PMDS602P - Advanced Machine Learning Lab

Slot: L7 + L8
Fall Semester 2025 - 26
VIT - Vellore

Date: August 5, 2025

Answer all the questions.

1. Computing Mahalanobis distance:
```
import numpy as np
from scipy.spatial.distance import mahalanobis
A = [[3, 0, 0], [0, 5, 0], [0, 0, 7]]
B = np.linalg.inv(A)
print(B)
a = mahalanobis([1, 1, 0], [0, 1, 0], B)
b = mahalanobis([1, 1, 0], [0, 0, 0], B)
print(a)
print(b)
```

2. Finding the mean of a set of vectors:
```
import numpy as np
A = np.array([[1, 0, 0], [0, 2, 3], [0, 1, 1]])
mu = np.mean(A, axis = 0)
print(mu)
```

3. Finding the outlier of the set of data $\{[1,3],[3,1],[2,2],[2,4],[4,3],[200,100]\}$.

   HINT: Note that we require defining a notion of *distance*, and have to reject the point that has the maximum distance, that crosses a given *threshold value*. That is, we consider the *Mahalanobis distance* and the *chi-squared test* to find the outlier.

```
import numpy as np
from scipy.stats import chi2
teststat = np.sqrt(chi2.ppf(0.85, 2))
```