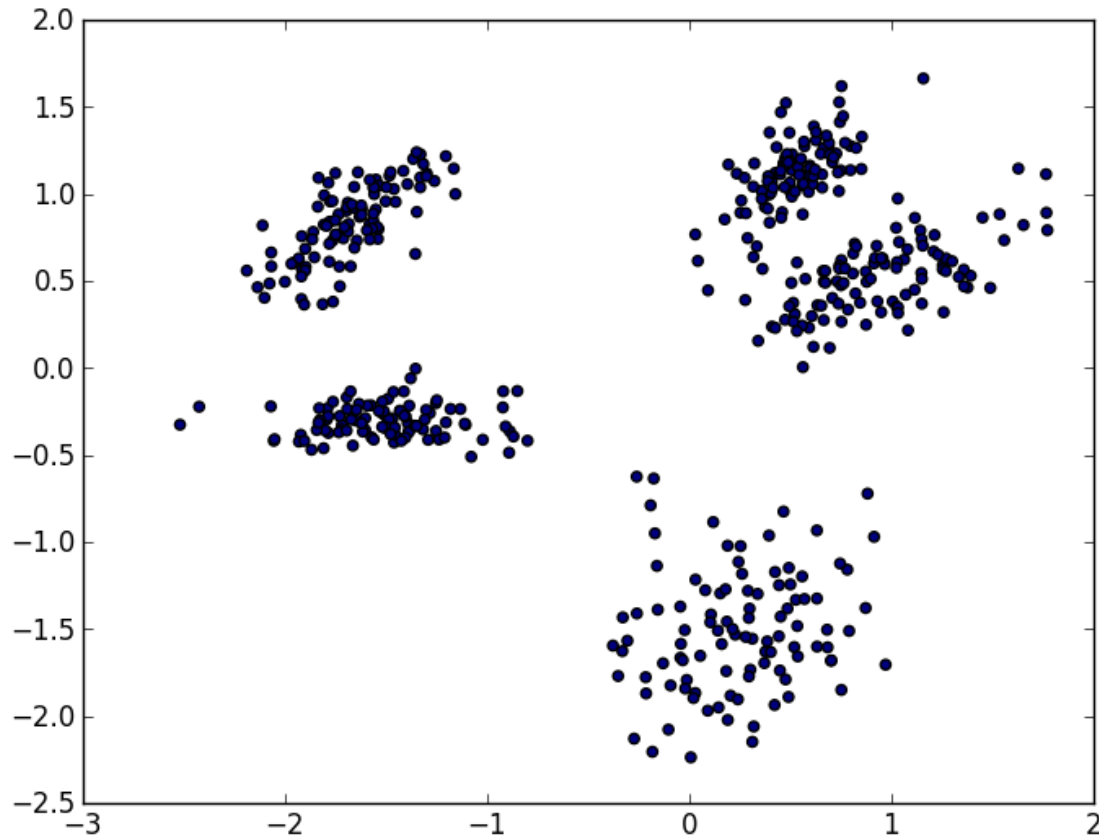


Lab Course

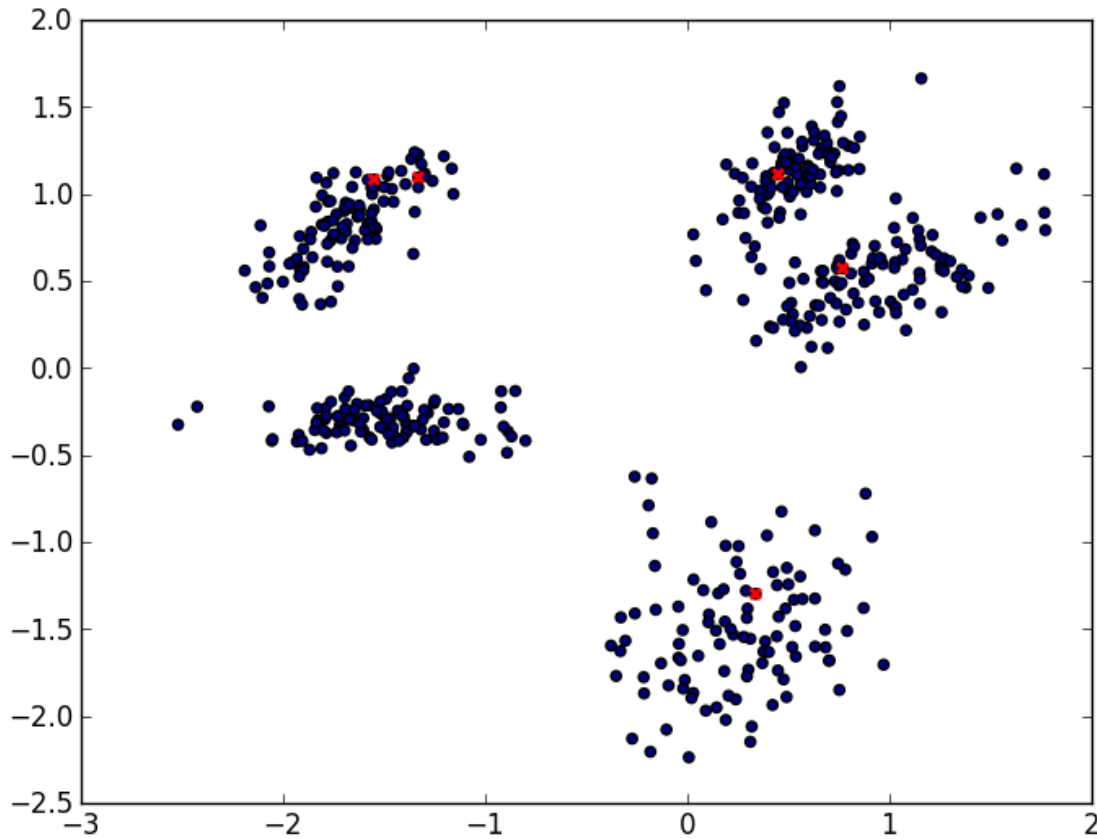
Clustering

K-Means



Goal: find clusters

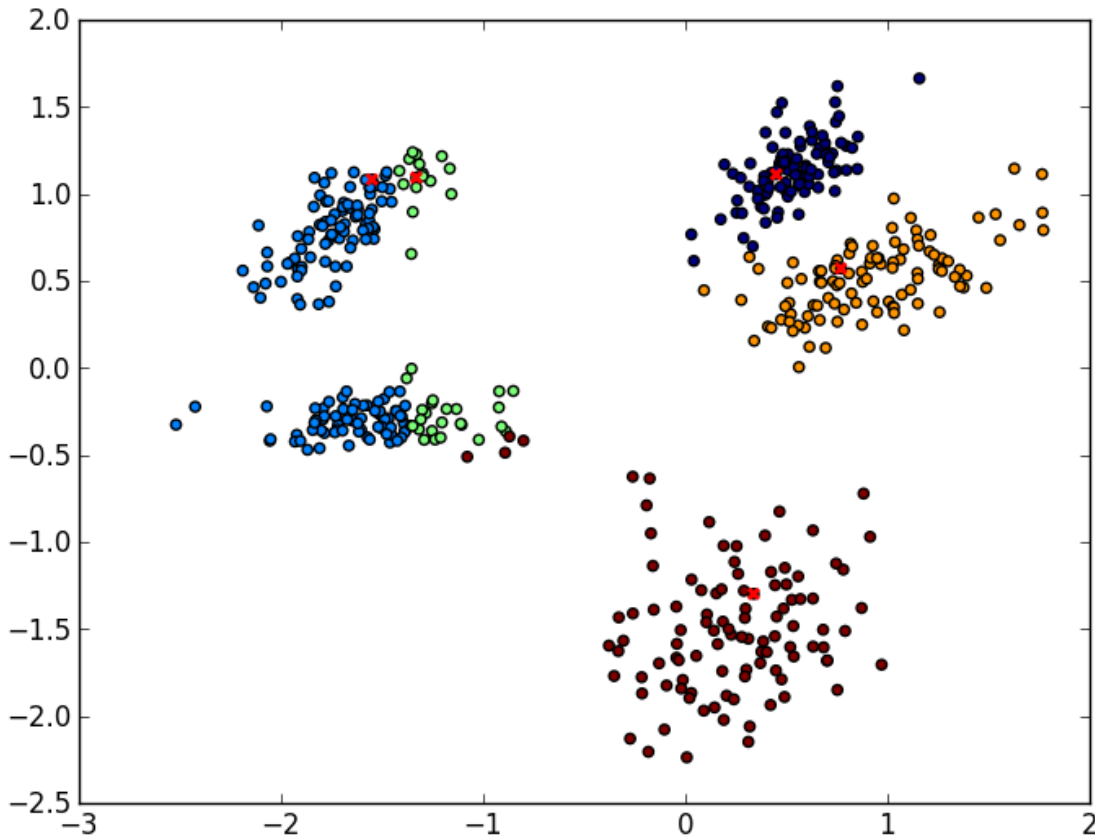
K-Means



Goal: find clusters

Start with random assignment
of cluster

K-Means



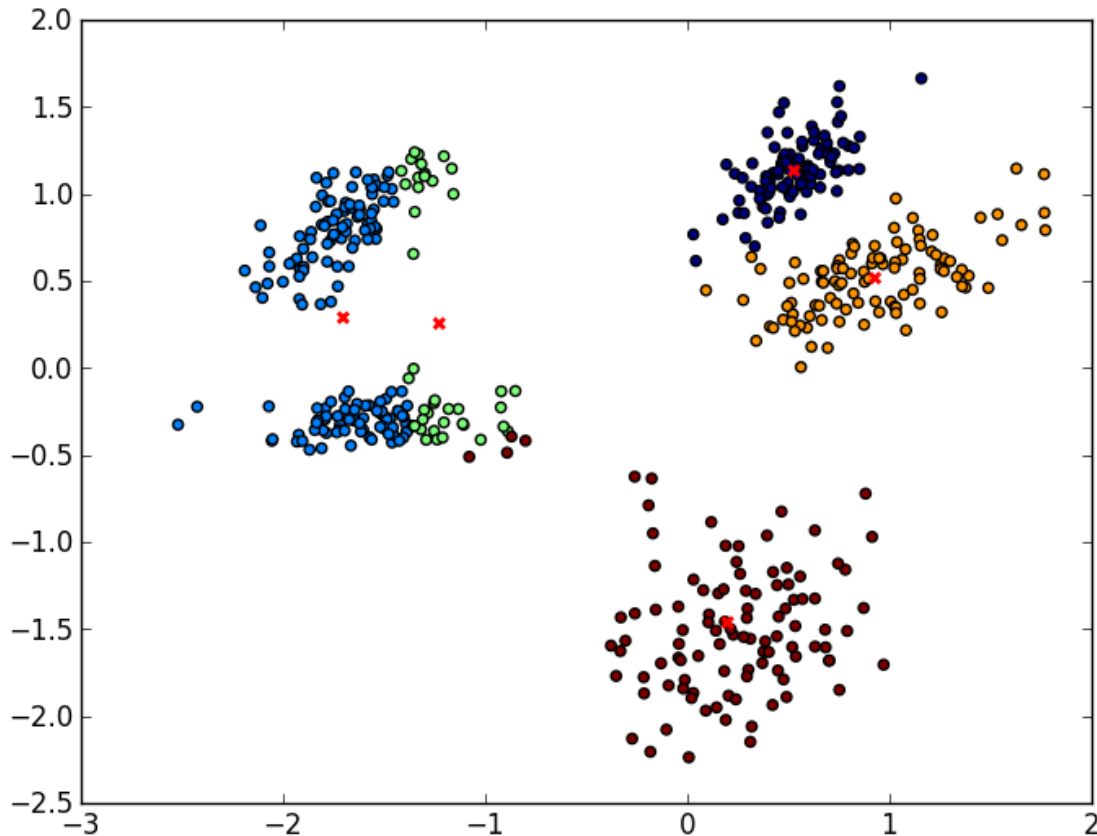
Goal: find clusters

Start with random assignment
of cluster

Till convergence:

Step 1: Reassign data points

K-Means



Goal: find clusters

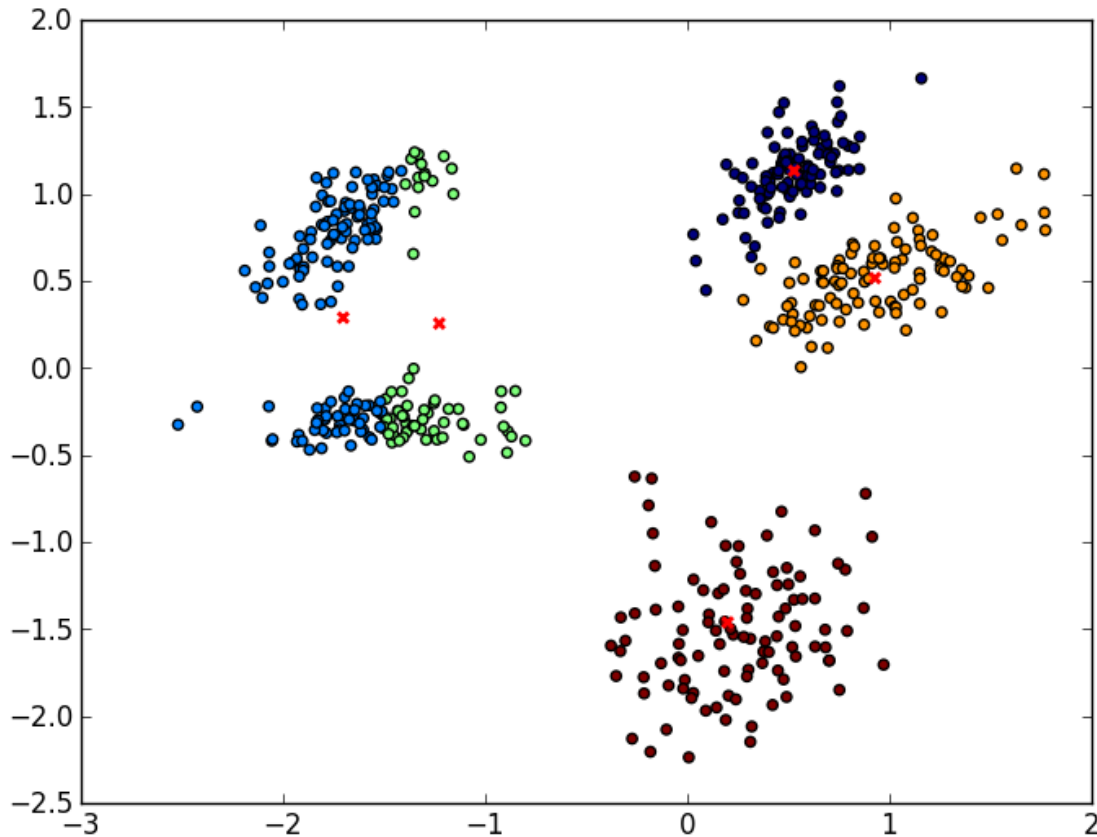
Start with random assignment
of cluster

Till convergence:

Step 1: Reassign data points

Step 2: Compute new cluster
centers

K-Means



Goal: find clusters

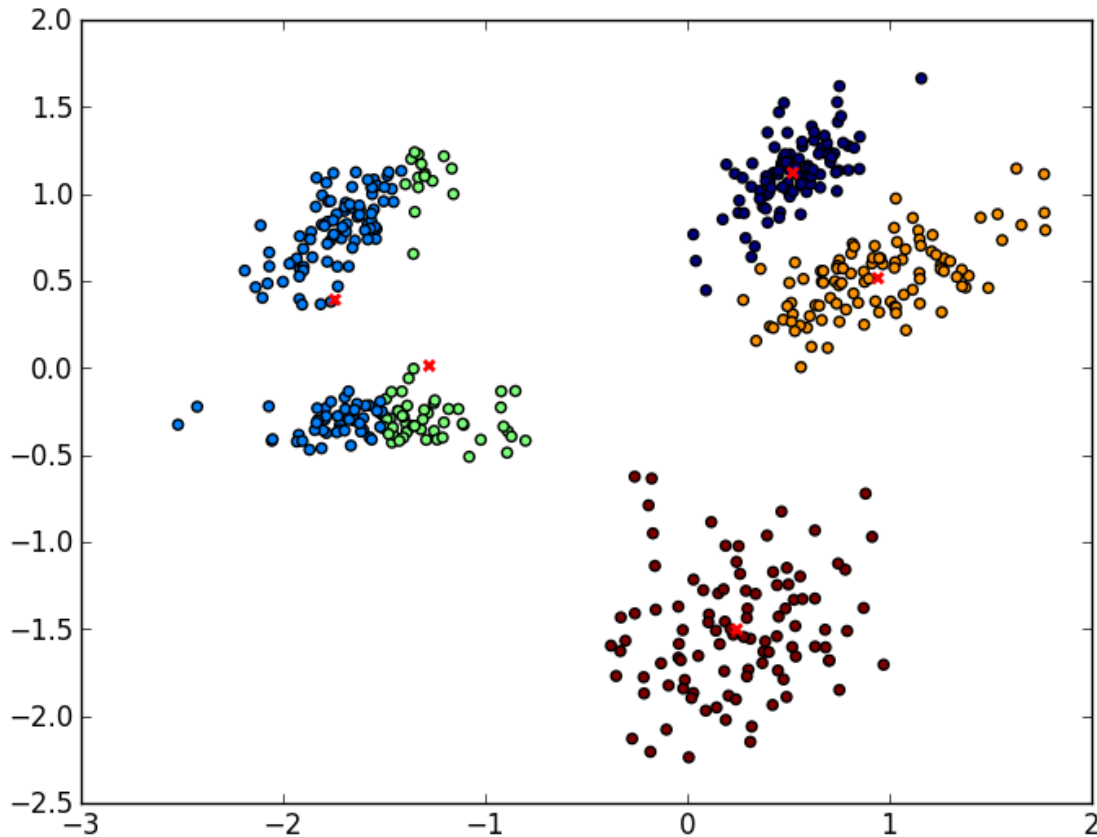
Start with random assignment
of cluster

Till convergence:

Step 1: Reassign data points

Step 2: Compute new cluster
centers

K-Means



Goal: find clusters

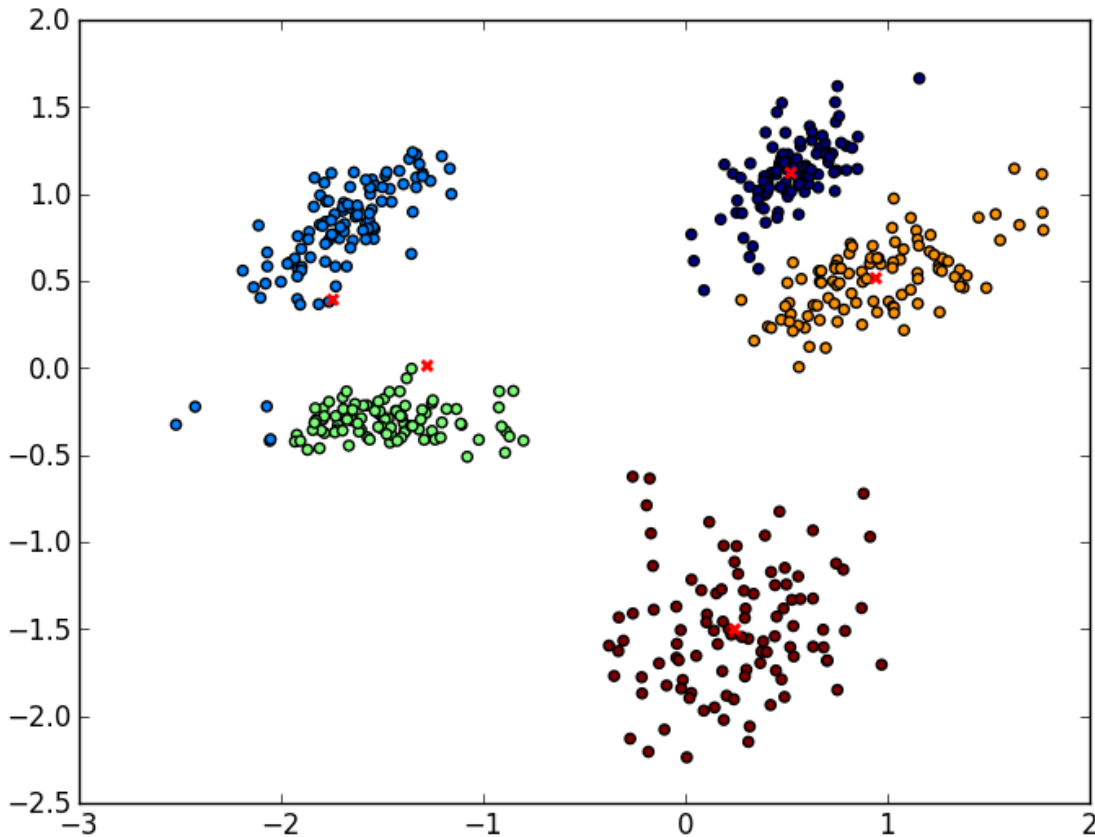
Start with random assignment
of cluster

Till convergence:

Step 1: Reassign data points

Step 2: Compute new cluster
centers

K-Means



Goal: find clusters

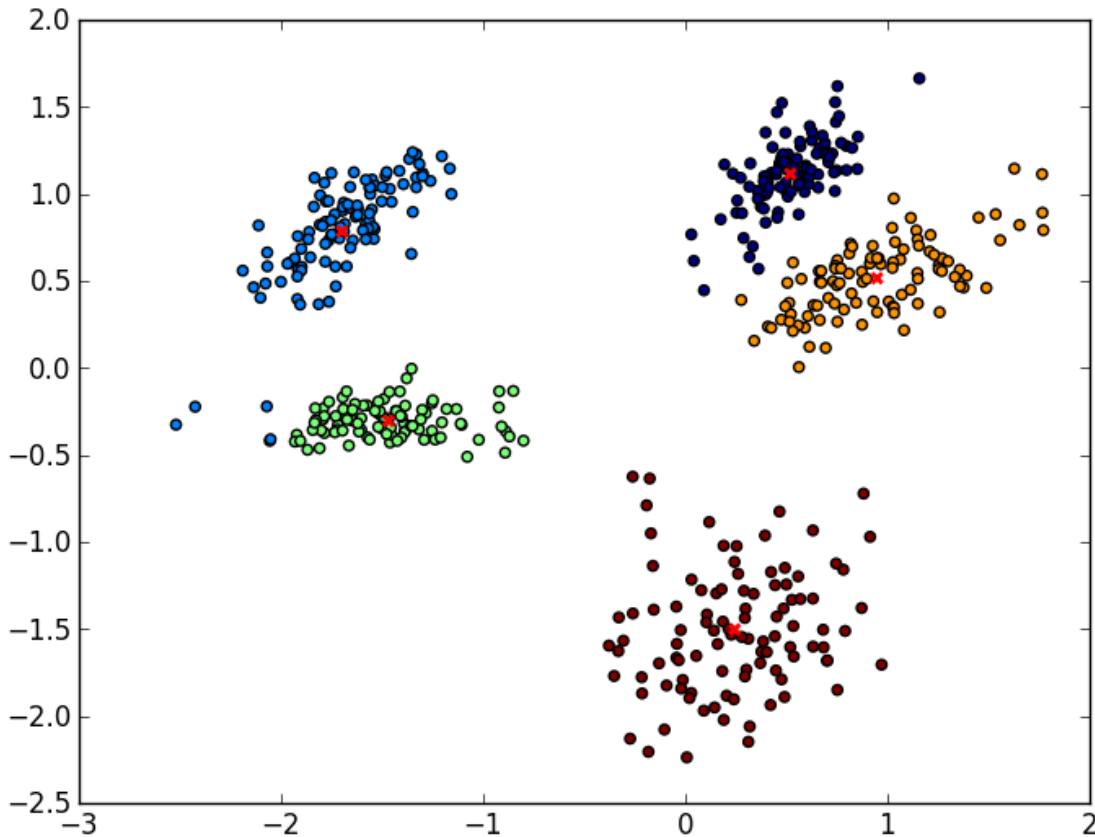
Start with random assignment
of cluster

Till convergence:

Step 1: Reassign data points

Step 2: Compute new cluster
centers

K-Means



Goal: find clusters

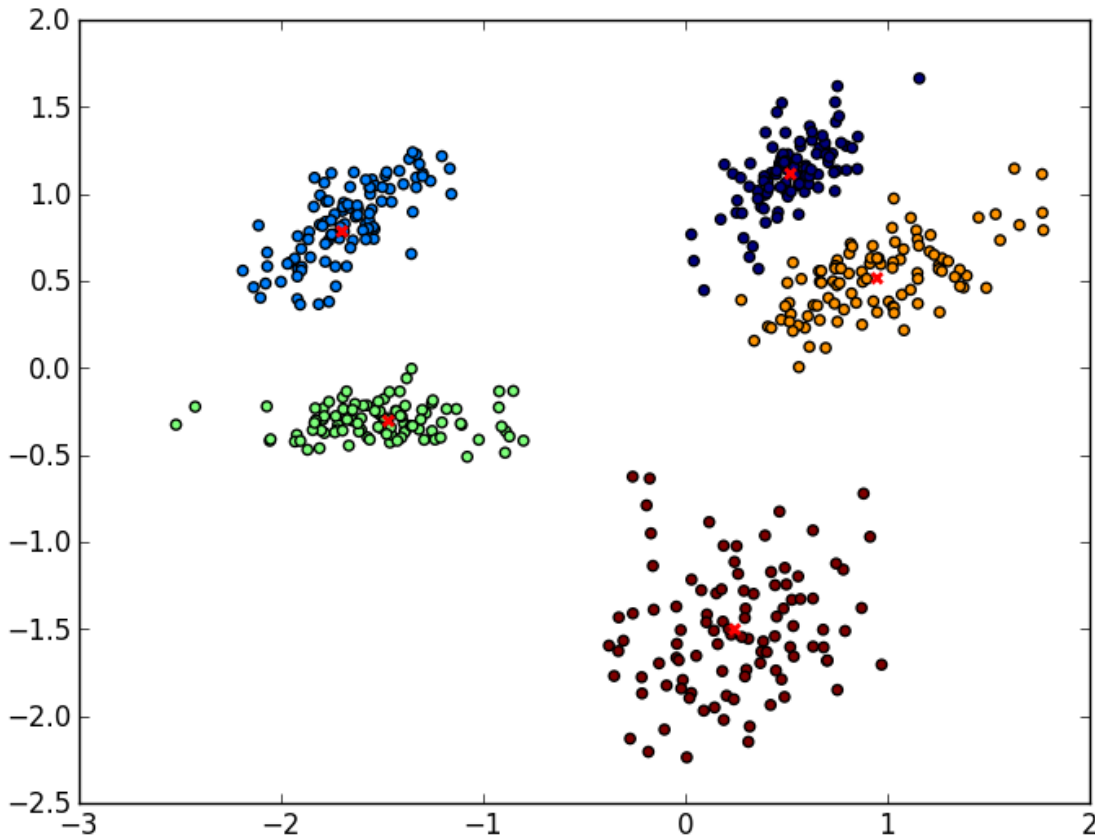
Start with random assignment
of cluster

Till convergence:

Step 1: Reassign data points

Step 2: Compute new cluster
centers

K-Means



Goal: find clusters

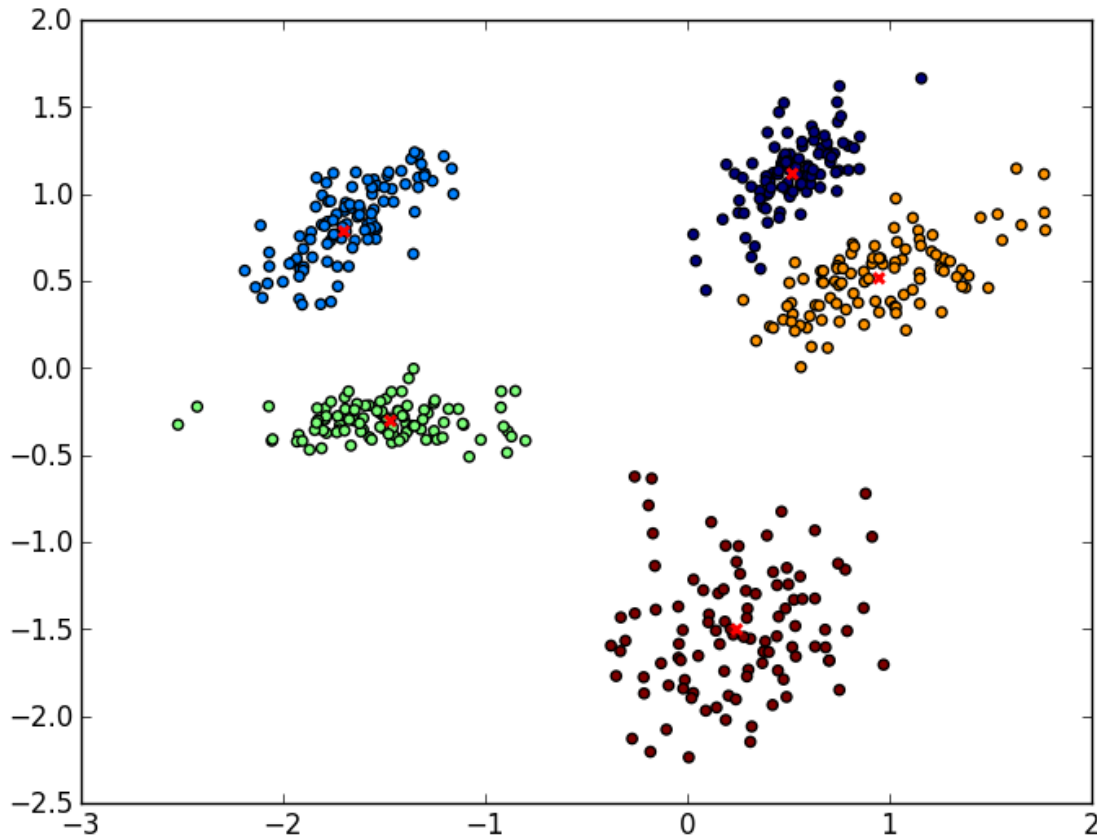
Start with random assignment
of cluster

Till convergence:

Step 1: Reassign data points

Step 2: Compute new cluster
centers

Hierarchical Clustering

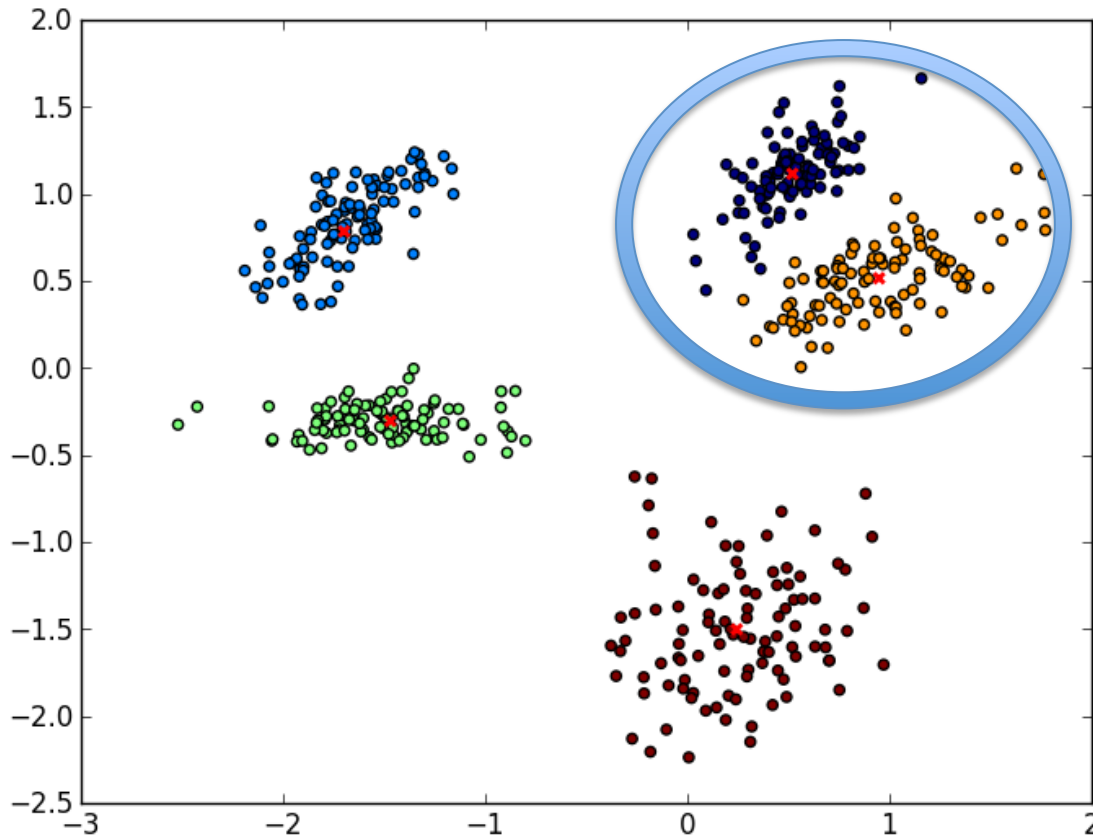


Continuously merge clusters

Loss function

E.g. K-Means criterion
(Average distance of points to
their cluster center)

Hierarchical Clustering

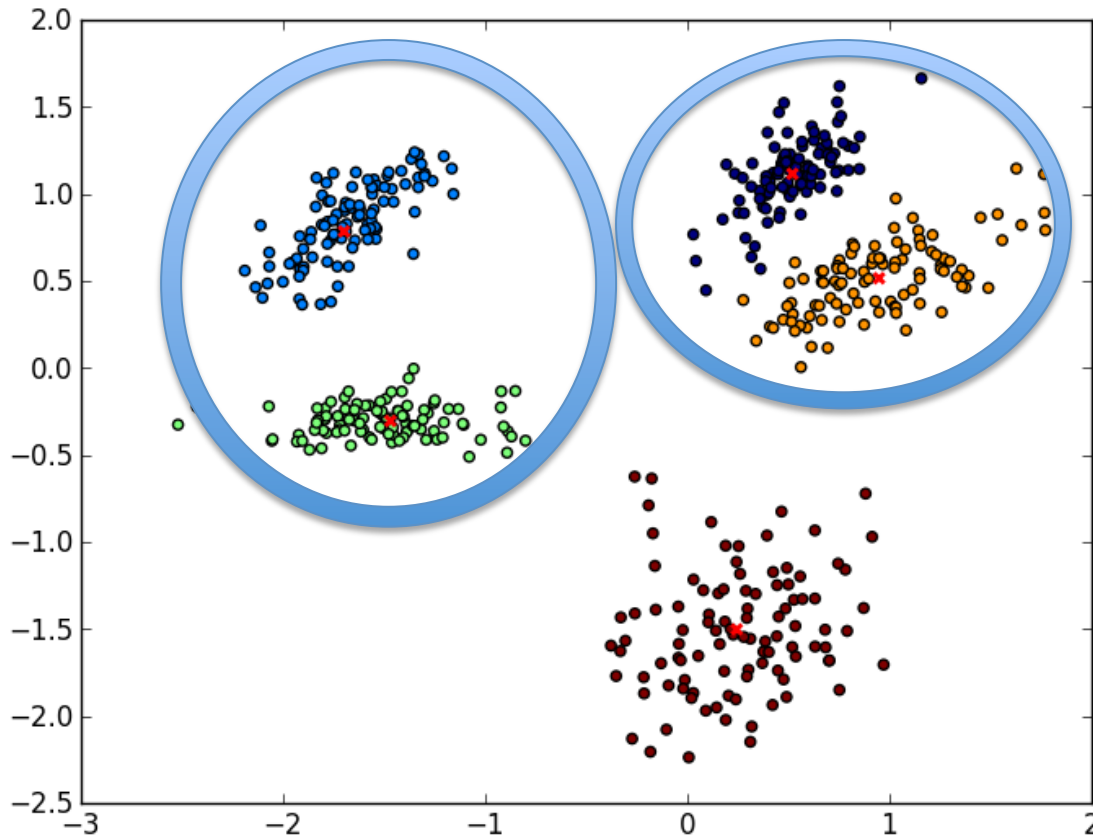


Continuously merge clusters

Loss function

E.g. K-Means criterion
(Average distance of points to
their cluster center)

Hierarchical Clustering

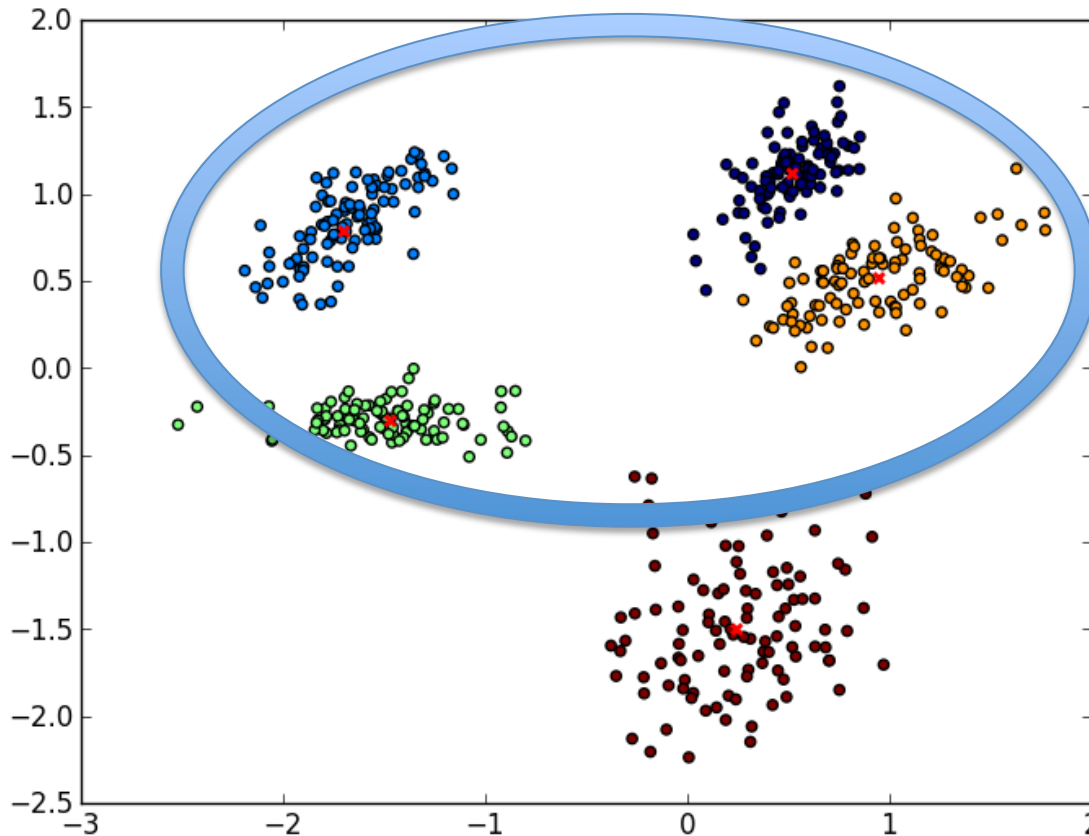


Continuously merge clusters

Loss function

E.g. K-Means criterion
(Average distance of points to
their cluster center)

Hierarchical Clustering

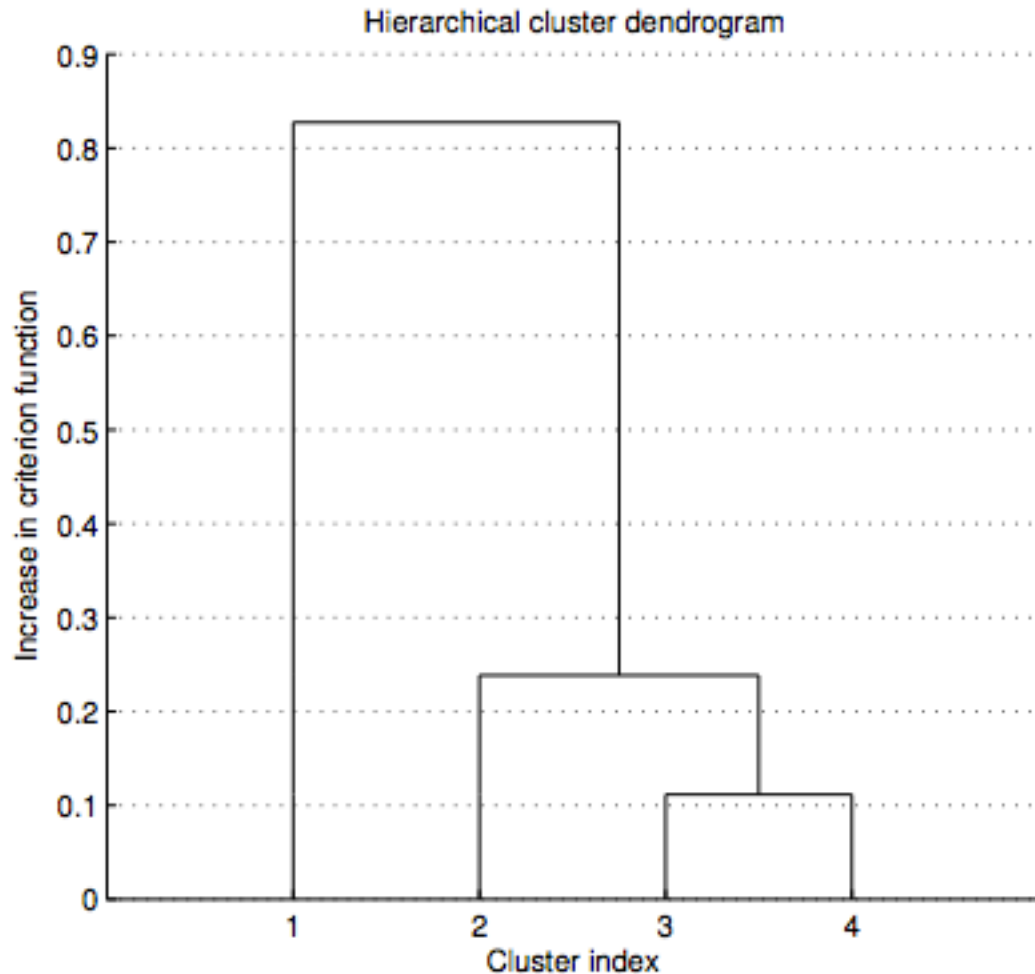


Continuously merge clusters

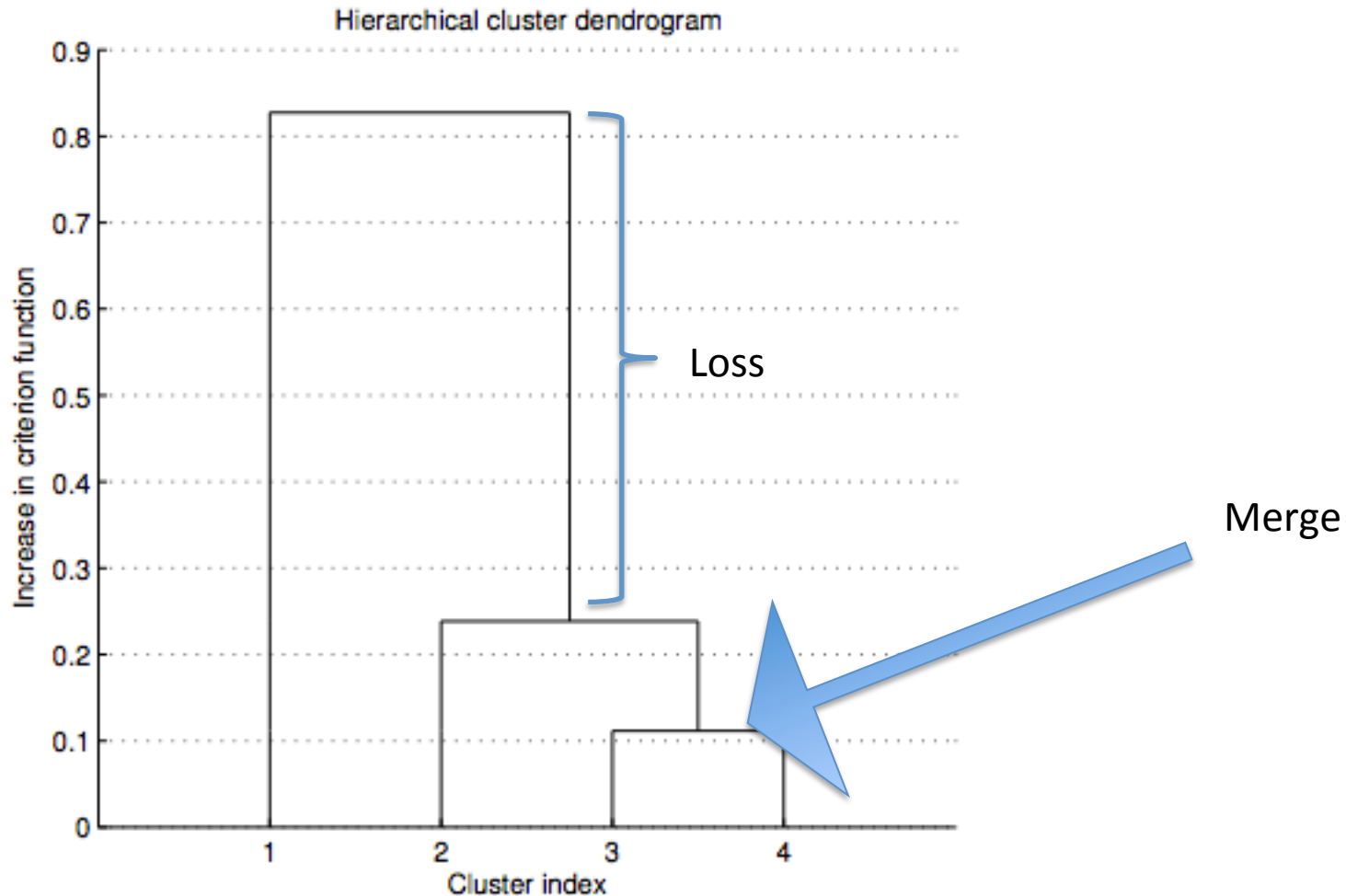
Loss function

E.g. K-Means criterion
(Average distance of points to
their cluster center)

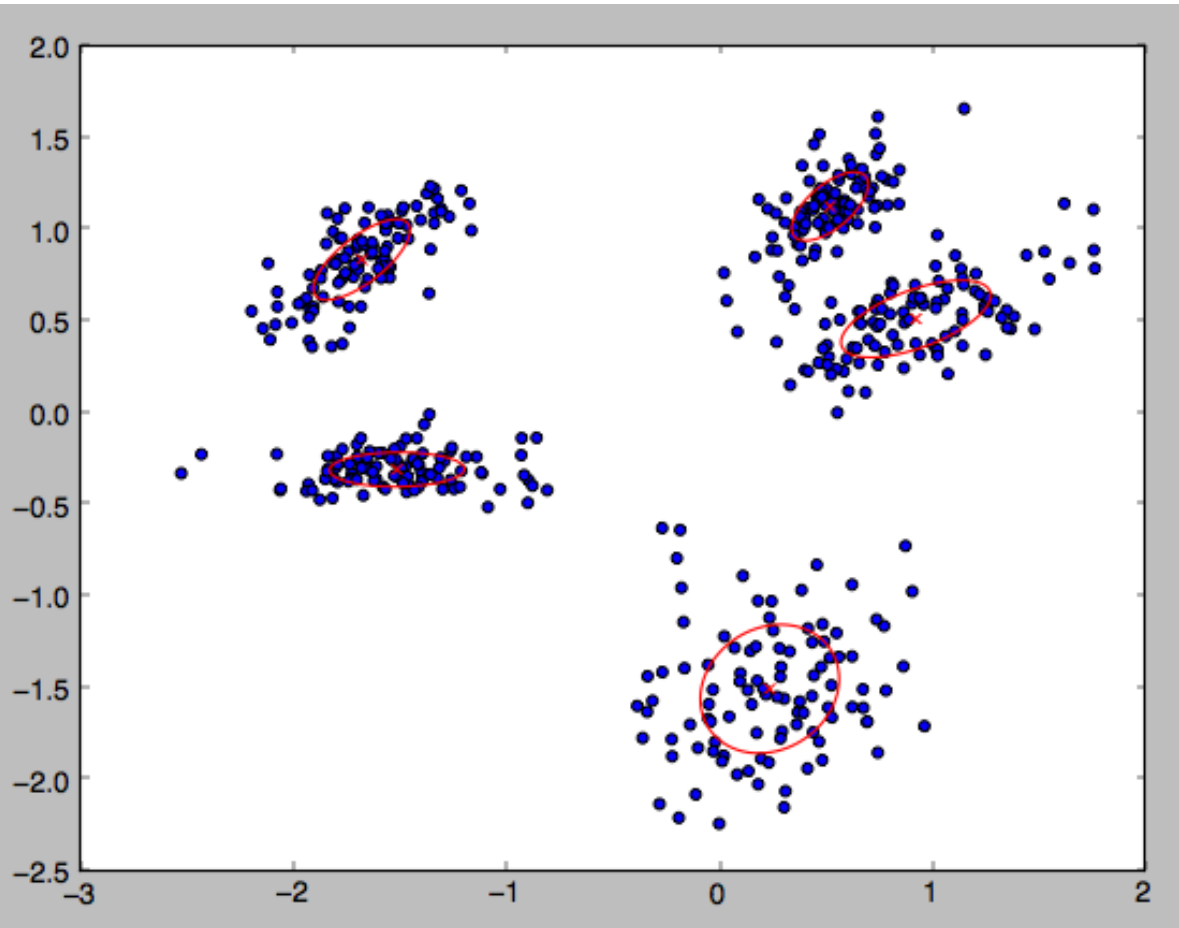
Visualization: Dendrogram plot



Visualization: Dendrogram plot



EM for Mixture of Gaussians



View clusters as mixtures of Gaussians

Consider not only cluster centers, but also Covariance matrices

EM for Mixture of Gaussians

Algorithm

$\hat{\pi}_k \leftarrow 1/K$ Prior distribution of cluster assignments

$\hat{\mu}_k \leftarrow$ random points out of X_1, \dots, X_n

$\hat{\Sigma}_k \leftarrow \mathbf{I}_d$

Step 1 (E-Step)

for $k \leftarrow 1$ to K **do**

for $n \leftarrow 1$ to N **do**

 Set $\gamma_{nk} \leftarrow \frac{\hat{\pi}_k g(X_n; \hat{\mu}_k, \hat{\Sigma}_k)}{\sum_{k'=1}^K \hat{\pi}_{k'} g(X_n; \hat{\mu}_{k'}, \hat{\Sigma}_{k'})}$

end for

end for

Compute likelihood that point n belongs to cluster k given the cluster centers and covariance matrices

g is the Gaussian probability density function

Step 2 (M-Step)

for $k \leftarrow 1$ to K **do**

$N_k \leftarrow \sum_{n=1}^N \gamma_{nk}$

$\hat{\pi}_k \leftarrow N_k / N$

$\hat{\mu}_k \leftarrow \frac{1}{N_k} \sum_{n=1}^N \gamma_{nk} X_n$

$\hat{\Sigma}_k \leftarrow \frac{1}{N_k} \sum_{n=1}^N \gamma_{nk} (X_n - \hat{\mu}_k)(X_n - \hat{\mu}_k)^\top$

end for

Computer new cluster centers + covariance matrices + priors