

The upper bound of Bayes error

7109056202 資工碩一 陳彥儒

1. Introduction

這次的作業是藉由HW1給定的Covariance Matrix 和 Mean來算Bayes error。

Bayes錯誤率 (Bayes error rate) 是應用Bayes分類規則的分類器的錯誤率。Bayes分類規則的一個性質是:在最小化分類錯誤率上是最優的。所以在分類問題中，Bayes錯誤率是一個分類器對某個類別所能達到的最低的分類錯誤率。

定義:

$$p_e = \sum_{i=1}^M \int_{R_i} \left(\sum_{j=1}^M \lambda_{ji} p(x|\omega_j) P(\omega_j) \right) dx$$

2. Implementation

- a. 首先要先計算covariance matrix，covariance matrix

公式如下圖:

$$\cdot \Sigma_i = \begin{bmatrix} 1 & \rho_i & \dots & \rho_i^{n-1} \\ \rho_i & 1 & \vdots & \vdots \\ \vdots & \dots & \ddots & \rho_i \\ \rho_i^{n-1} & \dots & \rho_i & 1 \end{bmatrix}$$

程式碼如下圖:

```
15 def cal_CovarianceMatrix(var):
16     arrayA = np.zeros((col,row))
17     for i in range(col):
18         for j in range(row):
19             arrayA[i][j]=(float)(var**(abs(i-j)))
20
21     return arrayA
```

以variance = 0.9的結果為例子，成果如下:

```
[[1.         0.9         0.81        ... 0.00706965 0.00636269 0.00572642]
 [0.9         1.         0.9         ... 0.00785517 0.00706965 0.00636269]
 [0.81        0.9         1.         ... 0.00872796 0.00785517 0.00706965]
 ...
 [0.00706965 0.00785517 0.00872796 ... 1.         0.9         0.81        ]
 [0.00636269 0.00706965 0.00785517 ... 0.9         1.         0.9         ]
 [0.00572642 0.00636269 0.00706965 ... 0.81        0.9         1.         ]]
```

b. Calculate Bhawttacharyya Distance

- i. 為了能進行矩陣運算，首要任務要先把兩個
mean都變成50*1的矩陣。因為我作業1的MEAN
為0和2.5，所以這裡照舊。

```
for i in range(col):
    array1[i] = m1
for i in range(col):
    array2[i] = m2
```

- ii. 接下來要算mean-difference，公式為

$\frac{1}{8}(M_2 - M_1)^T \left[\frac{\Sigma_1 + \Sigma_2}{2} \right]^{-1} (M_2 - M_1)$ ，程式碼如下:

```
30 meanDiff = ((1/8) * (array2- array1).T.dot(np.linalg.inv(0.5*(covMatrix1+covMatrix2))))).dot( array2- array1)
31 print("mean difference :",meanDiff[0][0])
32
```

iii. 再算covariance-difference，公式為 $\frac{1}{2} \ln \frac{|\frac{\Sigma_1 + \Sigma_2}{2}|}{\sqrt{|\Sigma_1||\Sigma_2|}}$ ，

程式碼如下：

```
temp = np.linalg.det(covMatrix1)*np.linalg.det(covMatrix2)
covDiff = 0.5 * np.log(np.linalg.det((covMatrix1+covMatrix2)*0.5)/(temp**0.5))
print("covariance difference: ",covDiff)
```

iv. 最後再把mean-difference+cov-difference相加

結果圖如下：

```
mean difference : 4.009096769535855
covariance difference: 3.5069801497545754
BhattacharyyaDistance: 7.51607691929043
```

c. Calculate Bhattacharyya error bound

題目給定 $P_1=P_2=0.5$ ，公式為 $\sqrt{P_1 P_2} e^{-\mu(\frac{1}{2})}$ ，程式碼如下：

```
bayesError = ((p1*p2)**0.5)*math.exp(-ans1)
```

結果圖如下：

```
Bayes error : 0.000272131786490041
```

3. Discussion

透過這次的作業讓我們知道Bayes error如何得出的，感謝老師及助教。

