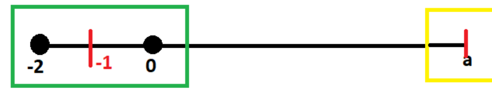
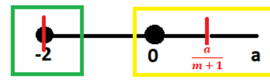
Ahmed Nouralla – B19-CS-01 – a.shaaban@innopolis.university

ML Assignment 2 – Theoretical Part

1. K-Means Clustering





- There are two possible clustering solutions that K-means will generate given the constraints $K = 2, a > 0, m > 0$
- Either create one cluster for the point a if it's too far from the others.
- Or include a in D_2 (the m-points of value 0)
- Let's calculate total Residual Sum of Squares for both cases:

$$J_1 = \sum_{i\in\{1,2\}} \sum_{x\in D_i} (x-\mu_i)^2 = m\big(-2-(-1)\big)^2 + m\big(0-(-1)\big)^2 + 0^2 = 2m$$

$$J_2 = \sum_{i\in\{1,2\}} \sum_{x\in D_i} (x-\mu_i)^2 = 0 + m\left(0-\frac{a}{m+1}\right)^2 + \left(a-\frac{a}{m+1}\right)^2 = \frac{a^2 m}{m+1}$$

- The global optimal solution is the one that minimizes the total heterogeneity for all clusters, which we known to be J_2, thus $J_2 < J_1$
- Doing the math:

$$\frac{a^2 m}{m+1} < 2m \Rightarrow a^2 < \frac{2m(m+1)}{m} \Rightarrow a^2 < 2m+2 \Rightarrow f(m) = 2m+2$$

2. Support Vector Machines (SVMs)

(I)

a. Yes.
b. No, because the hyperplane must pass through zero (since $\theta_0 = 0$).
c. No, because the margin is soft (some errors are tolerated), but there are no slack variables in the equation

(II)

a. No, because the hyperplane is passing through zero (although there is a non-zero bias term in the equation).
b. Yes.
c. No, because the margin is soft (some errors are tolerated), but there are no slack variables in the equation.

(III)

a. Yes.
b. No, because the hyperplane must pass through zero (since there is no bias term in the equation).
c. Yes.