STAT 215A Fall 2019
Week 10

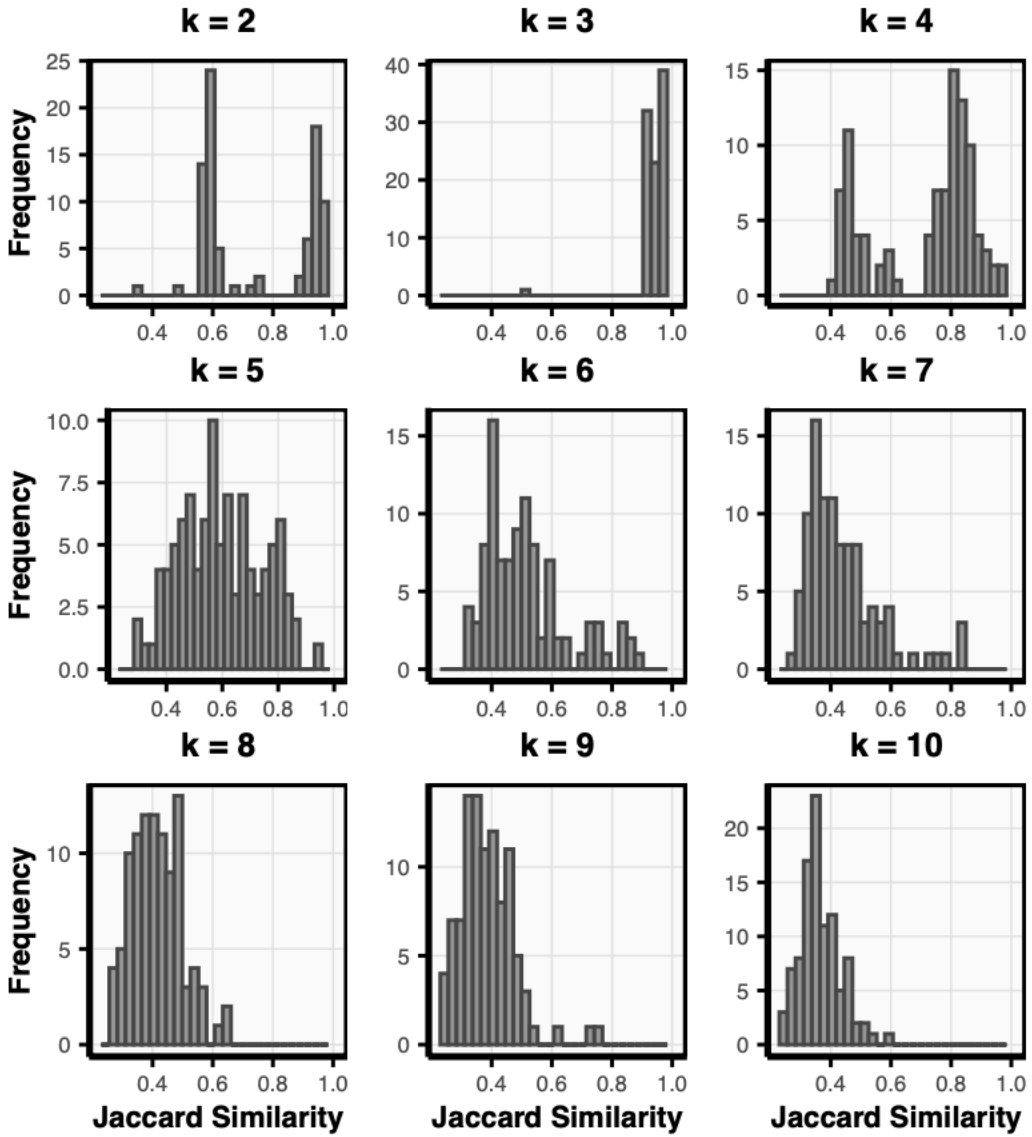
Tiffany Tang

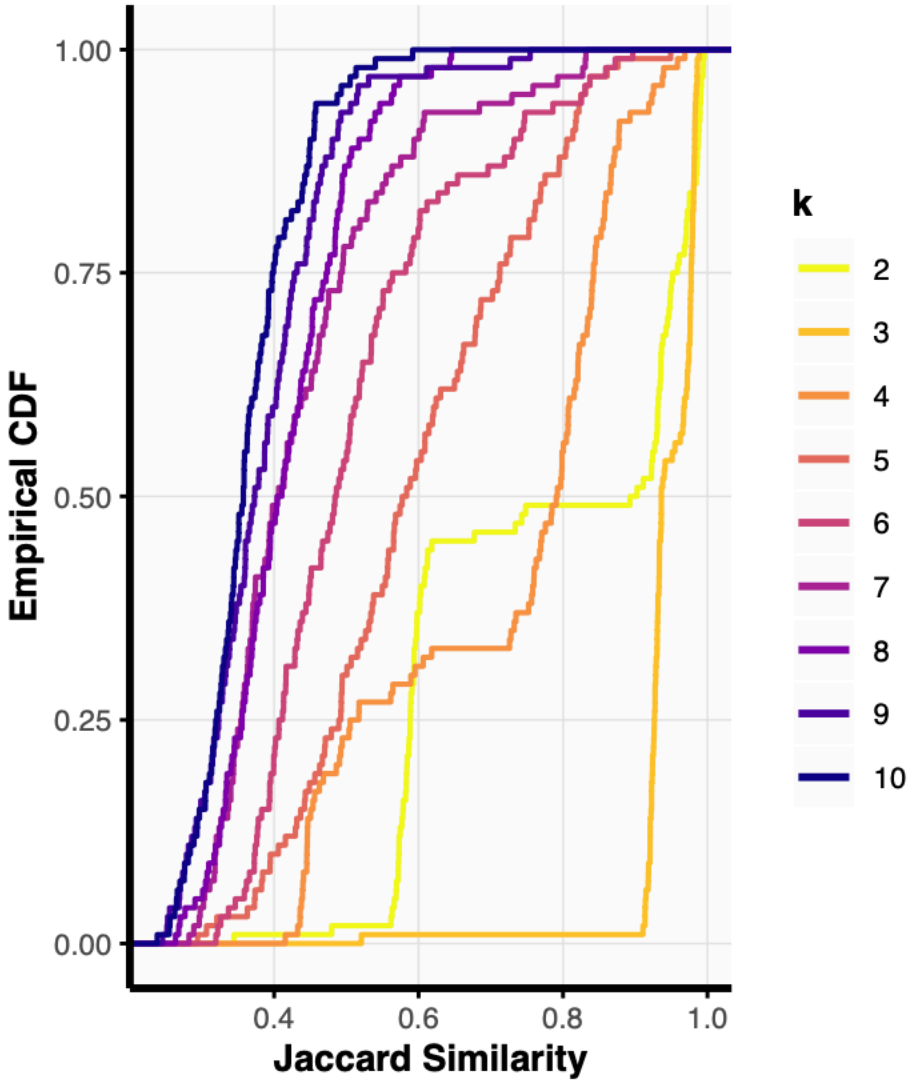
11/1/19

Announcements

- Midterm is over! No more exams for this class!
- Lab 4 has been postponed and will be released next Friday, November 8
  - Group project
  - I will tell you your groups during lab section next Friday, so if possible, please come
  - Have two weeks to complete
- Thank you for the feedback from last time!

Lab 3: Stability of K-means





Plan for Today

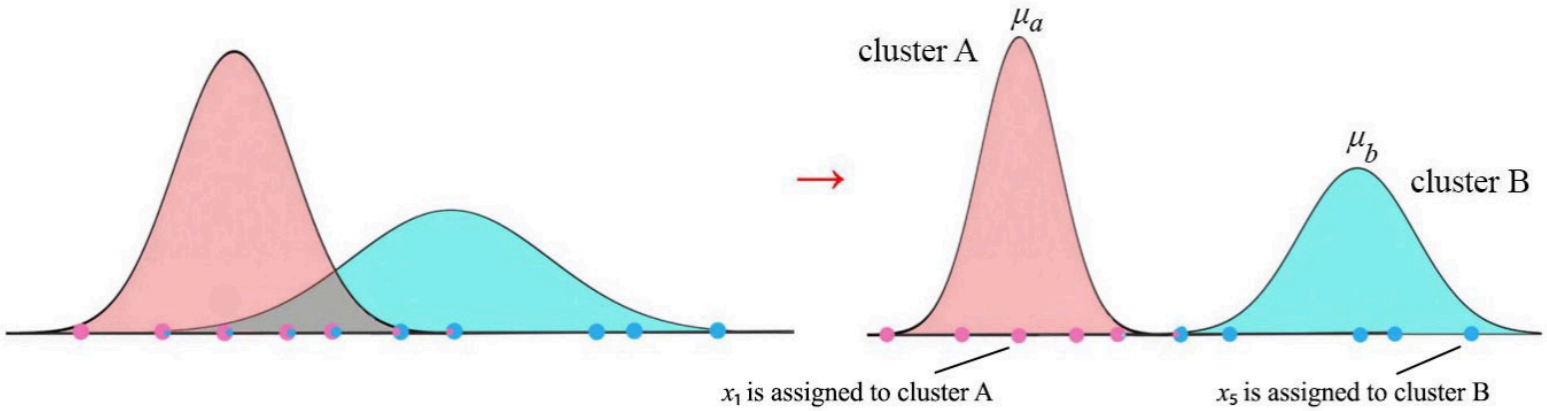
- EM Algorithm

EM Algorithm: Motivation

- **EM Algorithm = Expectation Maximization Algorithm**
- Can serve as a tool for
  - Finding (local) solution to “hard” maximum likelihood problem via “missing data” framework
    - Missing observations (at random)
    - Missing latent (unobserved) variables
  - Clustering

EM Algorithm: Motivation

- Consider the following scenario:
  - There are two clusters of data
  - We observe data points X_1, …, X_n but don't know which cluster each observation belongs to
  - Observed data: X_1, …, X_n
  - Latent variables (i.e., cluster assignments): Z_1, …, Z_n
  - **Goal:** want to solve some maximum likelihood problem to find the cluster labels Z_1, …, Z_n which are the most probable given some underlying model



EM Algorithm: General Setup

- **Observed data:** X
- **Goal:** want to estimate some parameter of interest $\widehat{\boldsymbol{\theta}}$ via maximum likelihood estimation:

$$argmax_\theta \; L(\theta; X)$$

or equivalently, if $X = (X_1, \dots, X_n) \sim p_\theta(X)$ iid,

$$argmax_\theta \; p_\theta(X) = \prod_{i=1}^{n} p_\theta(X_i)$$

or equivalently,

$$argmax_\theta \; \log(p_\theta(X)) = \sum_{i=1}^{n} \log(p_\theta(X_i))$$

EM Algorithm: General Setup

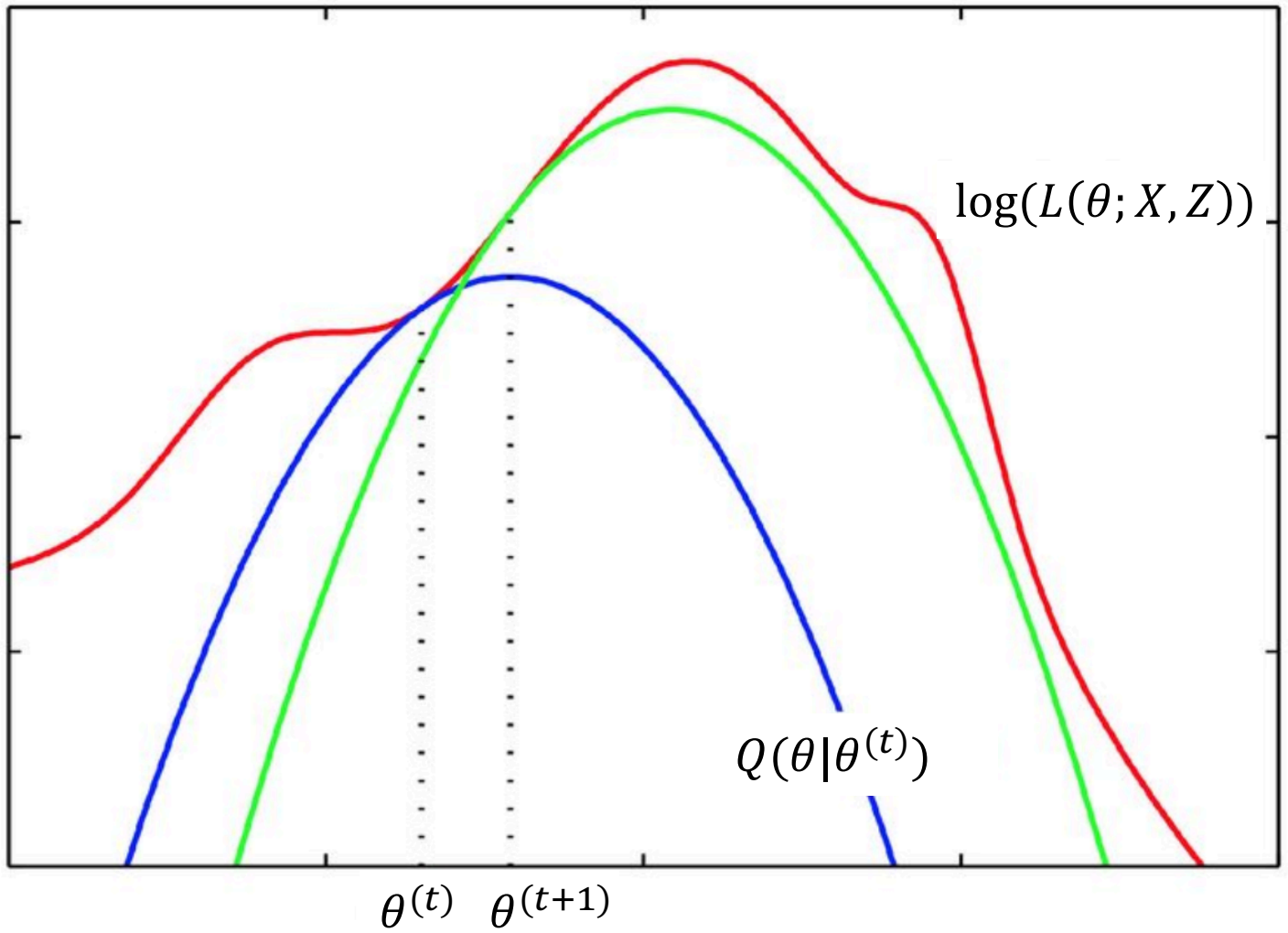
- **Observed Data:** X
- **Unobserved Data:** Z
- **Goal:** want to estimate some parameter of interest $\hat{\boldsymbol{\theta}}$ via maximum likelihood estimation:

$$argmax_\theta \; L(\theta; X, \; Z)$$

or equivalently,

$$argmax_\theta \; \log(p_\theta(X, Z)) = \sum_{i=1}^{n} \log(p_\theta(X_i, Z_i))$$

EM Algorithm: Some intuition



EM Algorithm

Initialize $\theta^{(0)}$

Repeat for t = 1, …, T (or until convergence):

1. **E Step:** Expectation Step
   - Given $\theta^{(t)}$, compute $Q\left(\theta\middle|\theta^{(t)}\right) = E\left[\log\left(L(\theta; X, Z)\right)\middle|\theta^{(t)}, X\right]$

2. **M Step:** Maximization Step
   - Update $\theta^{(t+1)}$ by maximizing

$$\theta^{(t+1)} = argmax_{\theta}\, Q\left(\theta\middle|\theta^{(t)}\right)$$

(i.e., take derivative of $Q\left(\theta\middle|\theta^{(t)}\right)$ with respect to θ and set equal to 0)