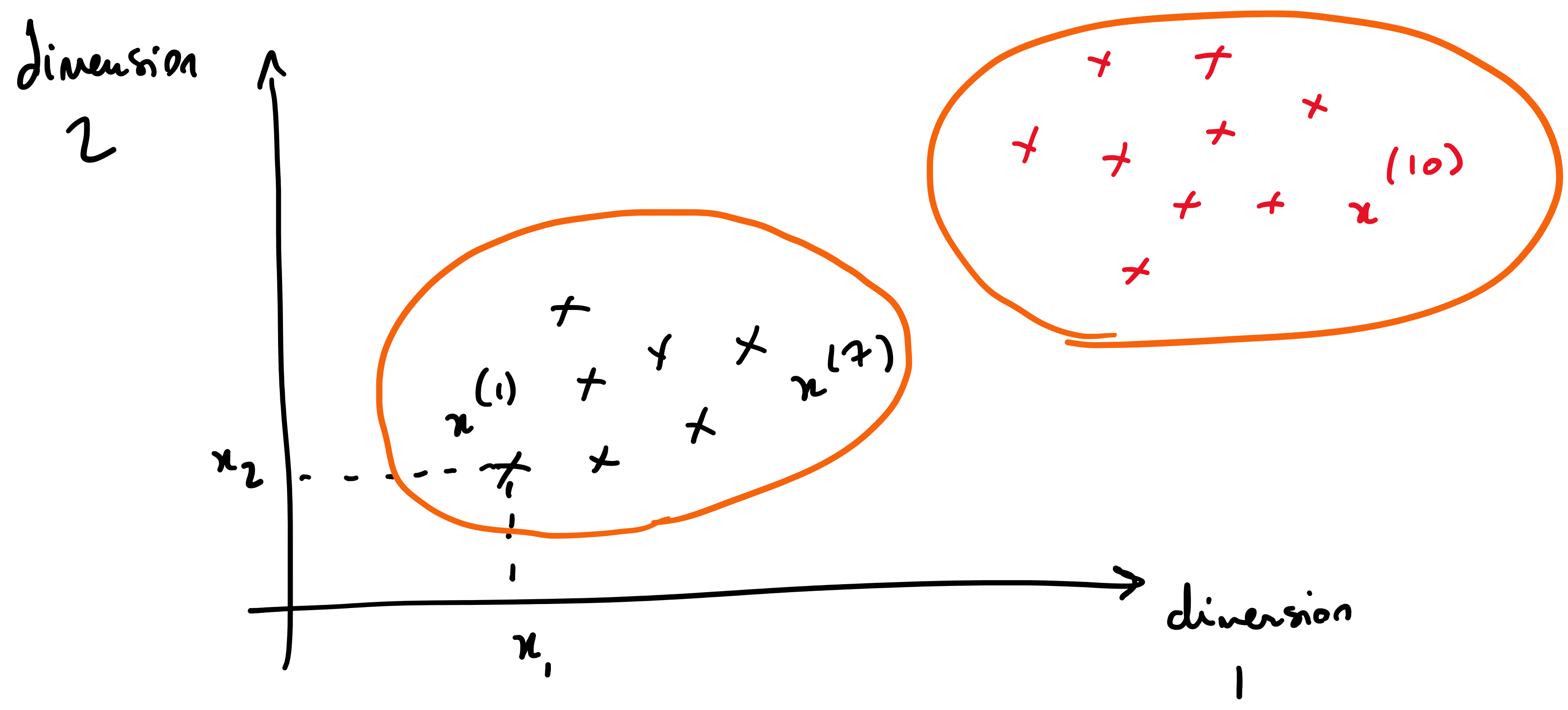
Video 3

Monday, March 30, 2020 1:28 PM



$x^{(1)} = (x_1^{(1)}, x_2^{(1)})$

$\{ x^{(1)}, x^{(2)}, x^{(3)}, \ldots, x^{(m)} \}$

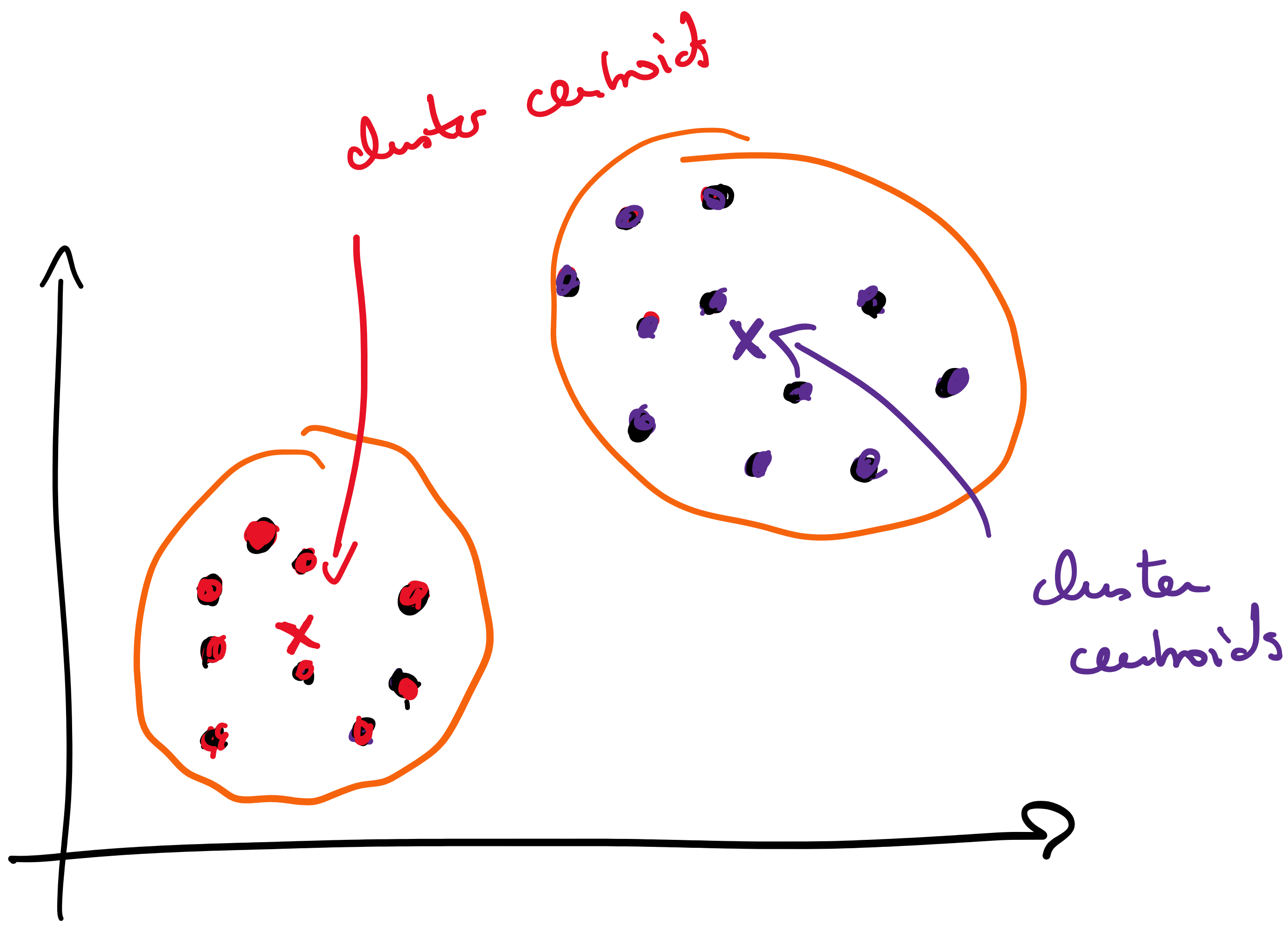
$x^{(3)} = (x_1^{(3)}, x_2^{(3)})$

m : nb points in our dataset

Examples of clustering applications:

- social network analysis
- market segmentation
- organize datacenters

k-means



k-means algorithm

- Inputs
  - k: # of clusters
  - Training set $\{x^{(1)}, \ldots, x^{(m)}\}$

- Random initialization
  k cluster centroids $\mu_1, \mu_2, \ldots, \mu_k$

- Repeat :

Cluster assignment step

for $i \in 1 .. m$

$c^{(i)}$:= index of cluster centroid closest to $x^{(i)}$

$c^{(i)} = \underset{k}{\operatorname{argmin}} \| x^{(i)} - \mu_k \|^2$

Move centroid

for $k \in 1 .. K$

μ_k := mean of points assigned to cluster k

$\mu_k = \text{mean}\{ x^{(i)} : c^{(i)} = k \}$

Non-seperated clusters

S, M, L