

日期: 2017/12/13(三) 上機考: Open Book  
授課教師: 吳漢銘 (臺北大學統計學系副教授)

## 請仔細閱讀每一個注意事項 (禁止討論)

### 1. 考試期間

- (a) 請按照平時上課之座位入座。
- (b) 可參考課本、上課講義 (包含電子檔) 及其它資料，但不能與別人討論。
- (c) 可使用計算機、自己的筆記型電腦及平板電腦，不可使用手機。
- (d) **全程可上網查詢，但不能用通訊軟體討論。**
- (e) 有問題者，請舉手發問。勿與同學交談。
- (f) 不按照規定作答者，酌量扣分。
- (g) 不可使用它人之隨身碟。「作弊」或「疑似作弊」，往後各項考試不予評分。

### 2. 下載題目卷，上傳答題檔案:

- (a) 於課程網站下載題目卷。
- (b) 於教師網站首頁登入 [作業考試上傳區]，帳號: hdda。密碼: xxx (上課教室號碼)。
- (c) 請上傳「學號-姓名-HDDA-exam3.docx」。(目錄: 「20171213-exam3」。)

### 3. 答題檔案原則:

- (a) **請依照「R 程式作業繳交方式」，複製 Console「程式執行及結果」至答案卷。** 圖形複製，請注意大小，內容數字文字需可辨識。
- (b) 程式設計題，若程式碼直接複製 (或照抄) 講義上的以不給分為原則。
- (c) 若上傳檔案格式錯誤，內容亂碼，空檔等等問題。請自行負責。
- (d) 若要重覆上傳 (第 2 次以上)，請在檔名最後加「-2」、「-3」，例如: 「學號-姓名-HDDA-exam3-2.docx」、「學號-姓名-HDDA-exam3-3.docx」等等。
- (e) 上傳兩次 (含) 以上、格式不合等等酌量扣分。

### 4. 完成考試

- (a) 上傳完畢，請通知教師確認。
- (b) 確認無誤，請刪除作答目錄及答案卷，清空資源回收筒，並關機。即可離席。

我已經仔細閱讀上述各注意事項，若有違背，會自行負責。

日期: 2017/12/13(三) 上機考: Open Book  
授課教師: 吳漢銘 (臺北大學統計學系副教授)

資料集來源: Wine Data Set, <https://archive.ics.uci.edu/ml/datasets/wine>

以下 ISOMAP 演算法中, 鄰居個數一律設定為 5, 若自覺得不合適, 請自行選一“合適”的個數。

1. 讀取資料, 以 MDS 及 ISOMAP 做維度縮減, 各得到前兩維的維度縮減資料, 畫出散佈圖, 其中圖上的點 (酒) 若為同一品種, 則以同顏色顯示。(酒彼此之間的距離為歐式空間距離)
2. 使用 ISOMAP 做維度縮減, 以 rgl 套件畫出前三維之 3D 散佈圖。(點的颜色顯示要求同上, 3D 散佈圖請轉三個不同角度貼上答案卷)
3. 於上小題之 3D 散佈圖中, 若是兩點鄰居, 則加一連線。(點的颜色顯示要求同上, 3D 散佈圖請轉三個不同角度貼上答案卷)
4. 將資料隨機分成 2/3 訓練集及 1/3 測試集, 利用訓練集造出線性 SVM 模型, 印出測試集的分類 cross table, 並算出測試集的分類錯誤率。(set.seed(12345))
5. 使用 ISOMAP 做維度縮減得到的投影資料記做  $\mathbf{Z}_{n \times k}$ , 其中  $n$  為酒的個數,  $k = 1, 2, \dots, 10$  為所取的低維度個數。當  $k = 1, 2, \dots, 10$  時, 將資料  $\mathbf{Z}_{n \times k}$  隨機分成 2/3 訓練集及 1/3 測試集, 利用訓練集造出線性 SVM 模型, 計算出測試集的分類錯誤率。(set.seed(12345))
6. 使用 MDS, 重覆上小題之步驟。
7. 將前 3 小題的三種方法所得到的錯誤率, 繪出線圖。橫軸為  $k = 1, 2, \dots, 10$ , 縱軸為錯誤率。(需加上方法的圖例說明 (legend) 。)
8. 算出 Wine Data Set 的 geodesic distance (即 IsoDistance), 記做  $D_G$ , 使用 hclust {stats} 在  $D_G$  上做群集分析, 並使用 cutree {stats} 將 hclust 結果分成 3 群, 將分群結果以顏色呈現在 ISOMAP 的前兩維散佈圖上。
9. 以 LCMC 及 *STRESS* 評估 MDS 及 ISOMAP 維度縮減的表現。(  $d_{ij}, \hat{d}_{ij}$  請參照講義)

$$STRESS = \sum_{i < j}^n (d_{ij} - \hat{d}_{ij})^2.$$

注意: 上傳檔案之後, 請刪除作答目錄 及 答案卷, 清空資源回收筒, 關機。交回題目卷。