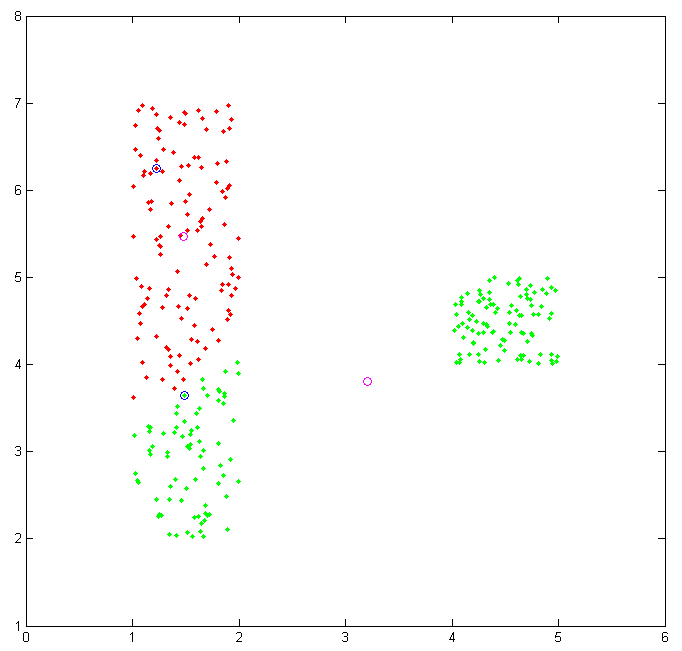
Machine Learning – Assignment 5

103062528 林玉山

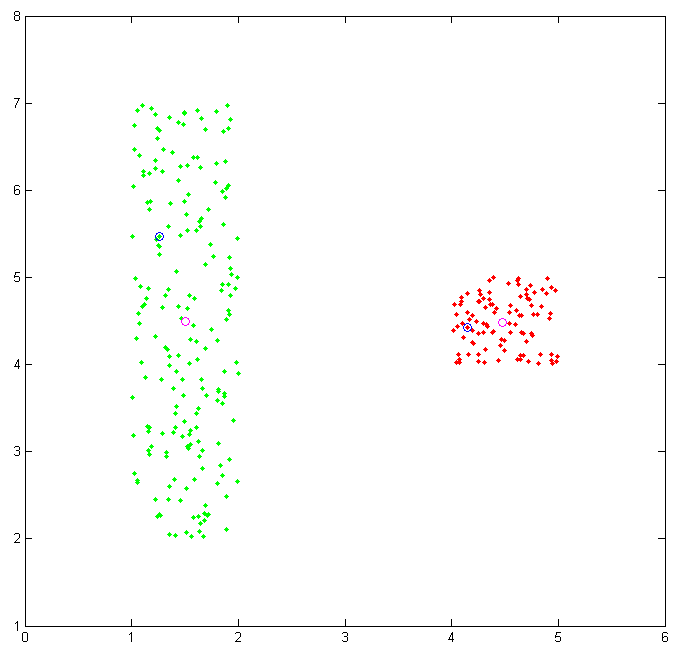
1. **喵**
2. **喵**
3. **令d=2, k=2，畫出一張一開始的mean選得不好的圖，以及一張選得好的圖**

下兩張圖為同一組data set，分為紅色與綠色兩群。藍色圈圈為一開始的means，紫色為最後收斂時的means。

第一張圖為一開始的兩個means都從左側那群點中取出。從最後收斂的結果可以看出來，左半部的下半段被分進了右邊那群中。很明顯就不是我們要的結果。



第二張圖則一開始就從左右兩群個挑一個點出來做為初始值。可以看見這樣的情況下，左右兩群就被正確地分開了。並且最後得到的means，也正好在兩群的正中央。正是我們期望的結果。由此可以得到「一開始挑選的點會大大影響結果」這個結論。



1. **喵**
2. **令有一 Graph Laplacian Matrix L，請證明下列性質**

**(a) 證明對於任意向量 ，**

首先，我們先計算L

接著，我們展開

當我們計算綠色部分乘上L時，我們令每一個column的結果為

此時 變為：

我們再將 各自代入，其中對於每一個

因此 為：

在上式中，我們讓左式(綠色)與右式之中，index相反的項相加，可以得到下式

又因為矩陣S是Symmetric的，所以對於

將t換成j即為原式，故得證。

**(b) 試證明**

目前已知 (注意 是純量)：

接著對於第i與第j個data point，一共會有四種情形：

1. i與j同屬於G
2. i屬於G，j不屬於G
3. i不屬於G，j屬於G
4. i與j皆不屬於G

因此我們可以依照四種將原式中的 拆開：

此時代入下式

因為 ，因此

得證。

1. **喵**