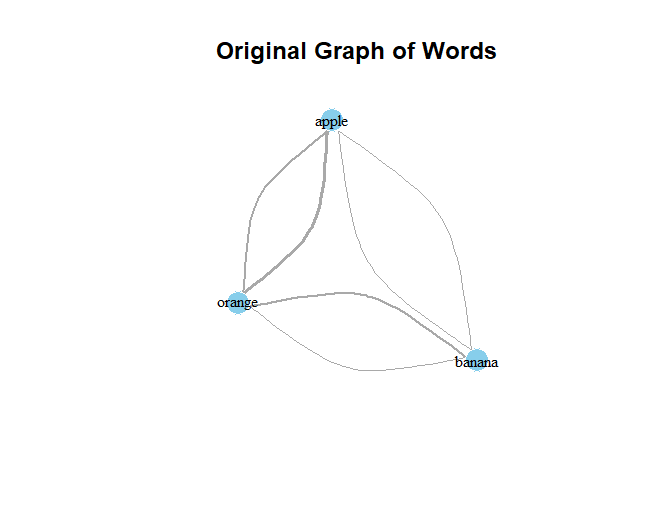
Laplacian eigenmap 是一種非線性降維技術，通常用於機器學習和數據可視化，特別是在處理具有複雜內在結構的數據時。它基於圖論中的拉普拉斯矩陣（Laplacian matrix），用於發掘數據點之間的低維表示，保持數據的局部特徵。

在文本分析或詞彙群聚的情境中，Laplacian eigenmap 可以用於探索詞彙之間的關係，尤其是在高維空間中的關係。這種方法可以幫助揭示詞彙的結構和模式，進而用於文本數據的降維或聚類分析。與前面提到的詞群分析使用詞雲的方法不同，Laplacian eigenmap 會進行更深層次的數據關係挖掘。

這裡是如何將 Laplacian eigenmap 應用到詞彙數據上的一個基本概念：

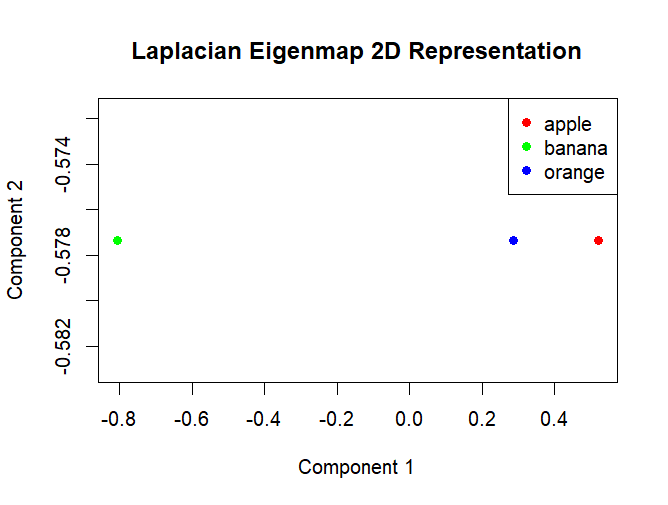
1. **建構圖模型**：在處理詞彙數據時，每個詞彙可以視為圖中的一個節點。節點之間的邊可以根據詞彙之間的共現關係或其他相似性度量來確定。
2. **計算拉普拉斯矩陣**：一旦圖形模型建立，可以計算其拉普拉斯矩陣。拉普拉斯矩陣是基於節點度（degree）和邊的權重來形成的，它反映了節點之間的連接強度。
3. **求解特徵值和特徵向量**：Laplacian eigenmap 的核心在於計算拉普拉斯矩陣的特徵值和特徵向量。這些特徵向量會根據特徵值的大小來排序，其中最小的非零特徵值對應的特徵向量（除了最小的零特徵值對應的均勻特徵向量）提供了數據最重要的低維表示。
4. **降維表示**：選擇幾個最小的非零特徵值對應的特徵向量，用它們來表達原始高維數據點在低維空間中的位置。這有助於保留數據點之間的局部鄰近關係。

將 Laplacian eigenmap 應用於詞彙群聚或任何文本數據分析中，可以幫助研究者發現詞彙之間更深層次的關係，進而在這些關係的基礎上進行更有效的聚類或其他形式的分析。這是一種相對於傳統詞頻方法更先進且富有洞察力的分析方式。

****

**原始圖形分析**

原始圖形展示了三個節點（apple, banana, orange）和它們之間的連接。邊的粗細可能代表了這些詞彙之間共現關係的強度或頻率。例如，apple和orange之間的連接看起來比其他連接要粗，這可能意味著apple和orange在數據集中更頻繁地一起出現。此圖是高維數據的一個表達，其中節點的位置是由圖形演算法決定，用以清晰地表示節點之間的關係。



**降維後的圖形分析**

降維後的圖形是使用Laplacian eigenmap技術將原始數據投影到二維空間的結果。在這個散點圖中，每個顏色代表一個詞彙，點的位置代表降維後的坐標。這個圖形顯示，即使在降維之後，這三個詞彙仍然能夠在二維空間中被區分開來，並且保持了一定程度的關係結構。

**比較分析**

比較這兩個圖形，我們可以看到原始數據中的關係在降維後有所保留。在降維後的圖形中，節點之間的相對距離依然體現了它們的相似性，雖然缺乏邊的信息。例如，如果apple和orange在原始數據中有較強的關係，我們希望在降維後的圖形中也看到它們彼此接近。

**結論**

降維過程似乎有效地將高維數據的結構轉化成了一個更為簡潔的二維表示，同時在某種程度上保留了詞彙之間的關係