

7. The K-Means Algorithm: the Big Picture

The K-Means Algorithm: the Big Picture



for every point in the cluster, we will compute its distance from the representative. So again, what are we doing here? We have the partitions. We will go and try to find the best assignments of the representatives in such a way that if for every point we compute its distance from the representative, we achieve the minimum.



[End of transcript. Skip to the start.](#)

Video

[Download video file](#)

Transcripts

[Download SubRip \(.srt\) file](#)

[Download Text \(.txt\) file](#)



The K-Means Algorithm: Step by Step

2/2 points (graded)

In the above lecture, given a set of feature vectors

$$S_n = \{x^{(i)} | i = 1, \dots, n\}$$

and the number of clusters k , we saw that we can use the K-Means algorithm to find reasonably good cluster assignments C_1, \dots, C_K and the representatives of each of the k clusters z_1, \dots, z_K . The algorithm was given like the following:

1. Randomly select z_1, \dots, z_K

2. Iterate

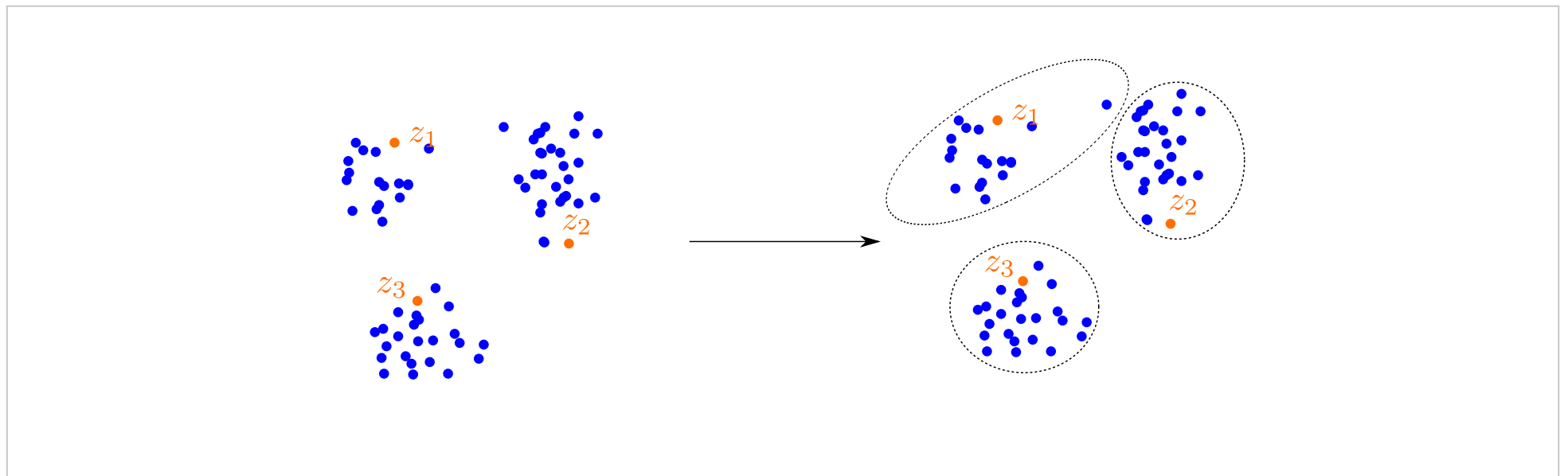
1. Given z_1, \dots, z_K , assign each data point $x^{(i)}$ to the closest z_j , so that

$$\text{Cost}(z_1, \dots, z_K) = \sum_{i=1}^n \min_{j=1, \dots, k} \|x^{(i)} - z_j\|^2$$

2. Given C_1, \dots, C_K find the best representatives z_1, \dots, z_K , i.e. find z_1, \dots, z_K such that

$$z_j = \operatorname{argmin}_z \sum_{i \in C_j} \|x^{(i)} - z\|^2.$$

1. The following figure depicts an example of one of the steps of K-means algorithm:



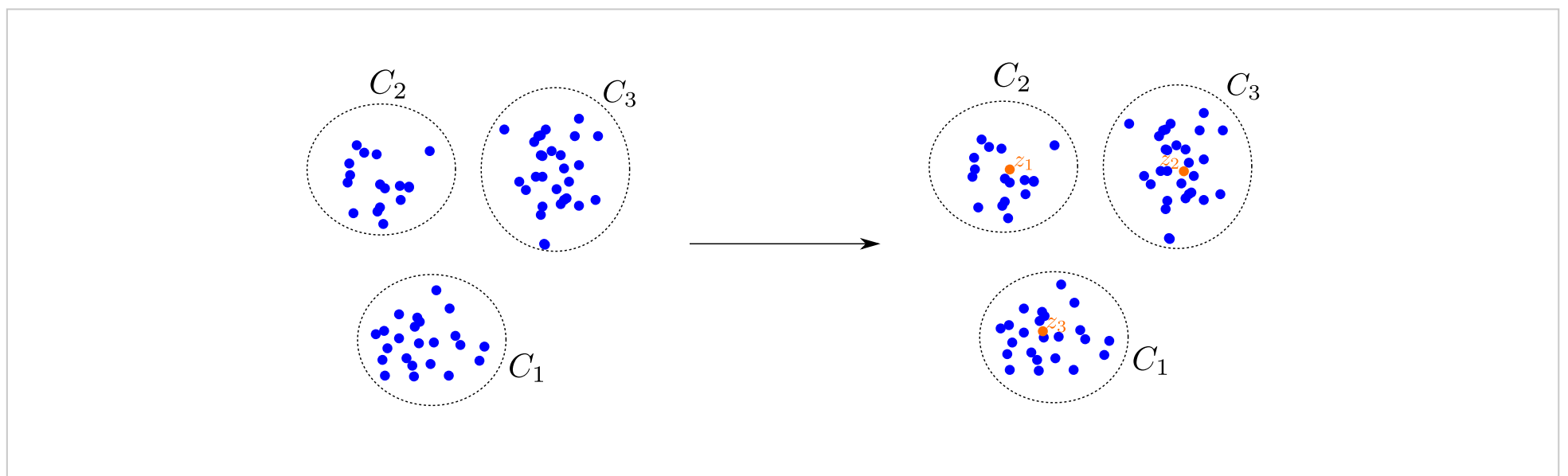
Which is it?

☐ Step 1

☒ Step 2.1 ✓

☐ Step 2.2

The following figure depicts an example of one of the steps of K-means algorithm:



Which step is it?

☐ Step 2.1

☒ Step 2.2 ✓

Solution:

Step 2.1 assigns each points to the best cluster, while step 2.2 selects out the representative of each cluster. Note that step 1 is random initialization of cluster assignments.

Submit

You have used 1 of 3 attempts