Clustering with variance minimization (k=2)

Initialization
Clustering with variance minimization (k=2)

Calculate centroids

\[ \mu = \left(\frac{4.3}{6.0}\right) \]

\[ \mu = \left(\frac{6.4}{5.0}\right) \]
Clustering with variance minimization (k=2)

Reassign points
Clustering with variance minimization (k=2)

Calculate centroids

\[ \mu = \left( \frac{2.7}{5.0} \right) \]

\[ \mu = \left( \frac{7.4}{5.6} \right) \]
Clustering with variance minimization (k=2)

Reassign points
Clustering with variance minimization (k=2)

Calculate centroids

\[ \mu = \left( \frac{3.25}{4.0} \right) \]

\[ \mu = \left( \frac{8.0}{6.75} \right) \]
Clustering with variance minimization (k=2)

Reassign points
Clustering with variance minimization (k=2)
Clustering with variance minimization (k=2): alternative initial clustering

Initialization

Calculate centroids

\[ \mu = \frac{4}{8.5} \]

\[ \mu = \frac{6.17}{4.33} \]
Clustering with variance minimization (k=2): alternative initial clustering

Reassign points
Clustering with variance minimization (k=2): alternative initial clustering

Calculate centroids

\[ \mu = \left( \frac{1}{10} \right) \]

\[ \mu = \left( \frac{6.3}{4.7} \right) \]
Clustering with variance minimization (k=2): alternative initial clustering

Reassign points
Clustering with variance minimization (k=2): alternative initial clustering
Costs?

Obtained clustering

\[
\mu = \frac{3.25}{4.0}
\]

\[
\mu = \frac{8.0}{6.75}
\]

\[
\text{Dist}(\mu, 1) = |8.0 - 7.0| + |6.75 - 7.0| = 1 + 0.25 = 1.25
\]

\[
\text{Dist}(\mu, 2) = |8.0 - 8.0| + |6.75 - 7.0| = 0.25
\]

\[
\text{Dist}(\mu, 4) = |8.0 - 9.0| + |6.75 - 7.0| = 1.25
\]

\[
\text{Dist}(\mu, 3) = |8.0 - 8.0| + |6.75 - 6.0| = 0.75
\]

\[
\text{Dist}(\mu, 5) = |3.25 - 5.0| + |4.0 - 1.0| = 1.75 + 3 = 4.75
\]

\[
\text{Dist}(\mu, 6) = |3.25 - 3.0| + |4.0 - 2.0| = 0.25 + 2 = 2.25
\]

\[
\text{Dist}(\mu, 7) = |3.25 - 4.0| + |4.0 - 3.0| = 0.75 + 1 = 1.75
\]

\[
\text{Dist}(\mu, 8) = |3.25 - 1.0| + |4.0 - 10.0| = 2.75 + 6 = 8.75
\]

\[
\text{TD}^2(C1) = 3.75
\]

\[
\text{TD}^2(C2) = 107.25
\]

\[
\text{TD}^2 = 107.25 + 3.75 = 111
\]
Costs?

Alternative clustering

\[ \mu = (1.0 / 10.0) \]

\[ \mu = (6.3 / 4.7) \]

\[
\begin{align*}
\text{Dist}(\mu, 1) &= |6.3 - 7.0| + |4.7 - 7.0| = 0.7 + 2.3 = 3.0 \\
\text{Dist}(\mu, 2) &= |6.3 - 8.0| + |4.7 - 7.0| = 1.7 + 2.3 = 4.0 \\
\text{Dist}(\mu, 3) &= |6.3 - 9.0| + |4.7 - 7.0| = 2.7 + 2.3 = 5.0 \\
\text{Dist}(\mu, 4) &= |6.3 - 8.0| + |4.7 - 6.0| = 1.7 + 1.3 = 3.0 \\
\text{Dist}(\mu, 5) &= |6.3 - 5.0| + |4.7 - 1.0| = 1.3 + 3.7 = 5.0 \\
\text{Dist}(\mu, 6) &= |6.3 - 3.0| + |4.7 - 2.0| = 3.3 + 2.7 = 6.0 \\
\text{Dist}(\mu, 7) &= |6.3 - 4.0| + |4.7 - 3.0| = 2.3 + 1.7 = 4.0 \\
\end{align*}
\]

\[ \text{TD}^2(C1) = 136 \]

\[ \text{Dist}(\mu, 8) = |1.0 - 1.0| + |10.0 - 10.0| = 0 \]

\[ \text{TD}^2(C2) = 0 \]

\[ \text{TD}^2 = 136 \]
Costs?

Optimal Clustering

Dist(μ,1) = |6.6 – 7.0| + |7.4 – 7.0| = 0.4 + 0.4 = 0.8
Dist(μ,2) = |6.6 – 8.0| + |7.4 – 7.0| = 1.4 + 0.4 = 1.8
Dist(μ,3) = |6.6 – 8.0| + |7.4 – 6.0| = 1.4 + 1.4 = 2.8
Dist(μ,4) = |6.6 – 9.0| + |7.4 – 7.0| = 2.4 + 0.4 = 2.8
Dist(μ,8) = |6.6 – 1.0| + |7.4 – 10.0| = 5.6 + 2.6 = 8.2
TD²(C1) = 86.8

Dist(μ,5) = |4.0 – 5.0| + |2.0 – 1.0| = 2
Dist(μ,6) = |4.0 – 3.0| + |2.0 – 2.0| = 1
Dist(μ,7) = |4.0 – 4.0| + |2.0 – 3.0| = 1
TD²(C2) = 6
TD² = 86.8 + 6.0 = 92.8

µ = (4.0 / 2.0)
µ = (6.6 / 7.4)
Conclusion

• Variant k-Means: k-Means already recalculates the centroids after reassigning one node only, instead of recalculating after regarding all nodes.
• Note, that clustering with variance minimization...:
  – ...mostly converges only to a local minimum
  – ...depends on the order of the assignment and on the initial partitioning
  – ...is prone to noise

Nevertheless, clustering with variance minimization is at the moment the most popular clustering algorithm in industry and in large parts of science