Problem 1

1. Code

2. K_mean is my main function that combines all the other sub functions together. It takes in centroid and datas. 'K' isn't included in the parameter. 'K' is used in initialization stage. However, considering that I'll have the K_means algorithm in a while loop, I didn't include initialization in the logic. Inside k_means, using the datas and centroid given, it assigns points in data to closest centroid then it also updates centroids after.

3. code

4. I would pick 4. Up until $K=5$, there is rapid exponential decrease but after, it slows down. Considering other factors, for example cost in t-shirt example given in lecture, $K=4$ seemed to be the best option to maximize in every perspective.

Problem 2

$x^{(1)} = 1$, $x^{(2)} = 2$, $x^{(3)} = 3$, $x^{(4)} = 4$

Init_centroid : $\mu_1 = 2$ $\mu_2 = 4$

- $dist(x^{(1)}, \mu_1) = \sqrt{(1-2)^2} = 1 \Rightarrow \mu_1$

  $dist(x^{(1)}, \mu_2) = \sqrt{(1-4)^2} = 3$

- $dist(x^{(2)}, \mu_1) = \sqrt{(2-2)^2} = 0 \Rightarrow \mu_1$

  $dist(x^{(2)}, \mu_2) = \sqrt{(2-4)^2} = 2$

- $dist(x^{(3)}, \mu_1) = \sqrt{(3-2)^2} = 1 \Rightarrow \mu_1$

  $dist(x^{(3)}, \mu_2) = \sqrt{(3-4)^2} = 1$

- $dist(x^{(4)}, \mu_1) = \sqrt{(4-2)^2} = 2$

  $dist(x^{(4)}, \mu_2) = \sqrt{(4-4)^2} = 0 \Rightarrow \mu_2$

$\Rightarrow \mu_1' = mean\{1, 2, 3\}$ $\mu_2' = mean\{4\}$

$= 2$ $= 4$

$$Cost(X) = \frac{1}{m}\sum_{i=1}^{m} \| X^{(i)} - \mu_{x^{(i)}} \|^2$$

$$= \frac{1}{4}\left[(1-2)^2 + (2-2)^2 + (3-2)^2 + (4-4)^2\right]$$

$$= \frac{1}{2}$$

Consider different

If Init_centroid : $\mu_1 = 2.5 \quad \mu_2 = 4$

- $dist(x^{(1)}, \mu_1) = \sqrt{(1-2.5)^2}$ $\Rightarrow \mu_1$
  $dist(x^{(1)}, \mu_2) = \sqrt{(1-4)^2}$
- $dist(x^{(2)}, \mu_1) = \sqrt{(2-2.5)^2}$ $\Rightarrow \mu_1$
  $dist(x^{(2)}, \mu_2) = \sqrt{(2-4)^2}$
- $dist(x^{(3)}, \mu_1) = \sqrt{(3-2.5)^2}$ $\Rightarrow \mu_1$
  $dist(x^{(3)}, \mu_2) = \sqrt{(3-4)^2}$
- $x^{(4)} \Rightarrow \mu_2$

$\Rightarrow \mu_1' = \text{mean}\{1,2,3\} = 2 \qquad \mu_2' = \text{mean}\{4\} = 4$

$$cost(x) = \frac{1}{4}\left[(1-2.5)^2 + (2-2.5)^2 + (3-2.5)^2 + (4-4)^2\right]$$
$$= \frac{1}{4}\left(\frac{11}{4}\right)$$
$$= \frac{11}{16}$$

They both converge to $\mu_1, \mu_2 = [2, 4]$

However, as shown through calculation, they have different cost with one being 1/2 while the other being 11/16. This shows how the algorithm may converge to one solution and still there are multiple cost within. In other words, the converged solution may not be the global optimal.

Problem 3

① $A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix}$

Rank : # of independent columns

Find Rank by row reduction.

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix} \xrightarrow[R_1, R_2]{\text{switch}} \begin{bmatrix} 2 & 3 & 1 \\ 1 & 2 & 1 \end{bmatrix}$$

$$\xrightarrow{R_1 - R_2} \begin{bmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \end{bmatrix}$$

$$\xrightarrow{R_2 - R_1} \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}$$

$$\xrightarrow{R_1 - R_2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \end{bmatrix}$$

$\because$ There's two independent rows

$\therefore$ rank 2.

(2)

$$AA^T = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 1 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} (1)(1) + (2)(2) + (1)(1) & (1)(2) + (2)(3) + (1)(1) \\ (2)(1) + (3)(2) + (1)(1) & (2)(2) + (3)(3) + (1)(1) \end{bmatrix}$$

$$= \begin{bmatrix} 6 & 9 \\ 9 & 14 \end{bmatrix}$$

Eigenvalues:

$$(\lambda \cdot I_2 - AA^T)\bar{v} = \bar{0}$$

$$\begin{bmatrix} \lambda - 6 & -9 \\ -9 & \lambda - 14 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$\det\left(\begin{bmatrix} \lambda - 6 & -9 \\ -9 & \lambda - 14 \end{bmatrix}\right) = 0$$

$$\Rightarrow \quad (\lambda-6)(\lambda-14)-(-9)(-9)=0$$

$$\lambda^2-20\lambda+84-81=0$$

$$\lambda^2-20\lambda+3=0$$

$$\lambda=\frac{20\pm\sqrt{400-4(1)(3)}}{2}$$

$$=10\pm\sqrt{97}$$

$$\because \sigma_n=\sqrt{\lambda_n}$$

$$\therefore \sigma_1=\sqrt{10+\sqrt{97}} \qquad \sigma_2=\sqrt{10-\sqrt{97}}$$

(3) $A^T A$

$$= \begin{bmatrix} 1 & 2 \\ 2 & 3 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} (1)(1)+(2)(2) & (1)(2)+(2)(3) & (1)(1)+(2)(1) \\ (2)(1)+(3)(2) & (2)(2)+(3)(3) & (2)(1)+(3)(1) \\ (1)(1)+(1)(2) & (1)(2)+(1)(3) & (1)(1)+(1)(1) \end{bmatrix}$$

$$= \begin{bmatrix} 5 & 8 & 3 \\ 8 & 13 & 5 \\ 3 & 5 & 2 \end{bmatrix}$$

λ is an eigenvalue of A $\iff$ $A\bar{v} = \lambda \bar{v}$ for nonzero $\bar{v}$

$$\underbrace{(\lambda I_n - A^T A)}\bar{v} = \bar{0}$$

$$\Downarrow$$

$$\det(\lambda I_n - A^T A) = 0$$

$$\lambda I_3 - A^TA = \begin{bmatrix} \lambda-5 & -8 & -3 \\ -8 & \lambda-13 & -5 \\ -3 & -5 & \lambda-2 \end{bmatrix}$$

$$\det(\lambda I_3 - A^TA) = (\lambda-5)(\lambda-13)(\lambda-2) - 120 - 120 - (64(\lambda-2)) - (25(\lambda-5)) - (9(\lambda-13))$$

$$\left[\begin{array}{ccc|cc} \lambda-5 & -8 & -3 & \lambda-5 & -8 \\ -8 & \lambda-13 & -5 & -8 & \lambda-13 \\ -3 & -5 & \lambda-2 & -3 & -5 \end{array}\right.$$

$$= (\lambda-5)(\lambda-13)(\lambda-2) - 64(\lambda-2) - 25(\lambda-5) - 9(\lambda-13) - 240$$

$$= (\lambda^2 - 18\lambda + 65)(\lambda-2) \underbrace{-64\lambda + 128 - 25\lambda + 125 - 9\lambda + 117 - 240}$$

$$= \lambda^3 - 18\lambda^2 + 65\lambda - 2\lambda^2 + 36\lambda - \cancel{130} - 98\lambda + \cancel{130}$$

$$= \lambda^3 - 20\lambda^2 + 3\lambda$$

$$= \lambda(\lambda^2 - 20\lambda + 3)$$

$$\lambda = \frac{20 \pm \sqrt{400 - 4(1)(3)}}{2}, \quad \lambda_3 = 0$$

$$= \frac{20 \pm 2\sqrt{97}}{2}$$

$$= 10 \pm \sqrt{97}$$

$$\lambda_1 = 10 + \sqrt{97}, \quad \lambda_2 = 10 - \sqrt{97}$$

For $\lambda = 10+\sqrt{97}$,

$$\begin{bmatrix} 5+\sqrt{97} & -8 & -3 \\ -8 & -3+\sqrt{97} & -5 \\ -3 & -5 & 8+\sqrt{97} \end{bmatrix} \bar{v} = \bar{0}$$

Row reduction,

$$R_1 = \frac{R_1}{5+\sqrt{97}} \begin{bmatrix} 1 & \frac{-8}{3+\sqrt{97}} & \frac{-3}{5+\sqrt{97}} \\ -8 & -3+\sqrt{97} & -5 \\ -3 & -5 & 8+\sqrt{97} \end{bmatrix}$$

$$R_2 = R_2 + 8R_1 \begin{bmatrix} 1 & \frac{5-\sqrt{97}}{9} & \frac{5-\sqrt{97}}{24} \\ 0 & \frac{\sqrt{97}+13}{9} & -\frac{\sqrt{97}+10}{3} \\ -3 & -5 & 8+\sqrt{97} \end{bmatrix}$$

$$R_3 = R_3 + 3R_1 \begin{bmatrix} 1 & \frac{5-\sqrt{97}}{9} & \frac{5-\sqrt{97}}{24} \\ 0 & \frac{\sqrt{97}+13}{9} & -\frac{\sqrt{97}+10}{3} \\ 0 & -\frac{\sqrt{97}+10}{3} & \frac{7\sqrt{97}+69}{8} \end{bmatrix}$$

$$R_2 = \frac{9}{\sqrt{97}+13} R_2 \begin{bmatrix} 1 & \frac{5-\sqrt{97}}{9} & \frac{5-\sqrt{97}}{24} \\ 0 & 1 & -\frac{\sqrt{97}+11}{8} \\ 0 & \frac{-\sqrt{97}-10}{3} & \frac{7\sqrt{97}+69}{8} \end{bmatrix}$$

$$R_1 = R_1 + \frac{-5+\sqrt{97}}{9} R_2 \begin{bmatrix} 1 & 0 & \frac{-3-\sqrt{97}}{8} \\ 0 & 1 & \frac{-\sqrt{97}-11}{8} \\ 0 & \frac{-\sqrt{97}-10}{3} & \frac{7\sqrt{97}+69}{8} \end{bmatrix}$$

$$R_3 = R_3 + \frac{\sqrt{97}+10}{3} R_2 \begin{bmatrix} 1 & 0 & \frac{-3-\sqrt{97}}{8} \\ 0 & 1 & \frac{-\sqrt{97}-11}{8} \\ 0 & 0 & 0 \end{bmatrix}$$

For $\lambda = 10-\sqrt{97}$

$$\begin{bmatrix} 5-\sqrt{97} & -8 & -3 \\ -8 & -3-\sqrt{97} & -5 \\ -3 & -5 & 8-\sqrt{97} \end{bmatrix} \bar{v} = \bar{0}$$

Row reduction,

$$R_1 = \frac{R_1}{5-\sqrt{97}} \begin{bmatrix} 1 & \frac{-8}{5-\sqrt{97}} & \frac{-3}{5-\sqrt{97}} \\ -8 & -3-\sqrt{97} & -5 \\ -3 & -5 & 8-\sqrt{97} \end{bmatrix}$$

$$R_2 = R_2 + 8R_1 \begin{bmatrix} 1 & \frac{5+\sqrt{97}}{9} & \frac{5+\sqrt{97}}{24} \\ 0 & \frac{13-\sqrt{97}}{9} & \frac{-10+\sqrt{97}}{3} \\ -3 & -5 & 8-\sqrt{97} \end{bmatrix}$$

$$R_3 = R_3 + 3R_1 \begin{bmatrix} 1 & \frac{5+\sqrt{97}}{9} & \frac{5+\sqrt{97}}{24} \\ 0 & \frac{13-\sqrt{97}}{9} & \frac{-10+\sqrt{97}}{3} \\ 0 & \frac{-10+\sqrt{97}}{3} & \frac{69-7\sqrt{97}}{8} \end{bmatrix}$$

$$R_2 = \frac{9}{13-\sqrt{97}} R_2 \begin{bmatrix} 1 & 0 & \frac{-3+\sqrt{97}}{8} \\ 0 & 1 & \frac{-11+\sqrt{97}}{8} \\ 0 & \frac{-10+\sqrt{97}}{3} & \frac{69-7\sqrt{97}}{8} \end{bmatrix}$$

$$R_1 = R_1 - \frac{5+\sqrt{97}}{9} R_2 \begin{bmatrix} 1 & 0 & \frac{-3+\sqrt{97}}{8} \\ 0 & 1 & \frac{-11+\sqrt{97}}{8} \\ 0 & \frac{-10+\sqrt{97}}{3} & \frac{69-7\sqrt{97}}{8} \end{bmatrix}$$

$$R_3 = R_3 - \frac{-10+\sqrt{97}}{3} R_2 \begin{bmatrix} 1 & 0 & \frac{-3+\sqrt{97}}{8} \\ 0 & 1 & \frac{-11+\sqrt{97}}{8} \\ 0 & 0 & 0 \end{bmatrix}$$

For $\lambda = 0$

$$\begin{bmatrix} -5 & -8 & -3 \\ -8 & -13 & -5 \\ -3 & -5 & -2 \end{bmatrix} \bar{v} = \bar{0}$$

$$R_1 = -\frac{R_1}{5} \begin{bmatrix} 1 & \frac{8}{5} & \frac{3}{5} \\ -8 & -13 & -5 \\ -3 & -5 & -2 \end{bmatrix}$$

$$R_2 = R_2 + 8R_1 \begin{bmatrix} 1 & \frac{8}{5} & \frac{3}{5} \\ 0 & -\frac{1}{5} & -\frac{1}{5} \\ -3 & -5 & -2 \end{bmatrix}$$

$$R_3 = R_3 + 3R_1 \begin{bmatrix} 1 & \frac{8}{5} & \frac{3}{5} \\ 0 & -\frac{1}{5} & -\frac{1}{5} \\ 0 & -\frac{1}{5} & -\frac{1}{5} \end{bmatrix}$$

$$R_2 = -5R_2 \begin{bmatrix} 1 & \frac{8}{5} & \frac{3}{5} \\ 0 & 1 & 1 \\ 0 & -\frac{1}{5} & -\frac{1}{5} \end{bmatrix}$$

$$R_1 = R_1 - \frac{8R_2}{5} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & -\frac{1}{5} & -\frac{1}{5} \end{bmatrix}$$

$$R_3 = R_3 + \frac{R_2}{5} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

To find eigenvector,

$$\begin{bmatrix} 1 & 0 & \frac{-3-\sqrt{97}}{8} \\ 0 & 1 & \frac{-\sqrt{97}-11}{8} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$V_1 + \frac{-3-\sqrt{97}}{8} V_3 = 0 \Rightarrow V_1 = \frac{3+\sqrt{97}}{8} V_3$$

$$V_2 + \frac{-\sqrt{97}-11}{8} V_3 = 0 \Rightarrow V_2 = \frac{11+\sqrt{97}}{8} V_3$$

$\Rightarrow$ $\because V_3$ is trivial, $\therefore$ let $V_3 = 1$

$$\overline{V}_{\lambda_1} = \begin{bmatrix} \frac{3+\sqrt{97}}{8} \\ \frac{11+\sqrt{97}}{8} \\ 1 \end{bmatrix}$$

To find eigenvector,

$$\begin{bmatrix} 1 & 0 & \frac{-3+\sqrt{97}}{8} \\ 0 & 1 & \frac{-11+\sqrt{97}}{8} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$V_1 + \frac{-3+\sqrt{97}}{8} V_3 = 0 \qquad V_2 + \frac{-11+\sqrt{97}}{8} V_3 = 0$$

$$\Rightarrow V_1 = \frac{3-\sqrt{97}}{8} V_3 \qquad \Rightarrow V_2 = \frac{11-\sqrt{97}}{8} V_3$$

$\because V_3$ is trivial

$\therefore$ let's assume $V_3 = 1$

$\therefore$ The eigenvector for λ_2 is

$$\overline{V}_{\lambda_2} = \begin{bmatrix} \frac{3-\sqrt{97}}{8} \\ \frac{11-\sqrt{97}}{8} \\ 1 \end{bmatrix}$$

To find eigenvector,

$$\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$V_1 - V_3 = 0 \Rightarrow V_1 = V_3$$

$$V_2 + V_3 = 0 \Rightarrow V_2 = -V_3$$

$\because V_3$ is trivial $\therefore$ let $V_3 = 1$

$\Rightarrow$ Eigenvector for λ_3 is

$$\overline{V}_{\lambda_3} = \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}$$

From the eigenvectors found above and normalizing to length,

$$V = \begin{bmatrix} \frac{3+\sqrt{97}}{2\sqrt{7\sqrt{97}+97}} & \frac{3-\sqrt{97}}{2\sqrt{97-7\sqrt{97}}} & \frac{\sqrt{3}}{3} \\ \frac{\sqrt{97}+11}{2\sqrt{7\sqrt{97}+97}} & \frac{11-\sqrt{97}}{2\sqrt{97-7\sqrt{97}}} & -\frac{\sqrt{3}}{3} \\ \frac{4}{\sqrt{7\sqrt{97}+97}} & \frac{4}{\sqrt{97-7\sqrt{97}}} & \frac{\sqrt{3}}{3} \end{bmatrix}$$

To get U,

$$① \; u_1 = \frac{1}{\sigma_1} A v_1 = \frac{1}{\sqrt{10+\sqrt{97}}} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} \frac{3+\sqrt{97}}{2\sqrt{7\sqrt{97}+97}} \\ \frac{\sqrt{97}+11}{2\sqrt{7\sqrt{97}+97}} \\ \frac{4}{\sqrt{7\sqrt{97}+97}} \end{bmatrix}$$

$$= \frac{1}{\sqrt{10+\sqrt{97}}} \begin{bmatrix} (1)\left(\frac{3+\sqrt{97}}{2\sqrt{7\sqrt{97}+97}}\right) + (2)\left(\frac{\sqrt{97}+11}{2\sqrt{7\sqrt{97}+97}}\right) + (1)\left(\frac{4}{\sqrt{7\sqrt{97}+97}}\right) \\ (2)\left(\frac{3+\sqrt{97}}{2\sqrt{7\sqrt{97}+97}}\right) + (3)\left(\frac{\sqrt{97}+11}{2\sqrt{7\sqrt{97}+97}}\right) + (1)\left(\frac{4}{\sqrt{7\sqrt{97}+97}}\right) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{3(\sqrt{97}+11)}{2\sqrt{167\sqrt{97}+1649}} \\ \frac{47+5\sqrt{97}}{2\sqrt{167\sqrt{97}+1649}} \end{bmatrix}$$

$$② \; u_2 = \frac{1}{\sigma_2} \cdot A \cdot V_2 = \frac{1}{\sqrt{10-\sqrt{91}}} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & 1 \end{bmatrix} \begin{bmatrix} \frac{3-\sqrt{91}}{2\sqrt{91-7\sqrt{91}}} \\ \frac{11-\sqrt{91}}{2\sqrt{91-7\sqrt{91}}} \\ \frac{4}{\sqrt{91-7\sqrt{91}}} \end{bmatrix}$$

$$= \begin{bmatrix} \frac{33-3\sqrt{91}}{2\sqrt{1649-161\sqrt{91}}} \\ \frac{41-5\sqrt{91}}{2\sqrt{1649-161\sqrt{91}}} \end{bmatrix}$$

$$\therefore \; U = \begin{bmatrix} \frac{3(\sqrt{91}+11)}{2\sqrt{161\sqrt{91}+1649}} & \frac{33-3\sqrt{91}}{2\sqrt{1649-161\sqrt{91}}} \\ \frac{41+5\sqrt{91}}{2\sqrt{161\sqrt{91}+1649}} & \frac{41-5\sqrt{91}}{2\sqrt{1649-161\sqrt{91}}} \end{bmatrix}$$