Aufgabe 17: kMeans per Hand

Population: $\vec{x}_1 = \binom{1}{4}$, $\vec{x}_2 = \binom{1}{5}$, $\vec{x}_3 = \binom{1}{6}$, $\vec{x}_4 = \binom{3}{3}$, $\vec{x}_5 = \binom{3}{2}$

$\vec{x}_6 = \binom{4}{1}$, $\vec{x}_7 = \binom{5}{1}$, $\vec{x}_8 = \binom{6}{2}$, $\vec{x}_9 = \binom{6}{3}$, $\vec{x}_{10} = \binom{8}{4}$

$\vec{x}_{11} = \binom{8}{5}$, $\vec{x}_{12} = \binom{8}{6}$

a) Start-Clusterzentren: $S_1^{(0)} = \binom{3}{7}$, $S_2^{(0)} = \binom{3}{4}$, $S_3^{(0)} = \binom{7}{4}$

Punkte zuordnen:

Cluster 1: $\{\vec{x}_3\}$

Cluster 2: $\{\vec{x}_1, \vec{x}_2, \vec{x}_4, \vec{x}_5, \vec{x}_6\}$

Cluster 3: $\{\vec{x}_7, \vec{x}_8, \vec{x}_9, \vec{x}_{10}, \vec{x}_{11}, \vec{x}_{12}\}$

Alle eindeutig zuordenbar bis auf $\vec{x}_7$:

$\|\vec{x}_7 - S_2^{(0)}\|^2 = \|\binom{5}{1} - \binom{3}{4}\|^2 = 2^2 + (-3)^2 = 13$

$\|\vec{x}_7 - S_3^{(0)}\|^2 = \|\binom{5}{1} - \binom{7}{4}\|^2 = (-2)^2 + 3^2 = 13$

} identisch → suche einen Cluster aus ⇒ Cluster 3

1. Iteration:

Neue Clusterzentren berechnen

$S_1^{(1)} = \binom{1}{6}$

$S_2^{(1)} = \frac{1}{5}\left[\binom{1}{4} + \binom{1}{5} + \binom{3}{3} + \binom{3}{2} + \binom{4}{1}\right] = \frac{1}{5}\binom{12}{15} = \binom{2{,}4}{3}$

$S_3^{(1)} = \frac{1}{6}\left[\binom{5}{1} + \binom{6}{2} + \binom{6}{3} + \binom{8}{4} + \binom{8}{5} + \binom{8}{6}\right] = \frac{1}{6}\binom{41}{21} \approx \binom{6{,}833}{3{,}5}$

Punkte neu zuordnen:

Cluster 1: $\{\vec{x}_2, \vec{x}_3\}$

Cluster 2: $\{\vec{x}_1, \vec{x}_4, \vec{x}_5, \vec{x}_6\}$

Cluster 3: $\{\vec{x}_7, \vec{x}_8, \vec{x}_9, \vec{x}_{10}, \vec{x}_{11}, \vec{x}_{12}\}$

$\vec{x}_7$ nicht offensichtlich zuordenbar:

$\|\vec{x}_7 - S_2^{(1)}\|^2 = \|\binom{5}{1} - \binom{2{,}4}{3}\|^2 = 2{,}6^2 + (-2)^2 = 10{,}76$

$\|\vec{x}_7 - S_3^{(1)}\|^2 \approx \|\binom{5}{1} - \binom{6{,}833}{3{,}5}\|^2 \approx 3{,}361 + 6{,}25 = 9{,}611$

} $\vec{x}_7$ gehört zu Cluster 3

b) 2. Iteration:

$S_1^{(2)} = \frac{1}{2}\left[\binom{1}{5} + \binom{1}{6}\right] = \binom{1}{5{,}5}$

$S_2^{(2)} = \frac{1}{4}\left[\binom{1}{4} + \binom{3}{3} + \binom{3}{2} + \binom{4}{1}\right] = \binom{2{,}75}{2{,}5}$

$S_3^{(2)} = S_3^{(1)} \approx \binom{6{,}833}{3{,}5}$

$\vec{x}_7$ nicht eindeutig zuordenbar:

$\|\vec{x}_7 - S_2^{(2)}\|^2 = \|\binom{5}{1} - \binom{2{,}75}{2{,}5}\|^2 = 2{,}25^2 + (-1{,}5)^2 = 7{,}3125$

$\|\vec{x}_7 - S_3^{(2)}\|^2 = 9{,}611$ (s.ob.)

} ⇒ $\vec{x}_7$ gehört zu Cluster 2

Cluster 1: $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$
Cluster 2: $\{\vec{x}_4, \vec{x}_5, \vec{x}_6, \vec{x}_7\}$
Cluster 3: $\{\vec{x}_8, \vec{x}_9, \vec{x}_{10}, \vec{x}_{11}, \vec{x}_{12}\}$

<u>3. Iteration:</u>
Neue Clusterzentren berechnen

$$S_1^{(3)} = \frac{1}{3}\left[\begin{pmatrix}1\\6\end{pmatrix}+\begin{pmatrix}1\\5\end{pmatrix}+\begin{pmatrix}1\\4\end{pmatrix}\right] = \begin{pmatrix}1\\5\end{pmatrix}$$

$$S_2^{(3)} = \frac{1}{4}\left[\begin{pmatrix}3\\3\end{pmatrix}+\begin{pmatrix}3\\2\end{pmatrix}+\begin{pmatrix}4\\1\end{pmatrix}+\begin{pmatrix}5\\1\end{pmatrix}\right] = \begin{pmatrix}3{,}75\\1{,}75\end{pmatrix}$$

$$S_3^{(3)} = \frac{1}{5}\left[\begin{pmatrix}6\\2\end{pmatrix}+\begin{pmatrix}6\\3\end{pmatrix}+\begin{pmatrix}8\\4\end{pmatrix}+\begin{pmatrix}8\\5\end{pmatrix}+\begin{pmatrix}8\\6\end{pmatrix}\right] = \begin{pmatrix}7{,}2\\4\end{pmatrix}$$

$\vec{x}_8$ nicht eindeutig zuordenbar

$$\left.\begin{aligned}\|\vec{x}_8 - S_2^{(3)}\|^2 &= \left\|\begin{pmatrix}6\\2\end{pmatrix}-\begin{pmatrix}3{,}75\\1{,}75\end{pmatrix}\right\|^2 = 2{,}25 + (-0{,}25)^2 = 5{,}125\\ \|\vec{x}_8 - S_3^{(3)}\|^2 &= \left\|\begin{pmatrix}6\\2\end{pmatrix}-\begin{pmatrix}7{,}2\\4\end{pmatrix}\right\|^2 = (-1{,}2)^2 + (-2)^2 = 5{,}44\end{aligned}\right\} \rightarrow \vec{x}_8 \text{ gehört zu Cluster 2}$$

Cluster 1: $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$
Cluster 2: $\{\vec{x}_4, \vec{x}_5, \vec{x}_6, \vec{x}_7, \vec{x}_8\}$
Cluster 3: $\{\vec{x}_9, \vec{x}_{10}, \vec{x}_{11}, \vec{x}_{12}\}$

<u>4. Iteration:</u>
Neue Clusterzentren berechnen

$$S_1^{(4)} = S_1^{(3)} = \begin{pmatrix}1\\5\end{pmatrix}$$

$$S_2^{(4)} = \frac{1}{5}\left[\begin{pmatrix}3\\3\end{pmatrix}+\begin{pmatrix}3\\2\end{pmatrix}+\begin{pmatrix}4\\1\end{pmatrix}+\begin{pmatrix}5\\1\end{pmatrix}+\begin{pmatrix}6\\2\end{pmatrix}\right] = \begin{pmatrix}4{,}2\\1{,}8\end{pmatrix}$$

$$S_3^{(4)} = \frac{1}{4}\left[\begin{pmatrix}6\\3\end{pmatrix}+\begin{pmatrix}8\\4\end{pmatrix}+\begin{pmatrix}8\\5\end{pmatrix}+\begin{pmatrix}8\\6\end{pmatrix}\right] = \begin{pmatrix}7{,}5\\4{,}5\end{pmatrix}$$

$\vec{x}_9$ nicht eindeutig zuordenbar:

$$\left.\begin{aligned}\|\vec{x}_9 - S_2^{(4)}\|^2 &= \left\|\begin{pmatrix}6\\3\end{pmatrix}-\begin{pmatrix}4{,}2\\1{,}8\end{pmatrix}\right\|^2 = 1{,}8^2 + 1{,}2^2 = 4{,}68\\ \|\vec{x}_9 - S_3^{(4)}\|^2 &= \left\|\begin{pmatrix}6\\3\end{pmatrix}-\begin{pmatrix}7{,}5\\4{,}5\end{pmatrix}\right\|^2 = (-1{,}5)^2 + (-1{,}5)^2 = 4{,}5\end{aligned}\right\} \rightarrow \vec{x}_9 \text{ gehört zu Cluster 3}$$

Cluster 1: $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$
Cluster 2: $\{\vec{x}_4, \vec{x}_5, \vec{x}_6, \vec{x}_7, \vec{x}_8\}$
Cluster 3: $\{\vec{x}_9, \vec{x}_{10}, \vec{x}_{11}, \vec{x}_{12}\}$

<u>5. Iteration</u>
→ keine Änderung mehr $S_1^{(5)} = S_1^{(4)}$, $S_2^{(5)} = S_2^{(4)}$, $S_3^{(5)} = S_3^{(4)}$

c) Der Algorithmus konvergiert in diesem Fall nach der 4. Iteration.
Das Ergebnis entspricht nicht unseren Erwartungen. Unter der Annahme, dass es 3 Cluster gibt, hätten wir folgendes erwartet:

Cluster 1: $\{\vec{x}_1, \vec{x}_2, \vec{x}_3\}$, Cluster 2: $\{\vec{x}_4, \vec{x}_5, \vec{x}_6, \vec{x}_7, \vec{x}_8, \vec{x}_9\}$, Cluster 3: $\{\vec{x}_{10}, \vec{x}_{11}, \vec{x}_{12}\}$

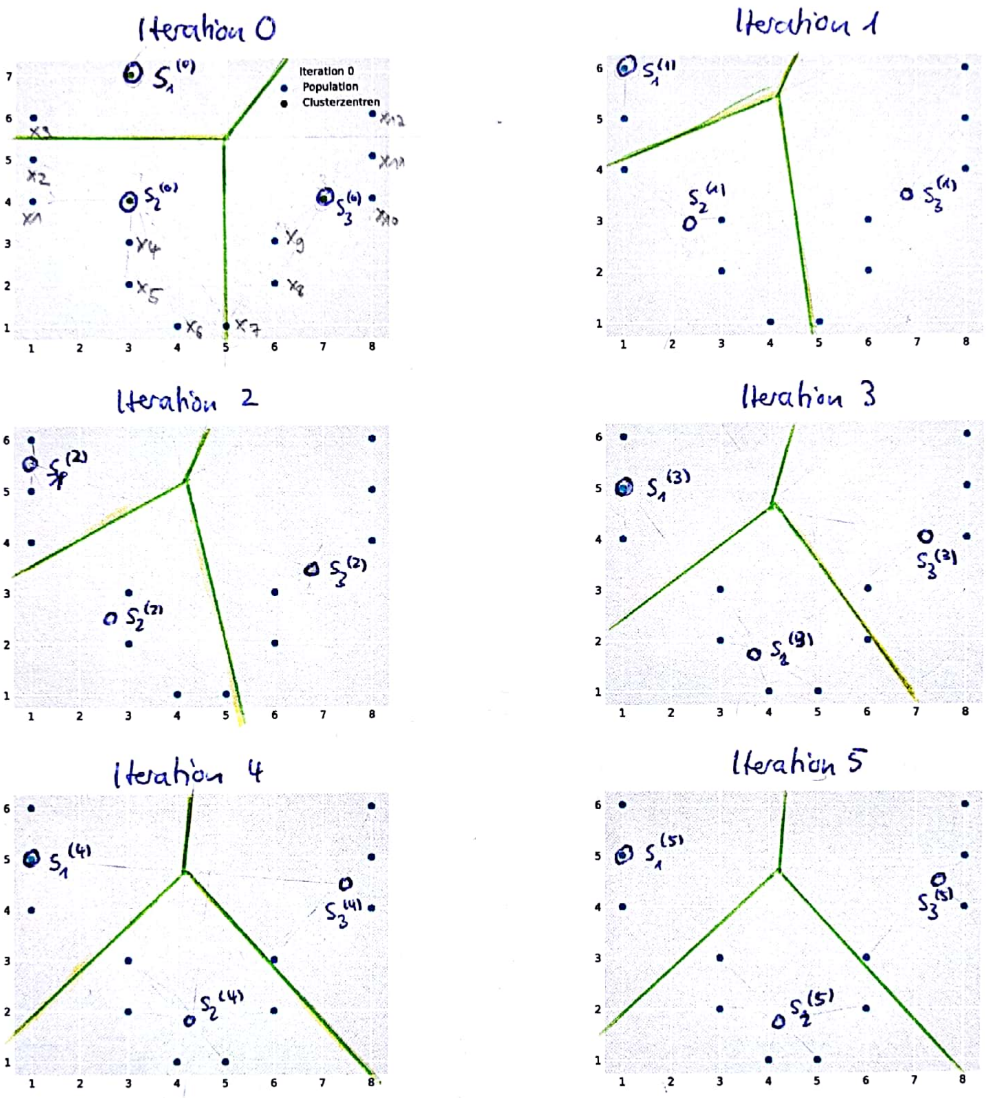

Abbildung 1: Population zum Einzeichnen der Clusterzentren und Clustergrenzen. Zu Aufgabe 17