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?

- **E**stimate cluster responsibilities, **M**aximize likelihood over parameters

- **E**stimate likelihood over parameters, **M**aximize cluster responsibilities
- **E**stimate number of parameters, **M**aximize likelihood over parameters
- **E**stimate likelihood over parameters, **M**aximize 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 are running the EM algorithm. After an E-step, we obtain the following responsibility matrix:

Cluster responsibilities	Cluster A	Cluster B	Cluster C
Data point 1	0.20	0.40	0.40
Data point 2	0.50	0.10	0.40
Data point 3	0.70	0.20	0.10

Which is the **most probable** cluster for data point 3?

- [ ] Cluster A
- [ ] Cluster B
- [ ] Cluster C

---

1 point

7.

Suppose we are running the EM algorithm. After an E-step, we obtain the following responsibility matrix:

Cluster responsibilities	Cluster A	Cluster B	Cluster C
Data point 1	0.20	0.40	0.40
Data point 2	0.50	0.10	0.40
Data point 3	0.70	0.20	0.10

Suppose also that the data points are as follows:

Dataset	X	Y	Z
Data point 1	3	1	2
Data point 2	0	0	3
Data point 3	1	3	7

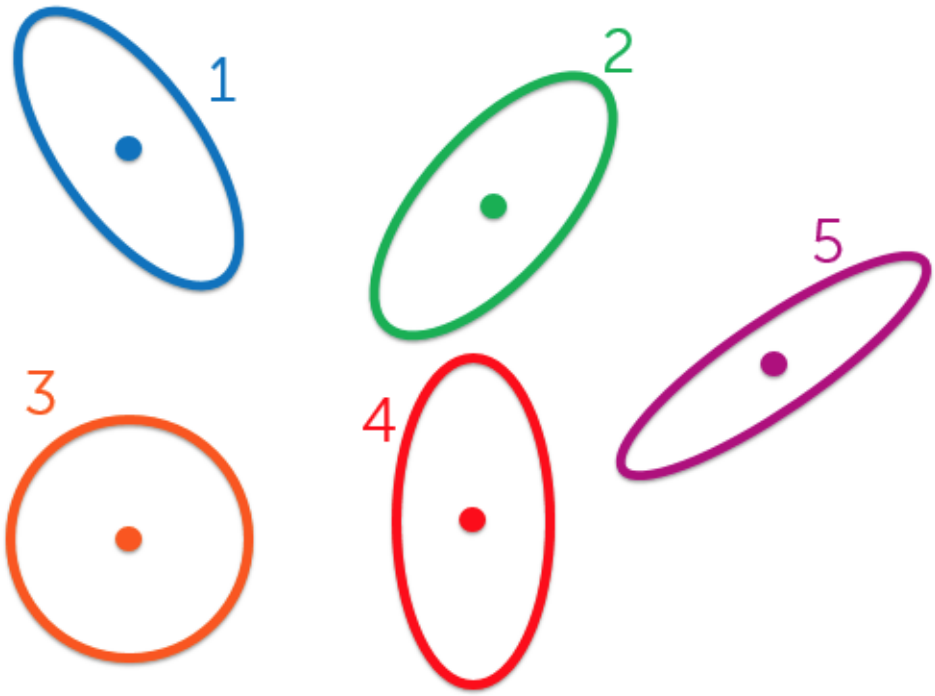
Let us compute the new mean for Cluster A. What is the **Z coordinate** of the new mean? Round your answer to 3 decimal places.

6.800

1 point

8.

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

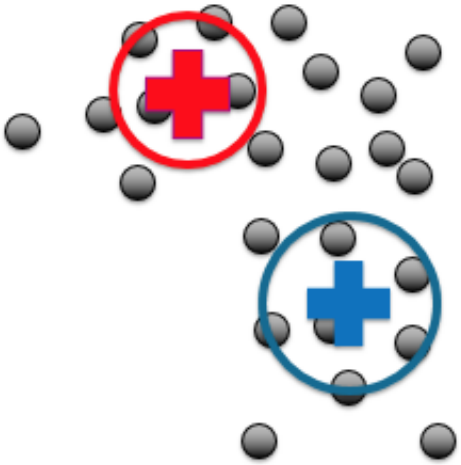


- [ ] (1)
- [ ] (2)
- [x] (3)
- [x] (4)
- [ ] (5)

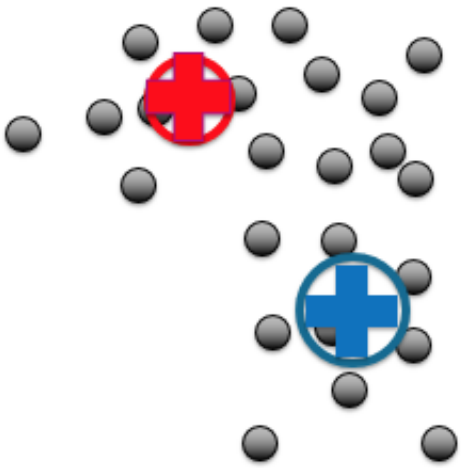
2 points

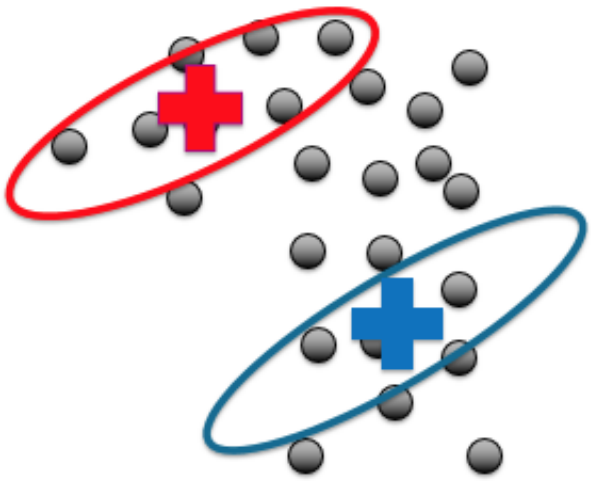
9.

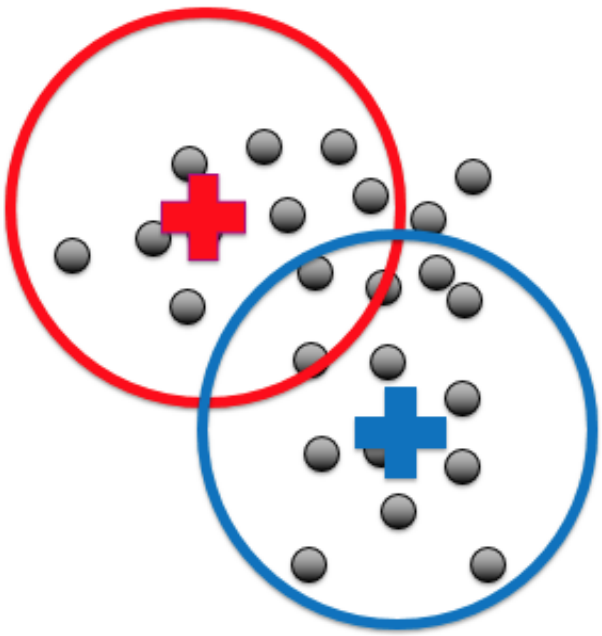
Suppose we initialize EM for mixtures of Gaussians (using full covariance matrices) with the following clusters:

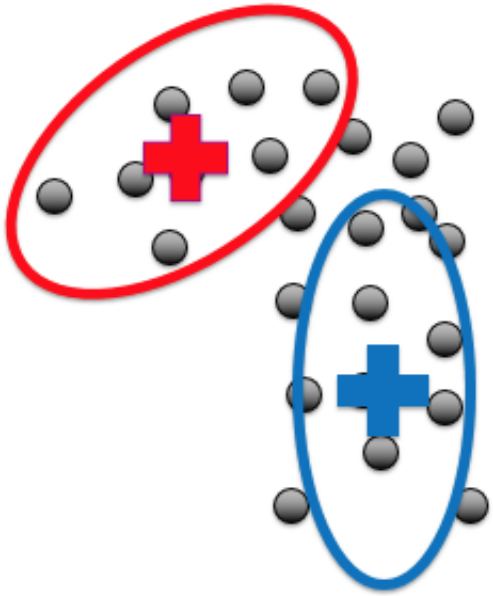


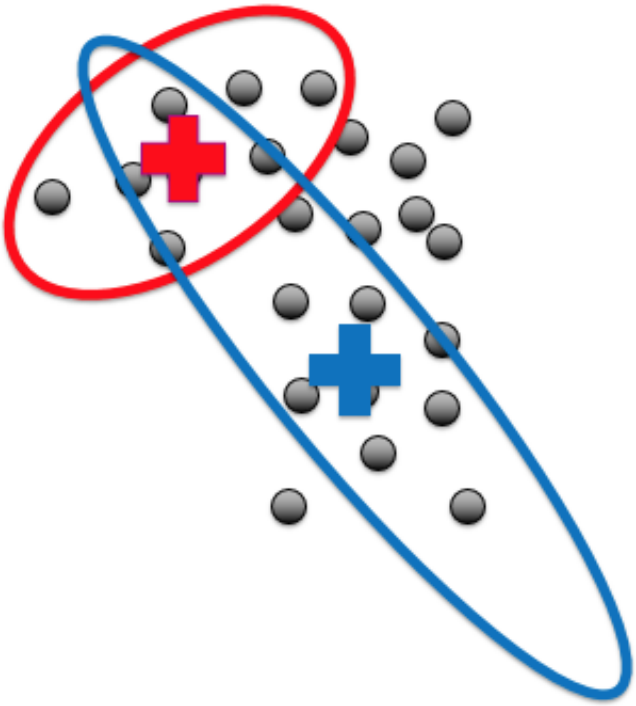
Which of the following best describes the updated clusters after the first iteration of EM?











I, **Devendra Singh Sachan**, understand that submitting work that isn't my own may result in permanent failure of this course or deactivation of my Coursera account.
Learn more about Coursera's Honor Code

Submit Quiz