

Homework 1

CSC 546

Syed Muhammad Hamza Shah

PART 1

HW part 1

$$1- \quad x = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad , \quad y = \begin{bmatrix} 10 \\ 18 \end{bmatrix}$$

L2

$$\begin{aligned} L2\text{-distance} &= \sqrt{(1-10)^2 + (2-18)^2} \\ &= 19.417 \text{ units.} \end{aligned}$$

L1

$$\begin{aligned} L1\text{-distance} &= |1-10| + |2-18| \\ &= 9 + 16 = 25 \text{ units} \end{aligned}$$

L∞

$$\begin{aligned} L\infty\text{-distance} &= \max \{ |1-10|, |2-18| \} \\ &= \max \{ 9, 16 \} = 16 \text{ units} \end{aligned}$$

I think L^∞ distance would not be good in determining the distance between the vectors with components [income, spend] because both the factors: income and spending are important to consider if we want to attain a better conclusion about the spending habits of a consumer. So I think L_2 norm would be better

$$2. \quad f(x) = x^T A x$$

$$f(x) = \begin{bmatrix} \alpha & \beta \end{bmatrix} \begin{bmatrix} a & c \\ c & b \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

$$= \begin{bmatrix} \alpha a + \beta c & \alpha c + \beta b \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

$$= \alpha^2 a + \beta c \alpha + \alpha c \beta + \beta^2 b$$

$$\frac{df}{d\alpha} = [2a\alpha + \cancel{\beta c} + c\beta + 0]$$

$$\frac{df}{d\beta} = [0 + c\alpha + \alpha c + 2b\beta]$$

$$2Ax = \begin{bmatrix} 2a & 2c \\ 2c & 2b \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}$$

$$= \begin{bmatrix} 2a\alpha + 2c\beta \\ 2c\alpha + 2b\beta \end{bmatrix} \cdot \{2c\alpha + 2b\beta\}$$

$$\frac{df}{dx} = \begin{bmatrix} 2a\alpha + 2c\beta \\ 2c\alpha + 2b\beta \end{bmatrix}$$

PART 2

K-Means Clustering

3. Briefly describe the two key steps in one iteration of the k-means algorithm.

There are 2 key steps which are center allocation and then updating the cluster points assigned to each center by assigning them to the closest center.

4. What is the distance measure used in k-means (implemented in sklearn)?

Euclidean center.

5. The k-means algorithm can converge in a finite number of iterations. Why?

After many iterations, the least avg distance between the centroids and the clusters reaches its minimum and after that the centers can be moved to a better place to minimize that distance so it converges.

6. The clustering result of k-means could be random. Why?

It could be random mainly to the fact that the clusters and the center points are chosen at different locations so they cluster at different locations and groups after many iterations.

7. The minimum value of the objective/loss function is zero for any dataset. What is the clustering result when the objective function is zero?

When the loss function is zero, that means each center is assigned to its own cluster and the kmeans cant converge further.