Gaussian

|  |  |  |
| --- | --- | --- |
| Lambda | Sigma | Accuracy |
| 100 | 0.1 | 0.5 |
| 10 | 0.1 | 0.5 |
| 1 | 0.1 | 0.5 |
| 0.1 | 0.1 | 0.5 |
| 0.01 | 0.1 | 0.5 |
| 100 | 1 | 0.5 |
| 10 | 1 | 0.5 |
| 1 | 1 | 0.5 |
| 0.1 | 1 | 0.5 |
| 0.01 | 1 | 0.5 |
| 100 | 10 | 0.87 |
| 10 | 10 | 0.96 |
| 1 | 10 | 0.97 |
| 0.1 | 10 | 0.98 |
| 0.01 | 10 | 0.99 |
| 100 | 100 | 0.5 |
| 10 | 100 | 0.92 |
| 1 | 100 | 0.93 |
| 0.1 | 100 | 0.95 |
| 0.01 | 100 | 0.95 |
| 100 | 500 | 0.5 |
| 10 | 500 | 0.61 |
| 1 | 500 | 0.92 |
| 0.1 | 500 | 0.93 |
| 0.01 | 500 | 0.94 |

As we see in the experiments smaller sigma is leading to random outputs (Every prediction is zero). This is most likely due to over fitting as the only in close proximity the influence of point is observed. As the sigma increases overfitting disappears. Sigma increases then also causes decrease in accuracy.

Lambda = 10e-8

|  |  |
| --- | --- |
| Sigma | Accuracy |
| 0.1 | 0.5 |
| 1 | 0.5 |
| 10 | 0.99 |
| 100 | 0.98 |
| 500 | 0.97 |

With lambda so small the accuracy are the best possible. This is similar to what we discussed in class – lambda does not really matter. Overfitting is sometimes good.

Laplace

|  |  |  |
| --- | --- | --- |
| Lambda | Sigma | Accuracy |
| 100 | 0.1 | 0.5 |
| 10 | 0.1 | 0.5 |
| 1 | 0.1 | 0.5 |
| 0.1 | 0.1 | 0.5 |
| 0.01 | 0.1 | 0.5 |
| 100 | 1 | 0.5 |
| 10 | 1 | 0.5 |
| 1 | 1 | 0.5 |
| 0.1 | 1 | 0.54 |
| 0.01 | 1 | 0.55 |
| 100 | 10 | 0.83 |
| 10 | 10 | 0.95 |
| 1 | 10 | 0.98 |
| 0.1 | 10 | 0.98 |
| 0.01 | 10 | 0.98 |
| 100 | 100 | 0.68 |
| 10 | 100 | 0.94 |
| 1 | 100 | 0.96 |
| 0.1 | 100 | 0.97 |
| 0.01 | 100 | 0.98 |
| 100 | 500 | 0.5 |
| 10 | 500 | 0.92 |
| 1 | 500 | 0.93 |
| 0.1 | 500 | 0.95 |
| 0.01 | 500 | 0.95 |

The results for Laplace is also similar. As sigma is small overfitting occurs. Smaller lambda is also leads to higher accuracy. The results are also comparable to gaussian in terms of accuracy. As sigma is increases the accuracy also decreases.

Lambda = 10e-8

|  |  |
| --- | --- |
| Sigma | Accuracy |
| 0.1 | 0.5 |
| 1 | 0.55 |
| 10 | 0.98 |
| 100 | 0.98 |
| 500 | 0.98 |

Similar to Gaussian the accuracy are high with really small lambda. Regularization does not seem to matter much.

3.

gaussian kernel

784 features

0.1 1 10 100

lambda 1.00E-08 0.5 0.5 0.5 0.9

1.00E-02 0.5 0.5 0.5 0.8

1.00E-01 0.5 0.5 0.5 0.89

1 0.5 0.5 0.5 0.86

10 0.5 0.5 0.5 0.67

100 0.5 0.5 0.5 0.5

laplace kernel

sigma 784 features

0.1 1 10 100

lambda 1.00E-08 0.5 0.5 0.5 0.98

1.00E-02 0.5 0.5 0.5 0.98

1.00E-01 0.5 0.5 0.5 0.98

1 0.5 0.5 0.5 0.97

10 0.5 0.5 0.5 0.93

100 0.5 0.5 0.5 0.51

gaussian kernel

sigma 64 features

0.1 1 10 100

lambda 1.00E-08 0.5 0.5 0.5 0.99

1.00E-02 0.5 0.5 0.5 0.99

1.00E-01 0.5 0.5 0.5 0.99

1 0.5 0.5 0.5 0.98

10 0.5 0.5 0.5 0.95

100 0.5 0.5 0.5 0.75

laplace kernel

sigma 64 features

0.1 1 10 100

lambda 1.00E-08 0.5 0.5 0.5 0.99

1.00E-02 0.5 0.5 0.5 0.99

1.00E-01 0.5 0.5 0.5 0.99

1 0.5 0.5 0.5 0.98

10 0.5 0.5 0.5 0.96

100 0.5 0.5 0.5 0.77

The Fourier transform also attends maximum accuracy at same parameters and follows same pattern. However, the accuracy seems more sensitive to parameters as the only good solutions start at sigma =100 everywhere else its 0.