Machine Learning 1, Fractal-2, 27/03/2021

$$C_i = \{x_1, x_2, x_3, x_4\} \Rightarrow \sum_{x \in C_i} d(x, \mu_i) = d(x_1, \mu_i) + d(x_2, \mu_i) + d(x_3, \mu_i) + d(x_4, \mu_i)$$

$$\mu_i = \frac{1}{|C_i|} \sum_{x \in C_i} x$$

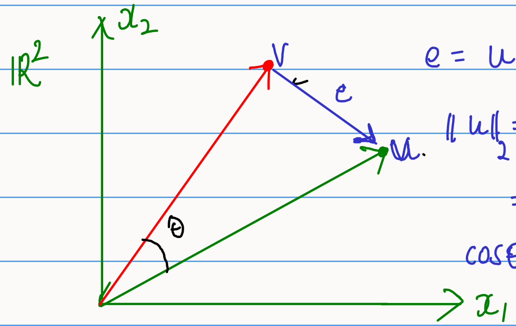
$|C_i| \leftarrow$ # Points in C_i

Here, $|C_i| = 4$

$$= \frac{1}{4}(x_1 + x_2 + x_3 + x_4)$$

$x \in \mathbb{R}^2$

$x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$



$e = u - v$

$\|u\|_2 = \sqrt{u_x^2 + u_y^2} \leftarrow L_2$ - Norm

$= \sqrt{u^T u}$

$\cos\theta = \frac{u^T v}{\|u\|_2 \|v\|_2}$

distance between two vectors $\underline{u}$ and $\underline{v}$

$= \|u - v\|_2 = \sqrt{(u-v)^T (u-v)}$

<u>Gradient</u> : let $f : \mathbb{R}^n \to \mathbb{R}$ be a function. Then, its gradient is defined as below:

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial v_1} \\ \frac{\partial f}{\partial v_2} \\ \vdots \\ \frac{\partial f}{\partial v_n} \end{bmatrix} \in \mathbb{R}^n$$

we assume that f is coordinatewise differentiable.

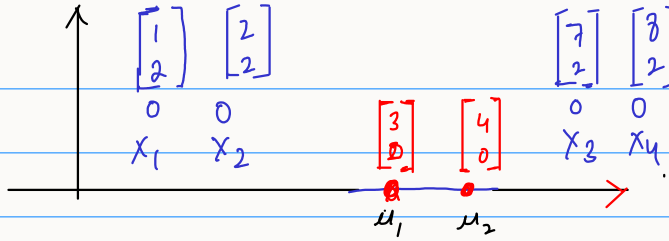
<u>Ex</u>: $f(v) = v^T v$, where $v \in \mathbb{R}^3$ $\quad v = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}$

$$= v_1^2 + v_2^2 + v_3^2$$

$$\frac{\partial f}{\partial v_1} = \frac{\partial}{\partial v_1}(v_1^2 + v_2^2 + v_3^2) = 2v_1$$

$$\frac{\partial f}{\partial v_2} = 2v_2 \quad \text{and} \quad \frac{\partial f}{\partial v_3} = 2v_3$$

$$\Rightarrow \nabla f = \begin{bmatrix} 2v_1 \\ 2v_2 \\ 2v_3 \end{bmatrix} = 2\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix} = 2v.$$



$\Rightarrow$ $C_1 = \{ x : d(x, \mu_1) < d(x, \mu_2) \}$

$\Rightarrow \quad = \{ x_1, x_2 \}$

$$d(x_1, \mu_1) = \sqrt{(1-3)^2 + (2-0)^2} = \sqrt{2^2 + 2^2} = \sqrt{8} = 2\sqrt{2}$$

distance between x_1 and μ_2

$$d(x_1, \mu_2) = \sqrt{(1-4)^2 + (2-0)^2} = \sqrt{8^2 + 2^2} = \sqrt{9+4} = \sqrt{13}$$

$\Rightarrow \quad d(x_1, \mu_1) < d(x_1, \mu_2) \Rightarrow x_1 \in C_1.$

$C_2 = \{ x : d(x, \mu_2) < d(x, \mu_1) \}$

$= \{ x_1, x_2 \}$

distance between x_4 and μ_1

$$d(x_4, \mu_1) = \| x_4 - \mu_1 \|_2 = \sqrt{(8-3)^2 + (2-0)^2} = \sqrt{5^2 + 2^2} = \sqrt{29}$$

distance between x_4 and μ_2

$$d(x_4, \mu_2) = \| x_4 - \mu_2 \|_2 = \sqrt{(8-4)^2 + (2-0)^2} = \sqrt{4^2 + 2^2} = \sqrt{20} < \sqrt{29}$$

$\Rightarrow \quad d(x_4, \mu_2) < d(x_4, \mu_1)$

$\Rightarrow \quad x_4 \in C_2$

Similarly we can show that $x_3 \in C_2$, and $x_2 \in C_1$

$\Rightarrow \quad C_1 = \{ x_1, x_2 \}, \quad C_2 = \{ x_3, x_4 \}.$

Initialize cluster centers μ_1 and μ_2 to some random locations.

We have found the clusters C_1 and C_2 given the initial center μ_1 and μ_2.

Now, given these clusters, we can update our cluster centers.

$$\Rightarrow \mu_1 = \frac{1}{|C_1|} \sum_{x \in C_1} x = \frac{1}{2} \left(\begin{bmatrix} 1 \\ 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right) = \frac{1}{2} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1.5 \\ 2 \end{bmatrix}$$

$$\Rightarrow \mu_2 = \frac{1}{|C_2|} \sum_{x \in C_2} x = \frac{1}{2} \left(\begin{bmatrix} 7 \\ 2 \end{bmatrix} + \begin{bmatrix} 8 \\ 2 \end{bmatrix} \right) = \frac{1}{2} \begin{bmatrix} 15 \\ 2 \end{bmatrix} = \begin{bmatrix} 7.5 \\ 2 \end{bmatrix}$$