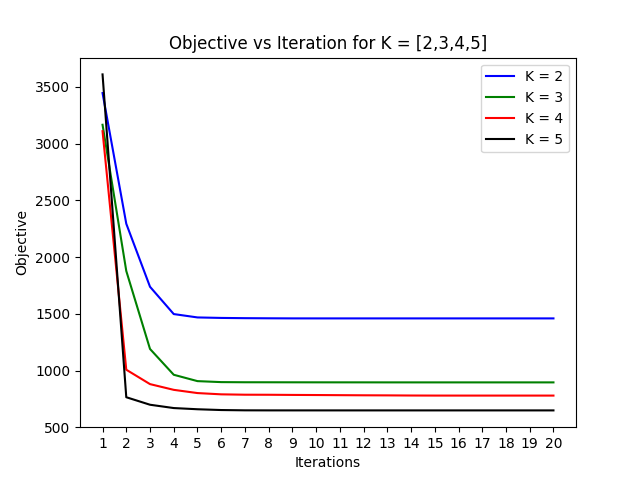
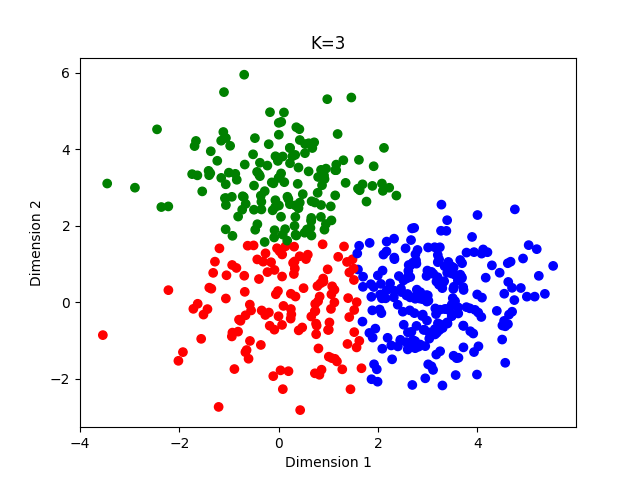
# Hw4 Report

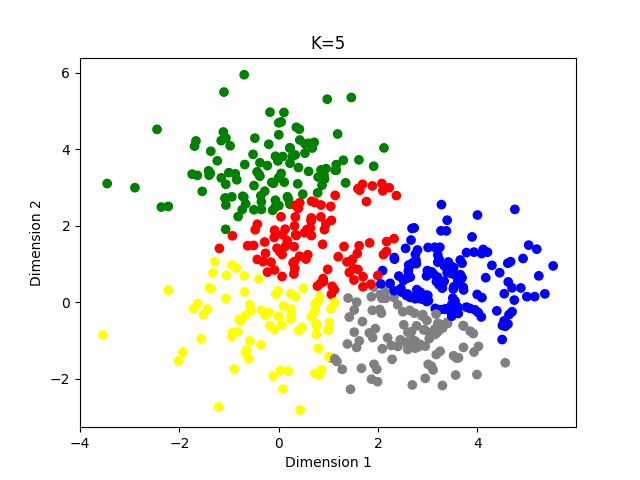
Wenbo Song

Ws2505

Problem 1

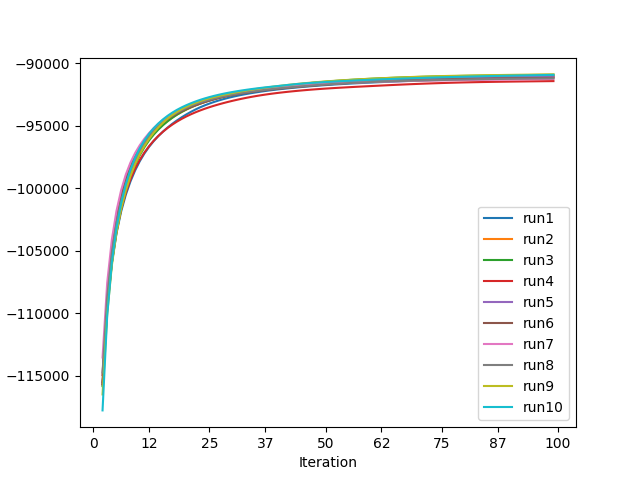
1. For K = 2,3,4,5, the value of K-means objective function per iteration for 20 iterations can be seen as follow:
2. 500 data points for K = 3, 5 is shown as follow:





Problem 2

1. The log joint likelihood for iterations 2 to 100 is shown as follow:



The training objective function and RMSE on the testing set for the 10 runs

|  |  |  |
| --- | --- | --- |
| In which run | Log joint likelihood | RMSE |
| 9 | -90889.06654 | 1.111649936 |
| 3 | -90918.16356 | 1.110751615 |
| 10 | -90928.72693 | 1.107983291 |
| 1 | -91044.9254 | 1.0948742 |
| 5 | -91044.9254 | 1.13924095 |
| 2 | -91082.05593 | 1.134378454 |
| 6 | -91163.65205 | 1.150539082 |
| 8 | -91204.67559 | 1.09554172 |
| 7 | -91256.348 | 1.106586519 |
| 4 | -91424.88786 | 1.126603782 |

1. Ten movies that are closest to movies “Star Wars”, “My Fair Lady” and “Goodfellas”

|  |  |  |
| --- | --- | --- |
| Star Wars | My Fair Lady | GoodFellas |
| Empire Strikes Back, The (1980) | Sound of Music, The (1965) | Once Upon a Time in the West (1969) |
| Return of the Jedi (1983) | Mary Poppins (1964) | Good, The Bad and The Ugly, The (1966) |
| Raiders of the Lost Ark (1981) | Shadowlands (1993) | Full Metal Jacket (1987) |
| Usual Suspects, The (1995) | Alice in Wonderland (1951) | Godfather: Part II, The (1974) |
| Princess Bride, The (1987) | Washington Square (1997) | Bonnie and Clyde (1967) |
| Sting, The (1973) | SubUrbia (1997) | Short Cuts (1993) |
| Terminator, The (1984) | Snow White and the Seven Dwarfs (1937) | Apocalypse Now (1979) |
| Blues Brothers, The (1980) | Gone with the Wind (1939) | Godfather, The (1972) |
| Indiana Jones and the Last Crusade (1989) | Ninotchka (1939) | 2001: A Space Odyssey (1968) |
| L.A. Confidential (1997) | Cinderella (1950) | Cool Hand Luke (1967) |

The distance of the query movie is shown as below:

|  |  |  |
| --- | --- | --- |
| Star Wars | My Fair Lady | GoodFellas |
| 0.328800011 | 0.676627739 | 0.554563938 |
| 0.616614867 | 0.928825099 | 0.594204485 |
| 0.621324346 | 0.932651283 | 0.624790998 |
| 0.807932304 | 1.050563988 | 0.724900652 |
| 0.81559754 | 1.05483451 | 0.780583509 |
| 0.861181757 | 1.074415031 | 0.812243327 |
| 0.867592219 | 1.082508022 | 0.896308002 |
| 0.870789908 | 1.100103024 | 0.926782276 |
| 0.896641271 | 1.151929637 | 0.93840313 |
| 0.902799374 | 1.192906191 | 0.99253423 |