**Part I (First 5k points)**

###### Linear Kernel ######

The linear kernel is:

[[582722.2 582576.94 581815.7 ... 594058.2 594323.7 606581.25]

[582576.94 582476.06 581783.7 ... 594052.94 594230. 606539.3 ]

[581815.7 581783.7 581308.1 ... 593473.94 593386.44 603798.44]

...

[594058.2 594052.94 593473.94 ... 607638.75 607792.06 620125.44]

[594323.7 594230. 593386.44 ... 607792.06 608322.5 622490.8 ]

[606581.25 606539.3 603798.44 ... 620125.44 622490.8 676781.4 ]]

In linear kernel PCA, the number of dimensions required to capture alpha=0.95 is:

r = 3.

The three dominant eigenvalues are 13415.404167, 216.683765 and 179.104588.

The two dominant PCs are:

u1:

[-5.8684702e-07 -5.8710685e-07 -7.3632071e-07 ... -5.9588058e-07

-4.4655312e-07 2.5336162e-06]

u2:

[ 8.2431470e-06 3.7953932e-06 -4.8961165e-06 ... 4.6421255e-06

1.9854502e-05 1.0461365e-05]

In linear kernel, the projected data onto the two dominant PCs is:

[[-39.36394847 8.93078051]

[-39.38137975 4.11200042]

[-49.39020139 -5.30454474]

...

[-39.96989313 5.02936607]

[-29.95345312 21.51074112]

[169.94742869 11.33403897]]

图表, 散点图

描述已自动生成

###### Regular PCA ######

In regular PCA, the number of dimensions required to capture alpha=0.95 is:

r = 3.

In regular PCA, the three largest eigenvalues are: 10528.138672, 1162.571045 and 210.181961.

The two dominant regular PCs are:

[[ 9.9910510e-01 -3.5806447e-02]

[ 1.5196445e-02 -4.1422203e-02]

[ 9.7922410e-04 2.7933549e-02]

[ 3.2187118e-03 -3.3016607e-02]

[ 2.7234582e-03 3.7279088e-02]

[-2.6158164e-03 -5.2810110e-02]

[ 1.8006051e-03 3.5746079e-02]

[ 9.6656405e-04 -5.0526921e-02]

[ 9.6277893e-04 4.0714059e-02]

[ 5.2660092e-04 -5.2762464e-02]

[ 4.8065174e-04 3.2355085e-02]

[ 3.5482997e-04 -7.5896680e-02]

[ 7.4813501e-03 1.2014513e-01]

[-2.8791005e-02 -9.0309447e-01]

[ 7.0487935e-04 4.5260817e-02]

[-2.9905306e-03 -5.7192661e-02]

[ 8.9343463e-04 3.6359467e-02]

[-5.0953915e-03 -7.8599781e-02]

[ 3.5446658e-04 4.1655693e-02]

[-2.2732015e-03 -4.9763251e-02]

[ 5.5620191e-03 1.0102244e-01]

[-2.4651964e-03 1.7509367e-02]

[-2.3576153e-02 -3.4020722e-01]

[ 2.0767075e-03 -5.2530617e-03]

[-1.2528883e-04 -4.2058475e-02]

[ 7.3431490e-04 2.7210914e-02]

[-1.6307642e-03 -6.1703809e-03]]

In regular PCA, the projected data onto the two dominant PCs is:

[[-38.393417 -35.34714 ]

[-38.401295 -35.06396 ]

[-48.385708 -33.792137]

...

[174.53008 53.118015]

[324.41605 47.573956]

[334.34076 46.71564 ]]

图表, 散点图

描述已自动生成

In linear kernel, the largest eigenvalue is much bigger than the second largest eigenvalue, compared with that in regular PCA.

###### Gaussian Kernel ######

The spread is sigma^2 = 10000:

The Gaussian kernel is:

[[1. 0.9977853 0.98025376 ... 0.89384145 0.88703656 0.09856404]

[0.9977853 1. 0.98922247 ... 0.9044368 0.8896513 0.09936648]

[0.98025376 0.98922247 1. ... 0.9048824 0.86685395 0.08008729]

...

[0.89384145 0.9044368 0.9048824 ... 1. 0.9813188 0.10986938]

[0.88703656 0.8896513 0.86685395 ... 0.9813188 1. 0.13451017]

[0.09856404 0.09936648 0.08008729 ... 0.10986938 0.13451017 1. ]]

The number of dimensions required to capture alpha=0.95 is:

r = 22.

In Gauss Kernel, the three largest eigenvalues are: 0.136403, 0.061578 and 0.020320.

In Gaussian kernel PCA, the two dominant PCs are:

u1:

[ 0.00018651 0.00020001 0.00024571 ... 0.00027318 0.00019724

-0.00147416]

u2:

[ 1.5130757e-04 1.5924961e-04 -8.8212531e-05 ... 1.0185257e-04

3.4680447e-04 9.3547751e-05]

In Gaussian Kernel, the projected data onto the two dominant PCs is:

[[ 0.12720171 0.04658604]

[ 0.13640822 0.04903131]

[ 0.16757728 -0.02715973]

...

[ 0.1863139 0.03135936]

[ 0.13452344 0.10677753]

[-1.00539452 0.02880239]]

图表

描述已自动生成

Sigma^2 = 10000

###### Gaussian Kernel ######

The spread is sigma^2 = 100000000:

The Gaussian kernel is:

[[1. 0.99999976 0.99999803 ... 0.9999888 0.999988 0.9997683 ]

[0.99999976 1. 0.9999989 ... 0.9999899 0.9999883 0.99976915]

[0.99999803 0.9999989 1. ... 0.99999 0.9999857 0.9997476 ]

...

[0.9999888 0.9999899 0.99999 ... 1. 0.9999981 0.99977916]

[0.999988 0.9999883 0.9999857 ... 0.9999981 1. 0.99979943]

[0.9997683 0.99976915 0.9997476 ... 0.99977916 0.99979943 1. ]]

The number of dimensions required to capture alpha=0.95 is:

r = 3.

In Gauss Kernel, the three largest eigenvalues are: 0.000134, 0.000002 and 0.000002.

In Gaussian kernel PCA, the two dominant PCs are:

u1:

[-0.00587218 -0.0058748 -0.0073672 ... -0.00596278 -0.00446908

0.02535421]

u2:

[ 0.08242832 0.03795722 -0.04899238 ... 0.04643349 0.19856478

0.10456307]

In Gaussian Kernel, the projected data onto the two dominant PCs is:

[[-0.00393515 0.00089292]

[-0.0039369 0.00041118]

[-0.00493701 -0.00053072]

...

[-0.00399586 0.000503 ]

[-0.00299488 0.002151 ]

[ 0.01699072 0.0011327 ]]

图表, 散点图

描述已自动生成

Sigma^2 = 100000000

**Part II**

###### train dataset ######

The Q matrix of the training data is:

[[ 1.00000000e+00 2.54480961e+01 -1.59256308e+00 ... 2.66600479e+01

-5.87841248e-02 -1.17638520e+01]

[ 1.00000000e+00 2.54480961e+01 -1.59256308e+00 ... 2.32128550e+01

-4.68891711e-02 -6.40291433e+00]

[ 1.00000000e+00 2.54480961e+01 -1.59256308e+00 ... 1.93908269e+01

-2.17593153e-02 3.63765989e+00]

...

[ 1.00000000e+00 -4.55190387e+00 1.50408200e+00 ... -3.88252889e+00

1.94021102e-01 8.23436869e+00]

[ 1.00000000e+00 -4.55190387e+00 1.50408200e+00 ... -6.01708014e+00

2.08188945e-01 2.60878298e-01]

[ 1.00000000e+00 -4.55190387e+00 1.50408200e+00 ... -4.42006388e+00

1.91006864e-01 -1.15889626e+01]]

The weight vectors are:

[-8.7043900e+01 2.1377640e+00 -1.2654609e+01 1.2472290e+01

-9.5476685e+00 -1.2751801e+01 3.1120508e+01 8.5772743e+00

-3.1249447e+00 -2.7383122e+00 -1.0130277e+00 5.6215283e-02

7.4349642e+00 2.6952463e-01 2.0259199e+00 -4.7592708e-01

8.2165165e+00 -5.2404060e+00 -1.5456933e+01 -2.0170922e+00

-1.2177909e+01 3.7924513e-01 -1.0511665e+00 1.5626988e+00

2.0598698e-01 7.5872192e+00 -3.7998717e-02]

The L2 norm of the weight vectors of the training data is:

99.08716

The predicted response of train dataset is:

[158.52832 144.81644 141.50781 ... 221.24728 214.95653 202.21909]

The SSE on the training data is:

130623220.74625452

The MSE on the training data is:

9455.857879416137

The R^2 on the training data is:

0.1718213143498506

###### test dataset ######

The predicted response of test dataset is:

[181.70778 173.64166 169.60117 ... 137.46739 126.258286 109.41677 ]

The SSE on the testing data is:

45469046.04874517

The MSE on the testing data is:

7679.284926320752

The R^2 on the testing data is:

0.08427943715438296