A gradient clustering algorithm

Manav Prabhakar

February 24, 2021

Our primary objective is to cluster a set of observations. The approach implemented here is moving each observation along the direction of the gradient at the observation points.

There can be two approaches for the same, one of which is moving a small step in the direction of the gradient and iteratively repeat the process until tight clusters are obtained.

The other approach involves shifting each observation by some amount proportional to the gradient at the observation point.

$$X_j^{i+1} = X_j^i + \left(a \times \nabla_x In\, p\left(X_j^i\right) \right) \tag{1}$$

Where,

$$\nabla_x In\, P(x) = \frac{\nabla_x P(x)}{P(x)} = \frac{n+2}{h} \times M_h(x) = \frac{n+2}{h^2} \times \frac{1}{n} \times \sum_{k=1}^{n} (X_k - X)$$

Each observation is transformed recursively using the Eq. 1 to cluster up the observations.

For experimentation purposes, a bivariate normal distribution was simulated with three classes each having 60 observations each.

Each class had identity covariance matrix, but different means.

The three means used were:

$$m1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, m2 = \begin{pmatrix} 4 \\ 0 \end{pmatrix}, m3 = \begin{pmatrix} 0 \\ 4 \end{pmatrix}$$
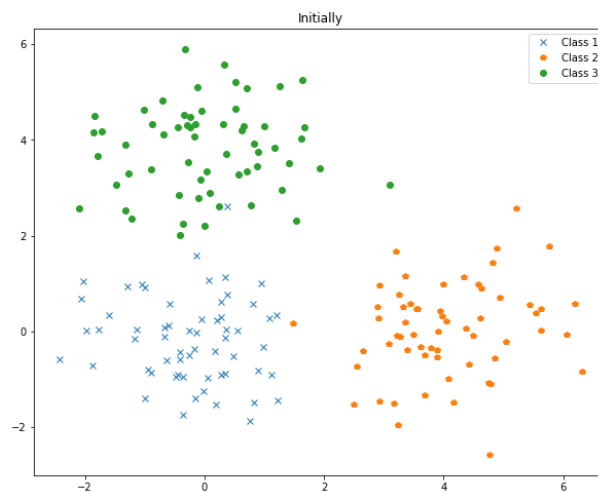

Figure 1: Sets obtained for 3 different classes using Bivariate distribution with different means but similar
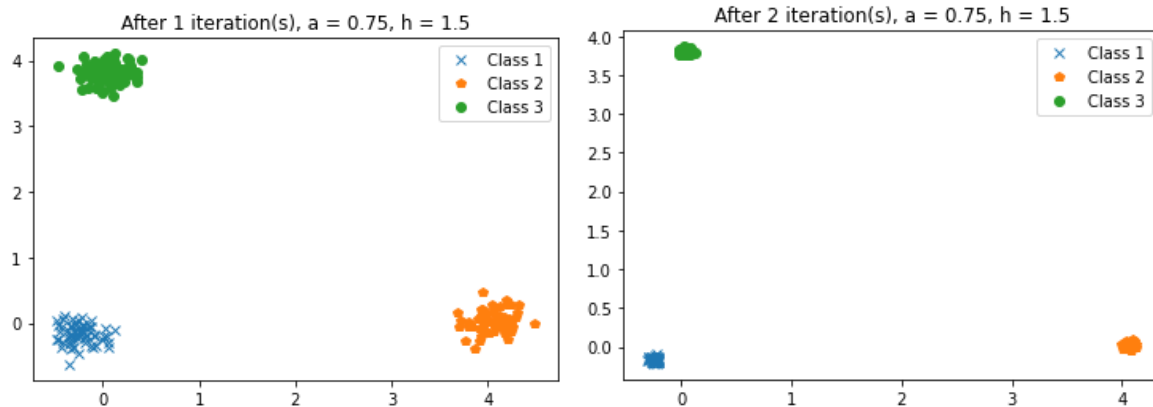.

Figure 2: Successive iterations of the clustering algorithm with a = 0.75, h = 1.5.

The next part of the experiment included changing the parameters a and h. The new values used for the same were a = 0.5, h = 0.8. A new distribution was generated for this.
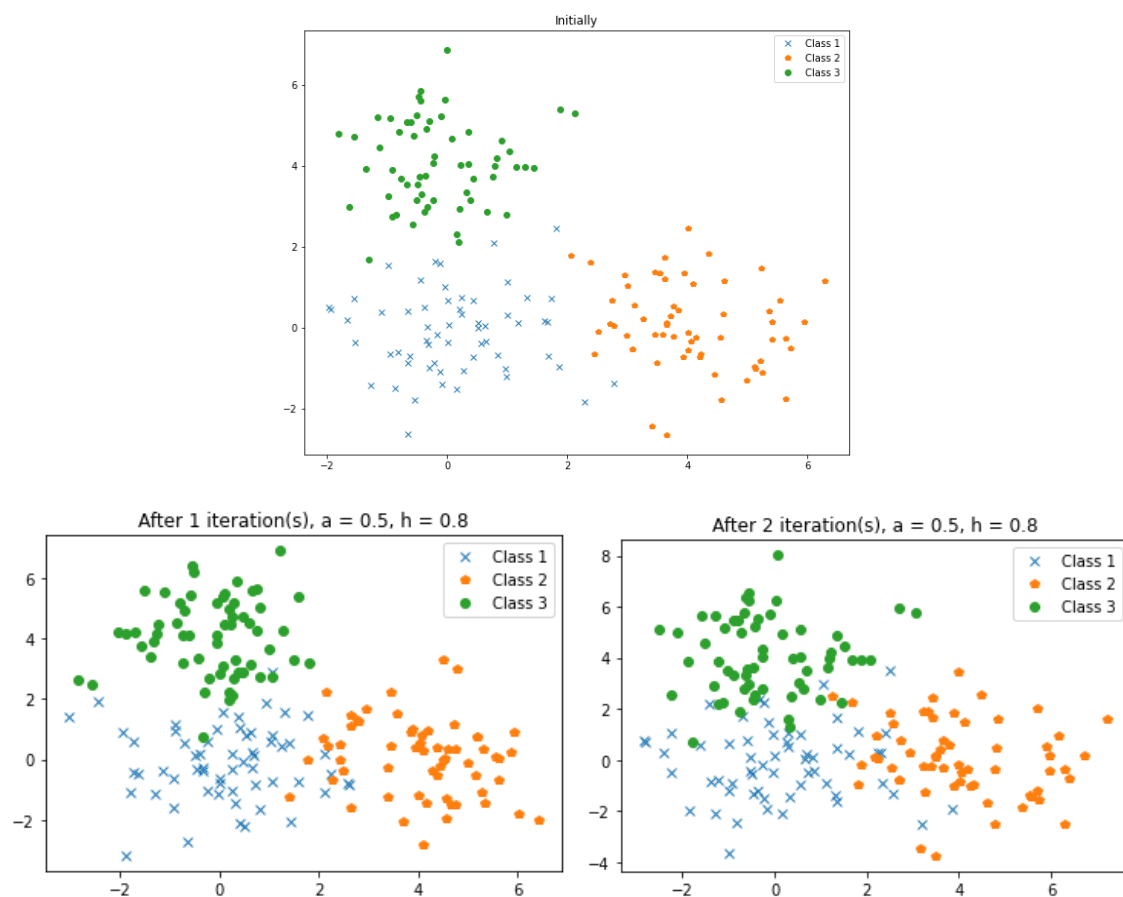

Figure 2: Successive iterations of the clustering algorithm with a = 0.5, h = 0.8.