Machine Learning (Homework 3)

0753420 郭家瑄

1. Gaussian Process

(1) Prediction result

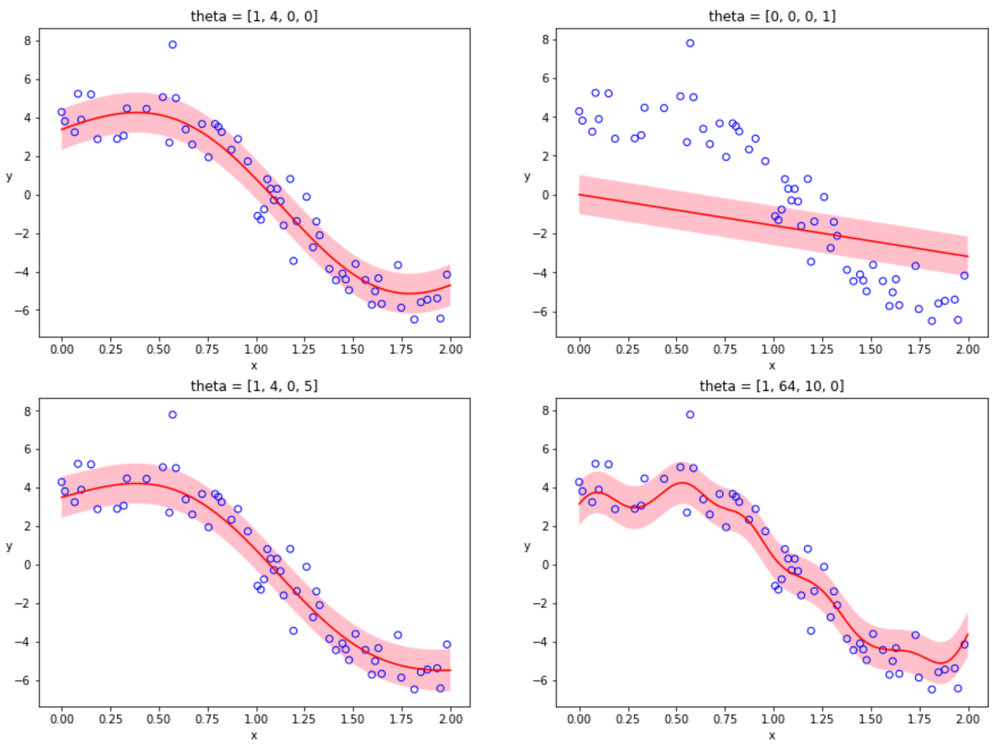
Using exponential-quadratic kernel function given by

$$k(\mathbf{x}_n, \mathbf{x}_m) = \theta_0 \exp\left\{-\frac{\theta_1}{2}||\mathbf{x}_n - \mathbf{x}_m||^2\right\} + \theta_2 + \theta_3 \mathbf{x}_n^\top \mathbf{x}_m$$

Where the hyperparameters θ have four different combinations:

- squared exponential kernel $\boldsymbol{\theta} = \{1, 4, 0, 0\}$
- linear kernel $\boldsymbol{\theta} = \{0, 0, 0, 1\}$
- exponential-quadratic kernel $\boldsymbol{\theta} = \{1, 4, 0, 5\}$
- exponential-quadratic kernel $\boldsymbol{\theta} = \{1, 64, 10, 0\}$

The prediction result is shown as below.



(2) Root-mean-square errors

The corresponding root-mean-square errors for both training and test sets with respect to four kernels are shown below.

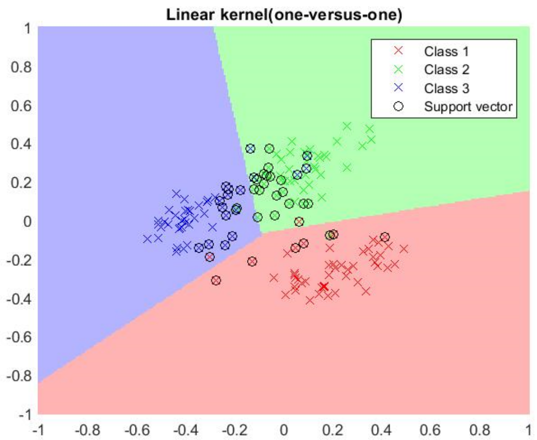
theta	train RMS	test RMS
[1, 4, 0, 0]	1.0859078251089593	1.0167599208810119
[0, 0, 0, 1]	3.375501292224377	3.8497751148340535
[1, 4, 0, 5]	1.076835972426699	1.057627032808463 6
[1, 64, 10, 0]	1.0606651654811656	1.243345786045694

(3) Discussion

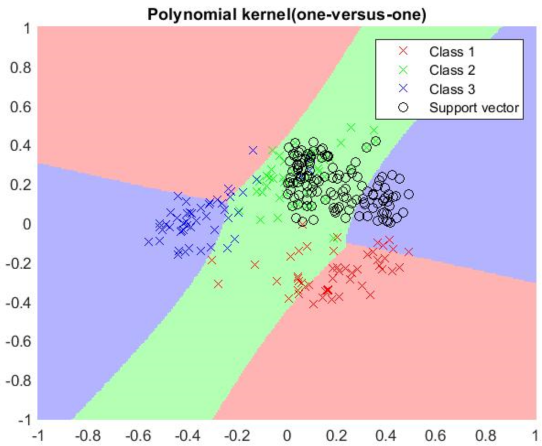
透過設定不同的θ可以得到不同的 kernel function，從實驗結果可以看的出來，若是 kernel function 過於簡單(如 theta=[0, 0, 0, 1])，會使得可以表達的 function 變的比較不靈活，因此無法很好的 fitting 資料的分布，也使得計算出的 error 值偏高。

2. Support Vector Machine

(1) Linear kernel with multi-class classification result



(2) Polynomial kernel with multi-class classification result



(3) Discussion

以上兩種不同 kernel 的 SVM 都是使用 one-versus-one 的方法來分類，從圖上的分群切割結果可發現，polynomial kernel 相較於 linear kernel 可以將分群空間切割的更為複雜。

3. Gaussian Mixture Model

(1) The estimated $\{\mu_k\}_{k=1}^{K}$

- $k = 2$

k-means mean value	r	g	b
0	182	203	230
1	107	87	35

- $k = 3$

k-means mean value	r	g	b
0	67	60	45
1	182	203	230
2	146	113	25

- $k = 5$

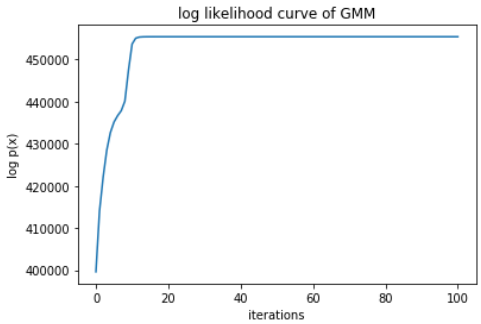
k-means mean value	r	g	b
0	155	120	16
1	38	60	89
2	185	206	234
3	138	134	144
4	87	63	16

- $k = 20$

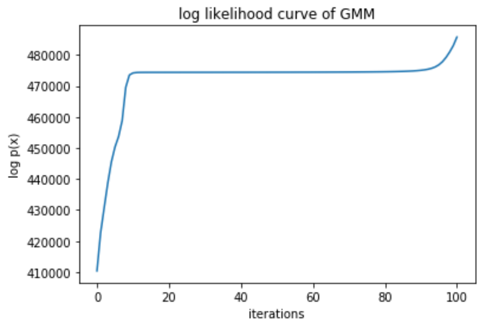
k-means mean value	r	g	b
0	8	47	82
1	124	99	43
2	112	100	106
3	43	30	6
4	58	76	107
5	139	131	143
6	86	67	61
7	99	7	11
8	165	190	221
9	199	218	244
10	187	146	18
11	151	160	184
12	104	76	10
13	183	190	198
14	171	152	77
15	182	207	237
16	244	247	249
17	133	99	6
18	159	123	12
19	77	55	9

(2) Log likelihood curve of GMM

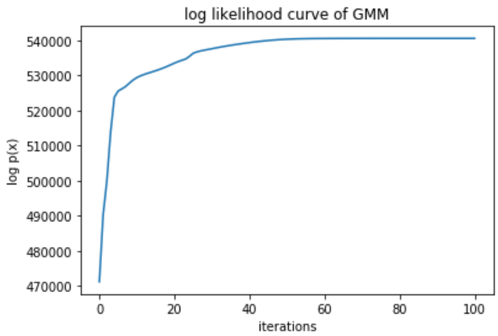
$k = 2$



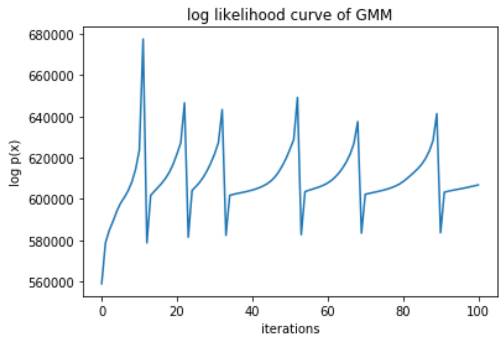
$k = 3$



$k = 5$

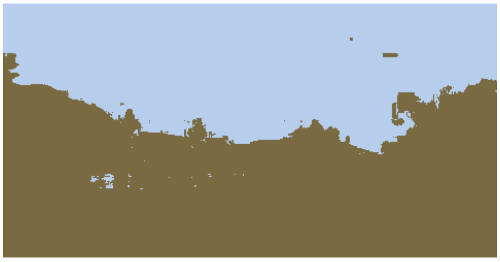


$k = 20$



(3) The resulting images

$k = 2$



$k = 3$

$k = 5$

$k = 20$