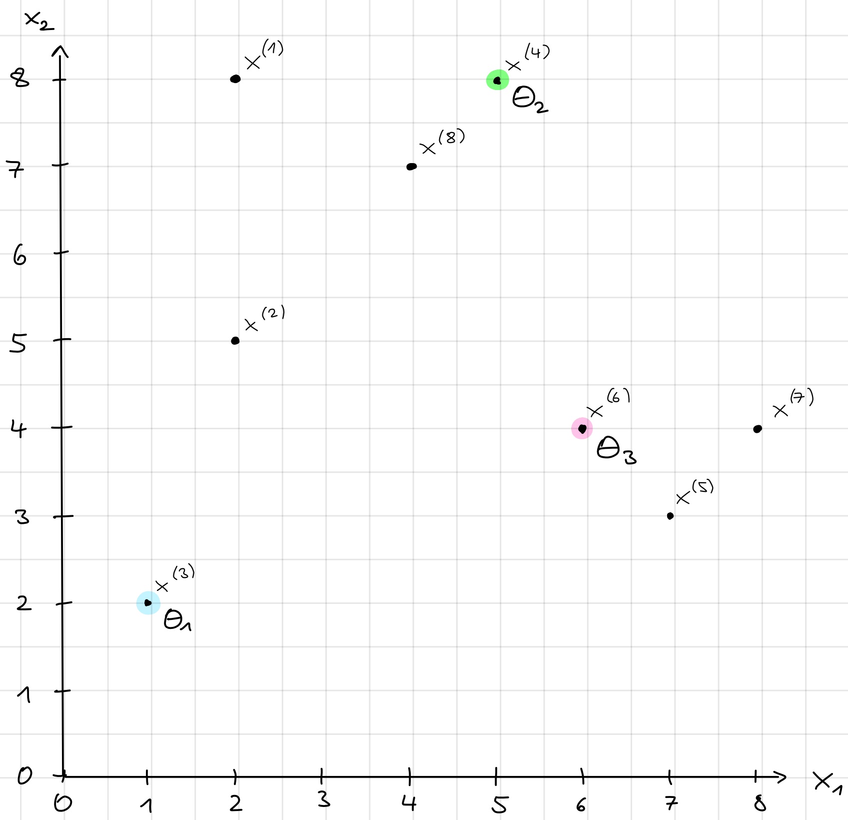
SDA - Sheet 09 - Exercise 1

Contributors:
* Annika Bätz: annika.baetz@uni-potsdam.de
* Max Serra: maximilia-manuel.serra.lasierra@uni-potsdam.de
* Adrián de Miguel: adrian.de.miguel.palacio@uni-potsdam.de
* Ignacio Llorca: ignacio.llorca.rodriguez@uni-potsdam.de

Forgy Initialization:



initial cluster centers:

$\theta_1 = x^{(3)}$

$\theta_2 = x^{(4)}$

$\theta_3 = x^{(6)}$

Notation: $x^{(i)} = (x_1^{(i)}, x_2^{(i)})$

$\theta_j = (\theta_{j1}, \theta_{j2})$

1. Iteration

$\mathcal{M}_1 = \mathcal{M}_2 = \mathcal{M}_3 = \{ \}$

The data points used for the initalization are easy to assign:

$\mathcal{M}_1 = \{x^{(3)}\}$, $\mathcal{M}_2 = \{x^{(4)}\}$, $\mathcal{M}_3 = \{x^{(6)}\}$

Let's compute the distance of each remaining datapoint to all cluster centers using the squared euclidian distance:

$$d(x^{(i)}, \theta_j) := \| x^{(i)} - \theta_j \|_2^2 = (x_1^{(i)} - \theta_{j1})^2 + (x_2^{(i)} - \theta_{j2})^2$$

$d(x^{(i)}, \theta_k)$	$\theta_1 = (1,2)$	$\theta_2 = (5,8)$	$\theta_3 = (6,4)$	$x^{(i)}$ is assigned to cluster:
$x^{(1)} = (2,8)$	37	9	32	θ_2
$x^{(2)} = (2,5)$	10	18	17	θ_1
$x^{(5)} = (7,3)$	37	29	2	θ_3
$x^{(7)} = (8,4)$	53	25	4	θ_3
$x^{(8)} = (4,7)$	34	2	13	θ_2

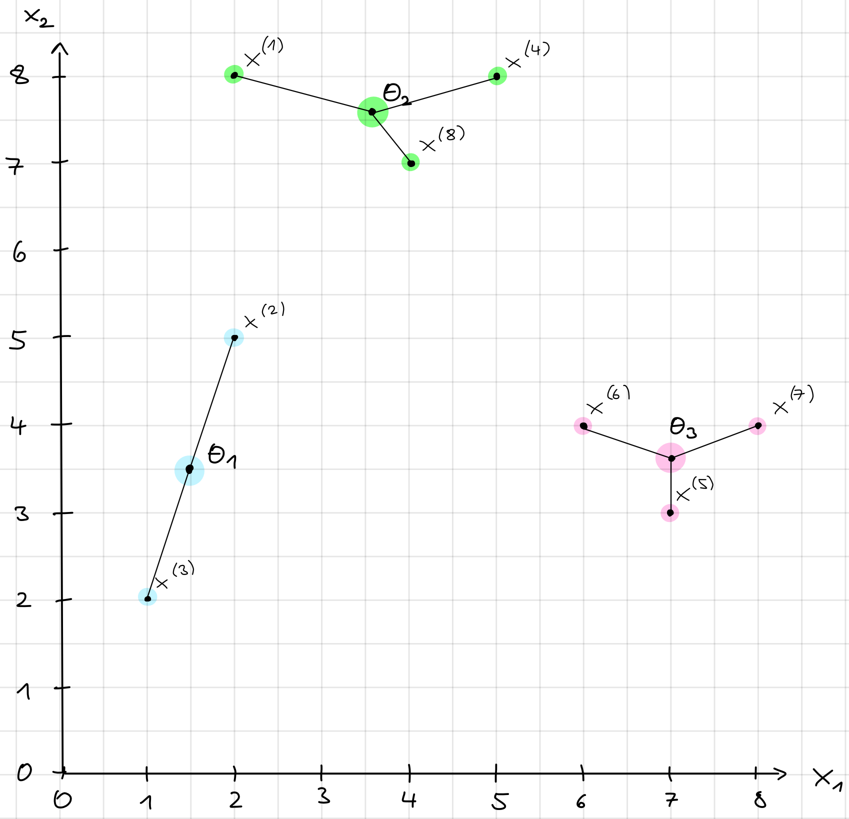
$$\Rightarrow \quad \mathcal{M}_1 = \{ x^{(3)}, x^{(2)} \}$$
$$\mathcal{M}_2 = \{ x^{(4)}, x^{(1)}, x^{(8)} \}$$
$$\mathcal{M}_3 = \{ x^{(6)}, x^{(5)}, x^{(7)} \}$$

Compute new cluster centers: $\quad \theta_k = \frac{1}{|\mathcal{M}_k|} \sum_{x^{(i)} \in \mathcal{M}_k} x^{(i)}$

$$\theta_1 = \frac{1}{2} \left(x^{(3)} + x^{(2)} \right) = (1.5, 3.5)$$

$$\theta_2 = \frac{1}{3} \left(x^{(4)} + x^{(1)} + x^{(8)} \right) = (11/3, 23/3)$$

$$\theta_3 = \frac{1}{3} \left(x^{(6)} + x^{(5)} + x^{(7)} \right) = (7, 11/3)$$



Loss: $\quad L = \sum_{k=1}^{3} \sum_{x^{(i)} \in \mathcal{M}_k} d(x^{(i)}, \theta_k) = \sum_{k=1}^{3} \sum_{x^{(i)} \in \mathcal{M}_k} \left((x_1^{(i)} - \theta_{k1})^2 + (x_2^{(i)} - \theta_{k2})^2 \right)$

$d(x^{(2)}, \theta_1) = 2.5$ $\quad d(x^{(1)}, \theta_2) \approx 2.89$ $\quad d(x^{(5)}, \theta_3) \approx 0.44$

$d(x^{(3)}, \theta_1) = 2.5$ $\quad d(x^{(4)}, \theta_2) \approx 1.89$ $\quad d(x^{(6)}, \theta_3) \approx 1.11$

$d(x^{(8)}, \theta_2) \approx 0.56$ $\quad d(x^{(7)}, \theta_3) \approx 1.11$

$L = 13.0$

2. Iteration

$\mathcal{M}_1 = \mathcal{M}_2 = \mathcal{M}_3 = \{\ \}$

$d(x^{(i)}, \theta_k)$	$\theta_1 = (1.5, 3.5)$	$\theta_2 = (\frac{11}{3}, \frac{23}{3})$	$\theta_3 = (7, \frac{11}{3})$	$x^{(i)}$ is assigned to cluster:
$x^{(1)} = (2,8)$	20.50	2.89	43.78	θ_2
$x^{(2)} = (2,5)$	2.50	9.89	26.78	θ_1
$x^{(3)} = (1,2)$	2.50	39.22	38.78	θ_1
$x^{(4)} = (5,8)$	32.50	1.89	22.78	θ_2
$x^{(5)} = (7,3)$	30.50	32.89	0.44	θ_3
$x^{(6)} = (6,4)$	20.50	18.89	1.11	θ_3
$x^{(7)} = (8,4)$	42.50	32.22	1.11	θ_3
$x^{(8)} = (4,7)$	18.50	0.56	20.11	θ_2

$\Rightarrow$ each datapoint is already assigned to its closest cluster center

$\Rightarrow$ $\mathcal{M}_1 = \{x^{(3)}, x^{(2)}\}$
$\mathcal{M}_2 = \{x^{(4)}, x^{(1)}, x^{(8)}\}$
$\mathcal{M}_3 = \{x^{(6)}, x^{(5)}, x^{(7)}\}$

remains unchanged

$\Rightarrow$ $L = 13.0$ unchanged