



EM for Gaussian mixtures

9 questions

1
point

1.

(True/False) While the EM algorithm maintains uncertainty about the cluster assignment for each observation via soft assignments, the model assumes that every observation comes from only one cluster.

☐ True

☐ False

1
point

2.

(True/False) In high dimensions, the EM algorithm runs the risk of setting cluster variances to zero.

☐ True

☐ False

1
point

3.

In the EM algorithm, what do the E step and M step represent, respectively?

☐ Estimate cluster responsibilities, **M**aximize likelihood over parameters

☐

- ☐ Estimate likelihood over parameters, **Maximize** cluster responsibilities
 - ☐ Estimate number of parameters, **Maximize** likelihood over parameters
 - ☐ Estimate likelihood over parameters, **Maximize** number of parameters
-

1
point

4.

Suppose we have data that come from a mixture of 6 Gaussians (i.e., that is the true data structure). Which model would we expect to have the highest log-likelihood after fitting via the EM algorithm?

- ☐ A mixture of Gaussians with 2 component clusters
 - ☐ A mixture of Gaussians with 4 component clusters
 - ☐ A mixture of Gaussians with 6 component clusters
 - ☐ A mixture of Gaussians with 7 component clusters
 - ☐ A mixture of Gaussians with 10 component clusters
-

1
point

5.

Which of the following correctly describes the differences between EM for mixtures of Gaussians and k-means? Choose all that apply.

- ☐ k-means often gets stuck in a local minimum, while EM tends not to
 - ☐ EM is better at capturing clusters of different sizes and orientations
 - ☐ EM is better at capturing clusters with overlaps
 - ☐ EM is less prone to overfitting than k-means
 - ☐ k-means is equivalent to running EM with infinitesimally small diagonal covariances.
-

1
point

6.

Suppose we have a Gaussian mixture model of 3-dimensional data with 4 component clusters, and we use a model with full covariance matrices. How many parameters are in the model?

40

1
point

7.

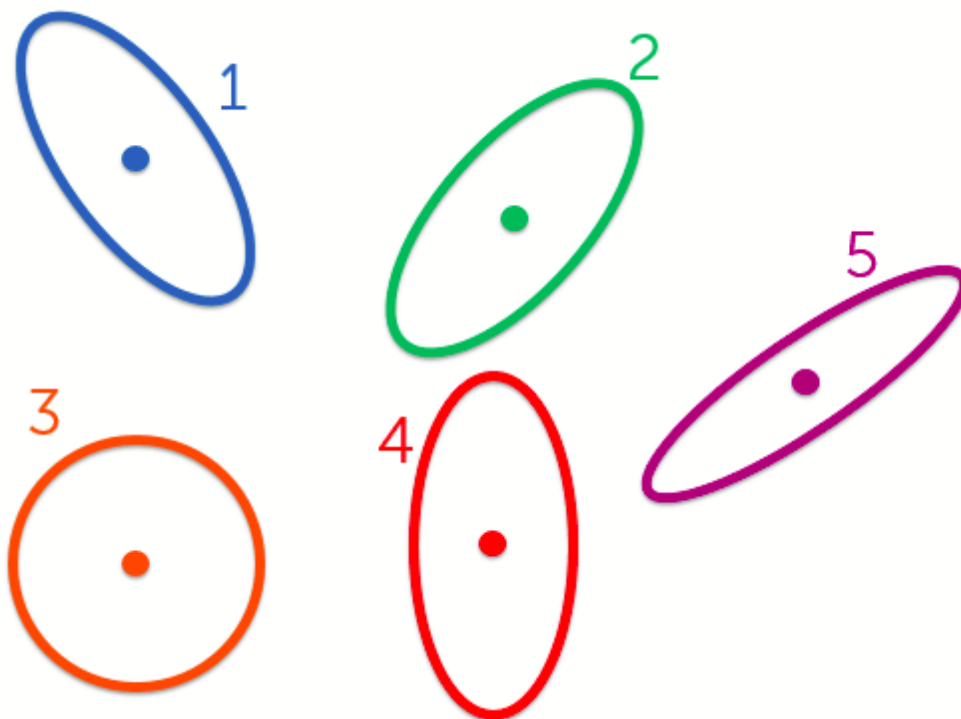
Suppose we have a Gaussian mixture model of 4-dimensional data with 5 component clusters, and we instead assume diagonal covariance matrices. How many parameters are in the model?

45

1
point

8.

Which of the following contour plots describes a Gaussian distribution with diagonal covariance? Choose all that apply.

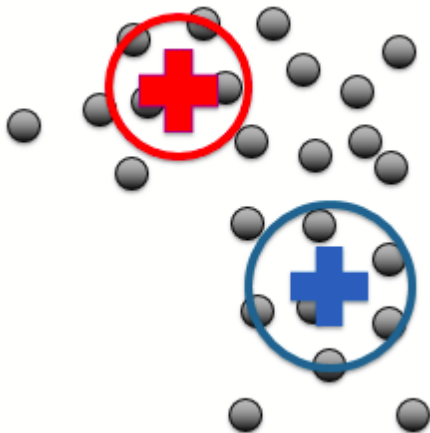


- ☐ (1)
- ☐ (2)
- ☐ (3)
- ☐ (4)
- ☐ (5)

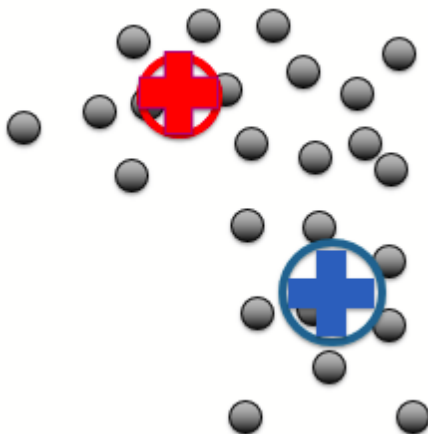
2
points

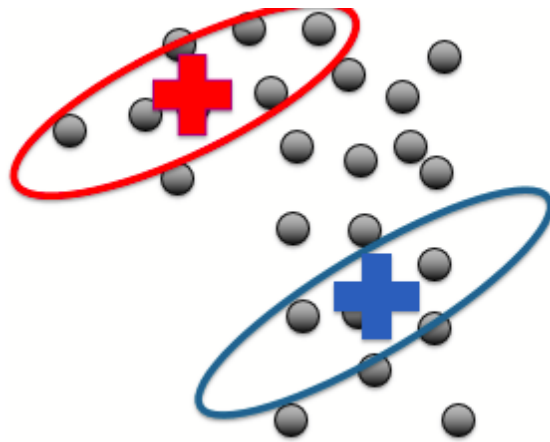
9.

Suppose we initialize EM for mixtures of Gaussians (using full covariance matrices) with the following clusters:

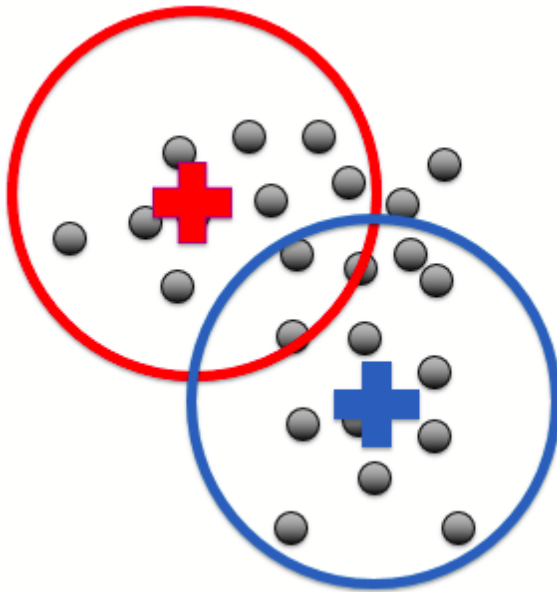


Which of the following best describes the updated clusters after the first iteration of EM?



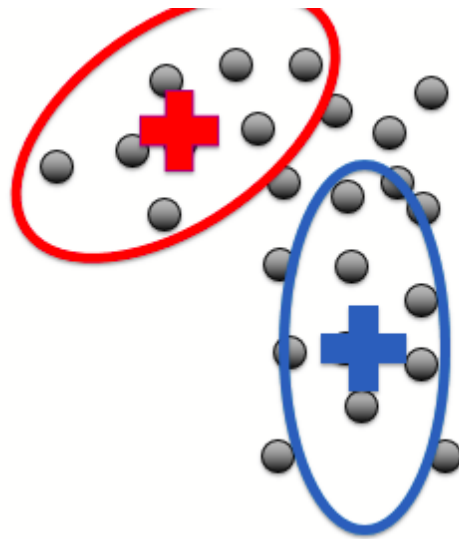


○

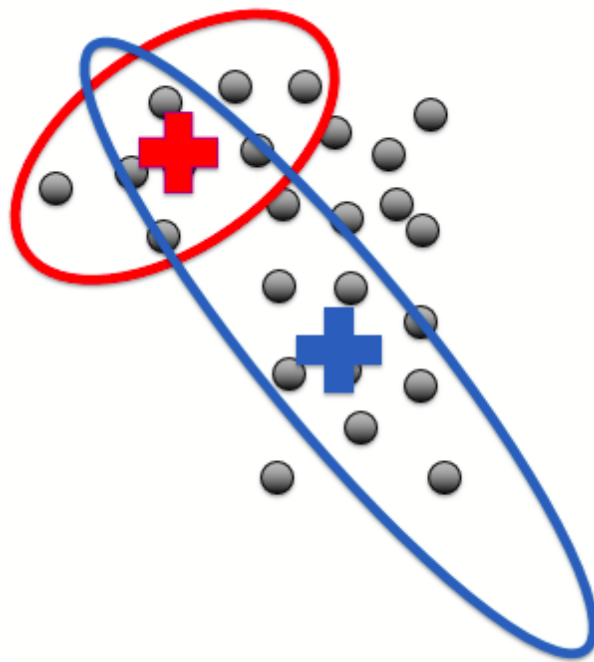


○





○



Submit Quiz

