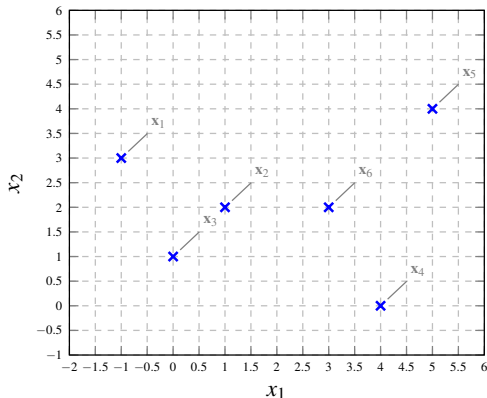


# K-Means Clustering

**Example:** Given the data set  $S$  and the initial centres as  $\mathbf{m}_1 = \mathbf{x}_1$  and  $\mathbf{m}_2 = \mathbf{x}_6$ , perform the K-means algorithm with  $K = 2$  until the means converge.

$$S = \{\underbrace{(-1, 3)}_{\mathbf{x}_1}, \underbrace{(1, 2)}_{\mathbf{x}_2}, \underbrace{(0, 1)}_{\mathbf{x}_3}, \underbrace{(4, 0)}_{\mathbf{x}_4}, \underbrace{(5, 4)}_{\mathbf{x}_5}, \underbrace{(3, 2)}_{\mathbf{x}_6}\}$$



**Example:** Initialisation:  $c = 2$ ;  $\mathbf{m}_1 = (-1, 3)$  and  $\mathbf{m}_2 = (3, 2)$

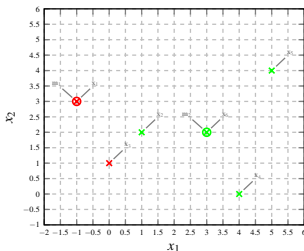
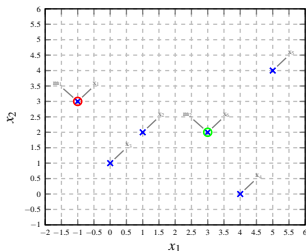
**First iteration:**

**Step 1:** Compute  $\|\mathbf{x}_i - \mathbf{m}_j\|$ ,  $i = 1, 2, \dots, 6$ ;  $j = 1, 2$  and classify samples.

$i$	$\mathbf{x}_i$	$\ \mathbf{x}_i - \mathbf{m}_1\ $	$\ \mathbf{x}_i - \mathbf{m}_2\ $	Assigned Cluster
1	$(-1, 3)$	0	4.1231	1
2	$(1, 2)$	2.2361	2.0000	2
3	$(0, 1)$	2.2361	3.1623	1
4	$(4, 0)$	5.8310	2.2361	2
5	$(5, 4)$	6.0828	2.8284	2
6	$(3, 2)$	4.1231	0.0000	2

**Step 2:** Recompute  $\mathbf{m}_j$ .

$$\mathbf{m}_1 = \frac{(-1, 3) + (0, 1)}{2} = (-0.5, 2), \quad \mathbf{m}_2 = \frac{(1, 2) + (4, 0) + (5, 4) + (3, 2)}{4} = (3.25, 2)$$



**Example:**

**Second iteration:**

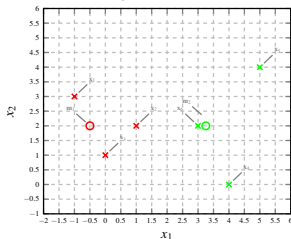
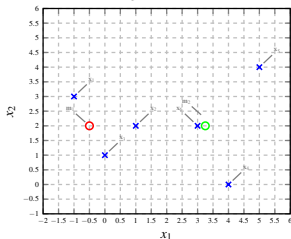
$$\mathbf{m}_1 = (-0.5, 2) \text{ and } \mathbf{m}_2 = (3.25, 2)$$

**Step 1:** Compute  $\|\mathbf{x}_i - \mathbf{m}_j\|$ ,  $i = 1, 2, \dots, 6$ ;  $j = 1, 2$  and classify samples.

$i$	$\mathbf{x}_i$	$\ \mathbf{x}_i - \mathbf{m}_1\ $	$\ \mathbf{x}_i - \mathbf{m}_2\ $	Assigned Cluster
1	$(-1, 3)$	1.1180	4.3661	1
2	$(1, 2)$	1.5000	2.2500	1
3	$(0, 1)$	1.1180	3.4004	1
4	$(4, 0)$	4.9244	2.1360	2
5	$(5, 4)$	5.8523	2.6575	2
6	$(3, 2)$	3.5000	0.25	2

**Step 2:** Recompute  $\mathbf{m}_j$ .

$$\mathbf{m}_1 = \frac{(-1, 3) + (1, 2) + (0, 1)}{3} = (0, 2), \mathbf{m}_2 = \frac{(4, 0) + (5, 4) + (3, 2)}{3} = (4, 2).$$



**Example:**

**Third iteration:**  $\mathbf{m}_1 = (0, 2)$  and  $\mathbf{m}_2 = (4, 2)$  do not change.

