Expectation Maximization (EM)

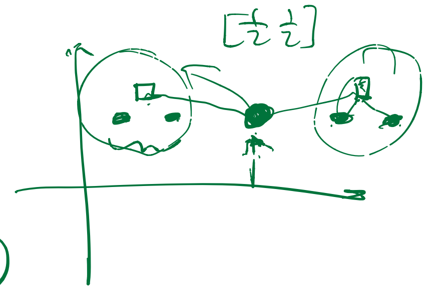
Qiang Liu
UT Austin

K-means vs. EM

- **Inputs**: n objects (data points) $\{x_i\}_{i=1}^{n}$ and a number K of clusters.

- **K-means Algorithm**:
  - Initialization: randomly place K points (as the centroids)
  - Iterate until convergence:
    - i) Assign each object to the group that has the closest centroid

$$z_i = \underset{k=1,\ldots,K}{\arg\min} ||x_i - \mu_k||^2$$



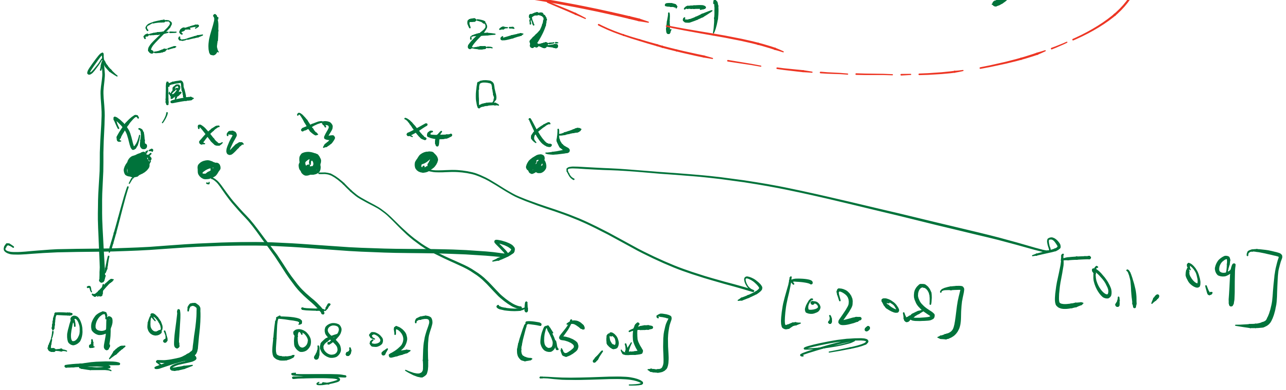
$$\text{prob}(z_i = k) = \frac{\exp(-||x_i - \mu_k||^2/\lambda)}{\sum_{\ell=1}^{K} \exp(-||x_i - \mu_\ell||^2/\lambda)}$$

    - ii) Recalculate the position of the K centroids

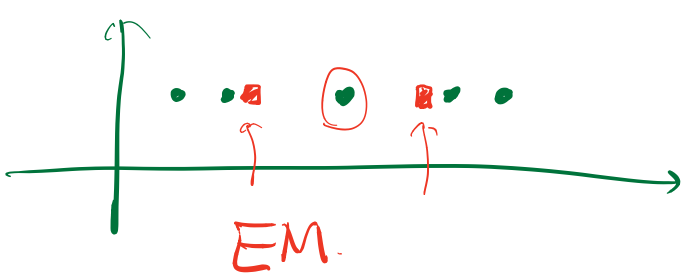
$$\mu_k = \frac{\sum_{i=1}^{n} \mathbb{I}(z_i = k) x_i}{\sum_{i=1}^{n} \mathbb{I}(z_i = k)}$$

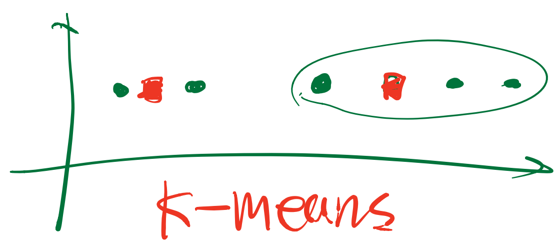
$$\mathbb{I}(z_i = k) = \begin{cases} 1, & z_i = k \\ 0, & z_i \neq k \end{cases}$$

$$\mu_k = \frac{\sum_{i=1}^{n} \text{prob}(z_i = k)\, x_i}{\sum_{i=1}^{n} \text{prob}(z_i = k)}$$

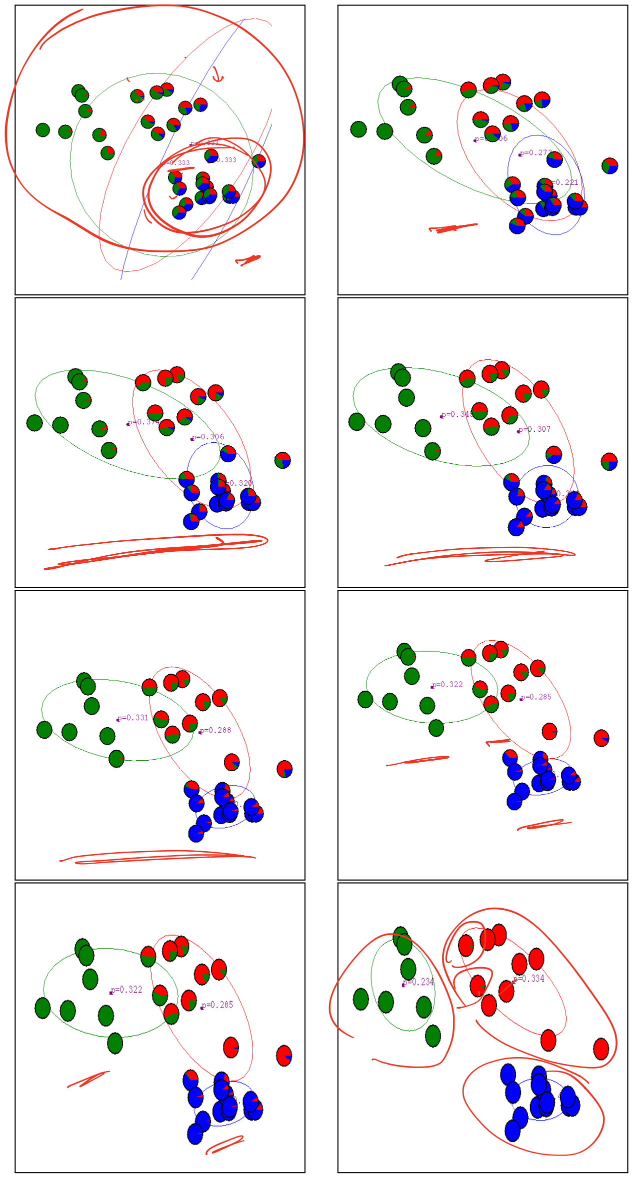


For
$$\begin{cases} \mu_1 = \dfrac{0.9x_1 + 0.8x_2 + 0.5x_3 + 0.2x_4 + 0.1x_5}{0.9 + 0.8 + 0.5 + 0.2 + 0.1} \\[2em] \mu_2 = \dfrac{0.1x_1 + 0.2x_2 + 0.5x_3 + 0.8x_4 + 0.9x_5}{0.1 + 0.2 + 0.5 + 0.8 + 0.9} \end{cases}$$





Expectation Maximization



Expectation Maximization as Maximum Likelihood Estimation

- More generally, EM algorithm is a special optimization algorithm for **maximum likelihood estimation of mixture (or latent variable) models**.
- Let us start with **Gaussian mixture models**...

Expectation Maximization as Maximum Likelihood Estimation

- More generally, EM algorithm is a special optimization algorithm for **maximum likelihood estimation of mixture (or latent variable) models**.
- Let us start with **Gaussian mixture models**...

Expectation Maximization as Maximum Likelihood Estimation

Expectation Maximization as Maximum Likelihood Estimation