Live Interaction #5:

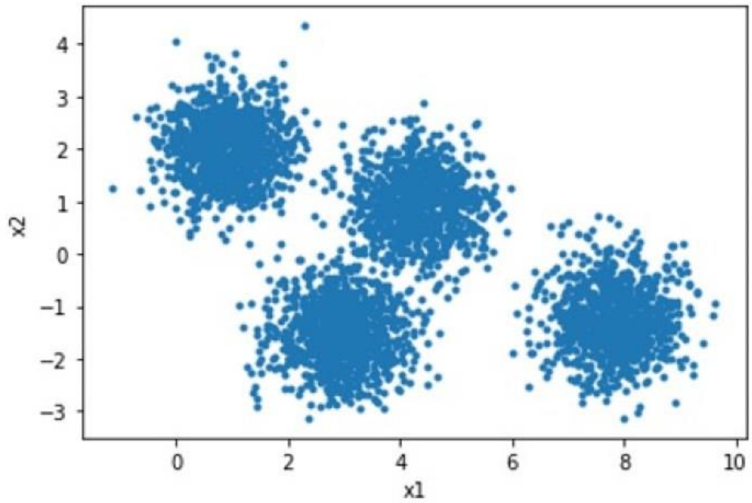
11th February 2024

E-masters Next Generation Wireless Technologies

EE902 Advanced ML Techniques for Wireless Technology

- **K-Means Clustering:**
- **Unsupervised learning: Data without labels!**

$$\underbrace{\bar{\mathbf{x}}(1), \bar{\mathbf{x}}(2), \dots, \bar{\mathbf{x}}(M)}_{\text{Training data}}$$



- K −clusters
- Cluster assignment:

$$\alpha_i(j) = \begin{cases} 1 & \bar{\mathbf{x}}(j) \in \mathcal{C}_i \\ 0 & \bar{\mathbf{x}}(j) \notin \mathcal{C}_i \end{cases}$$

- **Cost-function:**

$$\sum_{i=1}^{K}\sum_{j=1}^{M}\alpha_i(j)\|\bar{\mathbf{x}}(j)-\bar{\boldsymbol{\mu}}_i\|^2$$

- **Iterative procedure:**
- Consider the centroids after iteration $l-1$

$$\bar{\boldsymbol{\mu}}_1^{(l-1)}, \bar{\boldsymbol{\mu}}_2^{(l-1)}, \dots, \bar{\boldsymbol{\mu}}_K^{(l-1)}$$

- Given the centroids, we determine ***cluster assignment***.

$$\alpha_i^{(l)}(j) = ?$$

- Assign to cluster $\tilde{\imath}$

$$\underbrace{\tilde{\imath} = \arg\min_i \left\|\bar{\mathbf{x}}(j) - \bar{\boldsymbol{\mu}}_i^{(l-1)}\right\|^2}_{\text{Assign } \bar{\mathbf{x}}(j) \text{ to cluster } \tilde{\imath} \text{ with closest centroid}}$$

$$\alpha_i^{(l)}(j) = 1 \text{ if } i = \tilde{\imath}$$

- What is the property $\alpha_i(j)$ obey?

$$\alpha_i^{(l)}(j) \in \{0,1\}$$

$$\sum_{i=1}^{K}\alpha_i^{(l)}(j) = 1$$

- Next determine the centroids:

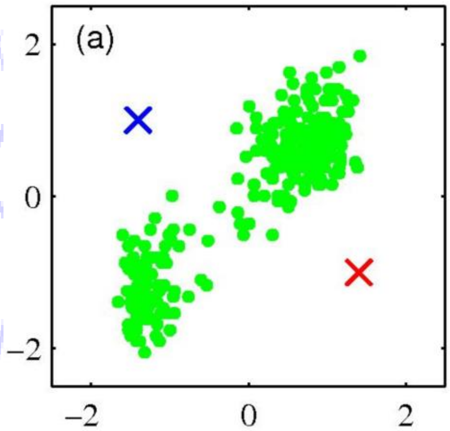
$$\bar{\boldsymbol{\mu}}_1^{(l)}, \bar{\boldsymbol{\mu}}_2^{(l)}, \dots, \bar{\boldsymbol{\mu}}_K^{(l)}$$

- How to determine the centroids?

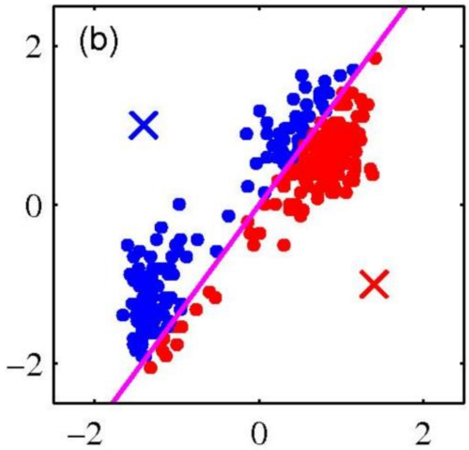
$$\bar{\boldsymbol{\mu}}_i^{(l)} = \frac{\text{Sum of all points in cluster } i}{\text{Number of points in cluster } i}$$

$$= \frac{\sum_{j=1}^{M} \alpha_i^{(l)}(j)\bar{\mathbf{x}}(j)}{\sum_{j=1}^{M} \alpha_i^{(l)}(j)}$$

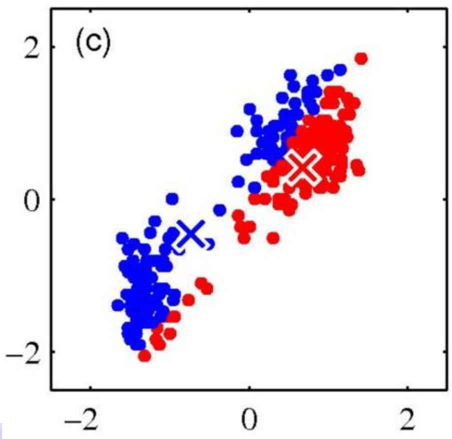
- Repeat until convergence!
- Example:
- Iteration 0: **Start with random placement of centroids**



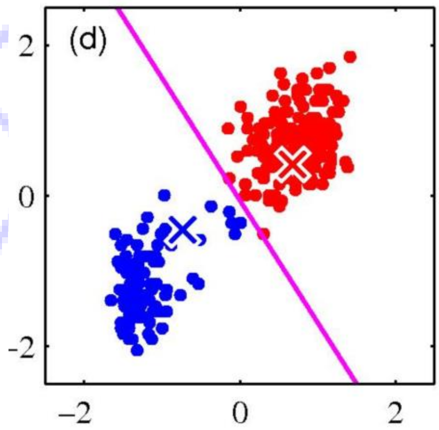
- **Iteration 1:**
- Cluster assignment:

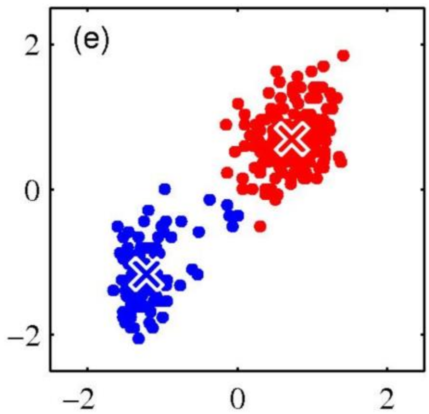


- Centroid computation



- **Iteration 2:**





- **Linear Discriminant Analysis:**
- Gaussian discriminant analysis:

$$\mathcal{C}_0 \sim \mathcal{N}(\bar{\boldsymbol{\mu}}_0, \mathbf{R})$$
$$\mathcal{C}_1 \sim \mathcal{N}(\bar{\boldsymbol{\mu}}_1, \mathbf{R})$$

- Given $\bar{\mathbf{x}}$, How to determine the class?

$$p(\bar{\mathbf{x}}; \mathcal{C}_0) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{R}|}} e^{-\frac{1}{2}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_0)}$$

$$p(\bar{\mathbf{x}}; \mathcal{C}_1) = \frac{1}{\sqrt{(2\pi)^n |\mathbf{R}|}} e^{-\frac{1}{2}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_1)^T \mathbf{R}^{-1}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_1)}$$

- **Multivariate Gaussian density**.
- Choose $\mathcal{C}_0$ if

$$p(\bar{\mathbf{x}}; \mathcal{C}_0) \geq p(\bar{\mathbf{x}}; \mathcal{C}_1)$$

$$\frac{\cancel{1}}{\cancel{\sqrt{(2\pi)^n |\mathbf{R}|}}} e^{-\frac{1}{2}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_0)} \geq \frac{\cancel{1}}{\cancel{\sqrt{(2\pi)^n |\mathbf{R}|}}} e^{-\frac{1}{2}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_1)^T \mathbf{R}^{-1}(\bar{\mathbf{x}}-\bar{\boldsymbol{\mu}}_1)}$$

$$(\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_0)^T \mathbf{R}^{-1}(\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_0) \leq (\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_1)^T \mathbf{R}^{-1}(\bar{\mathbf{x}} - \bar{\boldsymbol{\mu}}_1)$$

$$\Rightarrow \bar{\mathbf{h}}^T(\bar{\mathbf{x}} - \tilde{\boldsymbol{\mu}}) \geq 0$$

$$\bar{\mathbf{h}} = \mathbf{R}^{-1}(\bar{\boldsymbol{\mu}}_0 - \bar{\boldsymbol{\mu}}_1)$$

$$\tilde{\boldsymbol{\mu}} = \frac{\bar{\boldsymbol{\mu}}_0 + \bar{\boldsymbol{\mu}}_1}{2}$$

- Find the linear discriminant:

$$\mathcal{C}_0 \sim \mathcal{N}\left(\begin{bmatrix} -3 \\ 2 \end{bmatrix}, \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{6} \end{bmatrix}\right), \mathcal{C}_1 \sim \mathcal{N}\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}, \begin{bmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{6} \end{bmatrix}\right)$$

$$\bar{\mathbf{h}} = \mathbf{R}^{-1}(\bar{\boldsymbol{\mu}}_0 - \bar{\boldsymbol{\mu}}_1)$$

$$= \begin{bmatrix} 4 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} -5 \\ 3 \end{bmatrix} = \begin{bmatrix} -20 \\ 18 \end{bmatrix}$$

$$\tilde{\boldsymbol{\mu}} = \frac{\bar{\boldsymbol{\mu}}_0 + \bar{\boldsymbol{\mu}}_1}{2} = \begin{bmatrix} -\frac{1}{2} \\ \frac{1}{2} \end{bmatrix}$$

- Choose $\mathcal{C}_0$ if

$$\bar{\mathbf{h}}^T(\bar{\mathbf{x}} - \tilde{\boldsymbol{\mu}}) \geq 0$$

$$\Rightarrow \begin{bmatrix} -20 & 18 \end{bmatrix} \left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \frac{1}{2} \begin{bmatrix} -1 \\ 1 \end{bmatrix} \right) \geq 0$$

$$\Rightarrow -20x_1 + 18x_2 - 10 - 9 \geq 0$$

$$\Rightarrow -20x_1 + 18x_2 \geq 19$$

- Choose $\mathcal{C}_1$ if

$$-20x_1 + 18x_2 < 19$$

- What happens to classifier when prior probabilities are

$$p(\mathcal{C}_0) = p_0$$

$$p(\mathcal{C}_1) = p_1$$

- Determine the linear classifier for the above case.
- **Assignment #5 Deadline: 16th Feb Friday 11:59 PM.**
- **Live interaction #6: 18th February Sunday 2:00 – 3:00 PM.**
- **Assignment #6 Deadline: 23rd Feb Friday 11:59 PM.**
- **Assignment #5, 6 Discussion: 24th Feb Saturday 2:00 PM – 3:00 PM.**
- **Quiz #3: 24th February 3:30 - 4:30 PM.**
- **Live interaction #7: 25th February Sunday 2:00 – 3:00 PM.**
- **Final Exam: 10th March Sunday 9:00 AM – 12:00 PM. (Please check!!)**