

PRÁCTICA 7 - AGRUPAMIENTO

Fernando Aliaga Ramón - 610610

1. Datos Sintéticos 2D

Vamos a implementar el algoritmo KMedias. Para ello hay que completar tres funciones: `updateCentroids`, `updateClusters` y `kmeans`.

UpdateCentroids

La función actualiza la media de los puntos asignados a cada cluster. La implementación en matlab es la siguiente:

```
for i = 1:K
    idx = find(c==i);
    a = D(idx,1:n);
    newc = [newc ; mean(a)];
end
```

UpdateClusters

La función actualiza el cluster al que corresponde cada punto. La implementación en matlab es la siguiente:

```
for xi = 1:size(D,1)
    x = D(xi, :);
    best = Inf;
    for mui = 1:K
        mu = c(mui, :);
        d = dot(x - mu, x - mu);
        verf(xi,mui) = d;

        if d < best
            best = d;
            newc(xi) = mui;
            verf(xi,size(c, 1)+1) = mui;
            verf(xi,size(c, 1)+2) = best;
        end
    end
end
```

K-Means

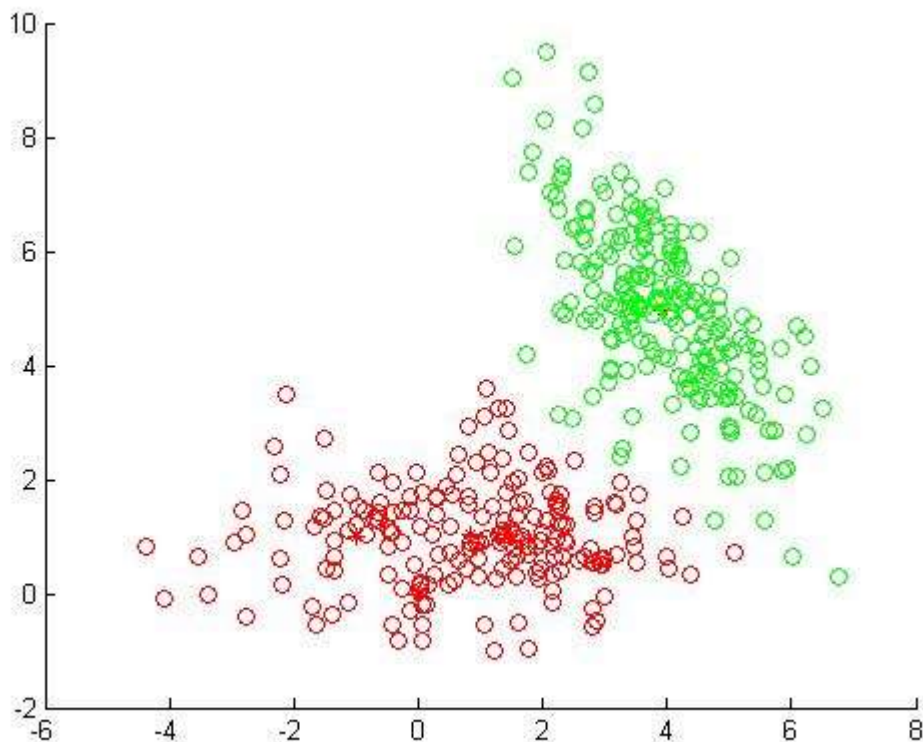
El algoritmo K-Means consiste en un bucle con el cual actualizamos los clusters mediante los dos métodos antes implementados hasta que estos no varíen.

La implementación en matlab es la siguiente:

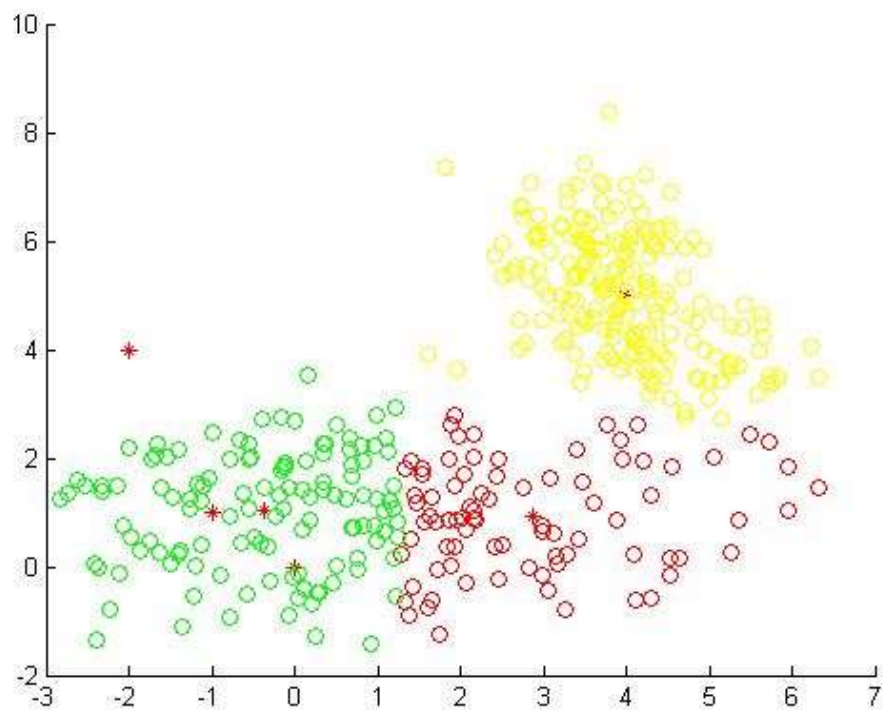
```
Z = updateClusters(D,c0);  
cAnterior = zeros(size(c0,1),size(c0,2));  
c = updateCentroids(D,Z,size(c0,1));  
i = 0;  
  
while (c~=cAnterior)  
    fprintf('Iteracion k-means : %d \n',i);  
    i = i + 1;  
    cAnterior = c;  
    Z = updateClusters(D,c);  
    c = updateCentroids(D,Z,size(c0,1));  
end  
fprintf('Fin k-means\n');
```

Vamos a entrenar ahora con diferente número de centroides:

K=2



K=3



K=4

