Rmeans Problem Given: D = ldi, ..., dn) data points d,=(d;2).-,d;d) , (I) Assume K clusters with 'centers' (certionals) M, M, M, ERd Z distance (di, le;) di is assynd to chorry $= \sum_{i=1}^{n} \sum_{j=1}^{n} ||a_{ij} - \mu_{jj}||^{2} = \int_{-\infty}^{\infty} ||a_{ij} - \mu_{jj}||^$ j=1 i such tn.t di is assigned to chaterj

RMEGNS Problem Given: D = (di), dn) data points d,=(d;1, .- ,d;d) Centrolls M., Mz, --., M. ERª E distance (di, lej) di is assand to choterj $= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} ||a_{i} - \mu_{j}||^{2} = \int_{0}^{\infty} =$ di is assand to choter; $= \sum_{i=1}^{k} \sum_{i=1}^{n} a_{ij} \cdot \|a_{i} - u_{j}\|^{2} \quad \text{where}$ azj = S | cssicned to cluster J otherwise

Rmeans Problem Given: D = (d1)..., dn) data points d,=(d;2,..,d;d) Centrolls M., Mz, --., M. EM (Unknown) $\int_{i=1}^{k} \sum_{i=2}^{n} a_{ij} \cdot ||a_{i} - u_{j}||^{2} \quad \text{where} \quad$ ajj = Si di is

ajj = Si assigned to

cluster J

otherwise If we were given the mi, lejsk, it's easy to compute 2 Nd, - us 12

RMEGNS Problem Given: D = ldis..., dn) data points d, = (d; 1, - , d; d) Centroils M, M2, -- , M2 & TRd (Unknown) $\int_{j=1}^{k} \sum_{i=1}^{n} a_{ij} \cdot ||a_{i} - u_{j}||^{2}$ Algo, item D'A-bitraily select centroids Mr. MK Assignment Step Assign each di to the closest controid Desompré controis me mu Repeat until the change in Lis sufficiently small.

kmeans Problem Given: D = ldis..., dn) data points d, = (d; 1, - , d; d) Centrolls M., Mz, --., M.z & TRd (Unknown) $\int_{j=1}^{k} \sum_{i=1}^{n} a_{ij} \cdot || d_{i} - u_{j}||^{2}$ Assignment Ster Dissignment Ster Assignment Ster di to the closest controid c; = { d; | ||d;, u; || < ||d;, 4p||, 1 < p < k } So Cj. is all the data points assigned to the jt cluster.

Remeans Problem

Given: $D \in [d_1, \dots, d_n]$ data points $d_1 = [d_{12}, \dots, d_{1d}]$ Centrolls $M_1, M_{23}, \dots, M_{2r} \in \mathbb{R}^d$ (Unknown) $\int_{j=1}^{k} \sum_{i=1}^{n} a_{ij} \cdot ||d_{i} - M_{j}||^2$

1) Rezompnte controids me ma

Updated $mj = \frac{1}{|C_3|}$. $\sum_{dj \in C_j} dj$

ussigned $d_1 = (4, 8, 16)$ to My $d_2 = (3, 9, 2)$ $d_3 = (4, 7, 5)$ $M_3 = \frac{1}{2} \cdot (11, 24, 23) = (11/3, 8, 23/3)$

Rearson Correlation

$$\begin{cases} 1 & 1 \\ 1 & 1 \\ 2 & 1 \end{cases} = \begin{cases} 2 & 1 \end{cases} = \begin{cases} 2 & 1 \\ 2 & 1 \end{cases} = \begin{cases} 2 & 1 \\ 2 & 1 \end{cases} = \begin{cases} 2 & 1 \\ 2 & 1 \end{cases} = \begin{cases} 2 & 1 \end{cases} = \begin{cases} 2 & 1 \\ 2 & 1 \end{cases} = \begin{cases} 2 & 1 \\ 2 & 1 \end{cases} = \begin{cases} 2 & 1 \end{cases} = \begin{cases} 2 & 1 \\ 2 & 1 \end{cases} = \begin{cases} 2 & 1 \end{cases} = \begin{cases} 2 & 1 \\ 2 & 1 \end{cases} = \begin{cases} 2 & 1 \end{cases} = \begin{cases} 2 & 1 \\ 2 & 1 \end{cases} = \begin{cases} 2 & 1 \end{cases} = \begin{cases} 2$$

$$x = (1, 2, 4, 6)$$

 $y = (2, 4, 8, 16)$

$$x = (1, 2, 4, 6)$$

 $y = (8, 16, 32, 48)$

$$x=(1, 2, 4, 6)$$

 $y=(-1, -2, -4, -7)$

Rearson Correlation

$$\begin{cases} 1 & \text{A.s.} & \text{A.s.} & \text{A.s.} \\ 2 & \text{A.s.} & \text{A.s.} & \text{A.s.} \\ 3 & \text{A.s.} & \text{A.s.} & \text{A.s.} & \text{A.s.} \end{cases}$$

$$\frac{d}{dx} = \frac{d}{dx} = \frac{d}{dx}$$

where X=1. d d i=1

RMEGNS Problem Given: D = ldis..., dn) data points Intuition: Points are assigned to a cluster so Het the distance to the centroid of the cluster is shorter than the distance to any other aluster Az () K clusters with 'centers' (certion d s) M, M2, -- , M, ERª Z distance (di, sej) di is assyned to chaterj