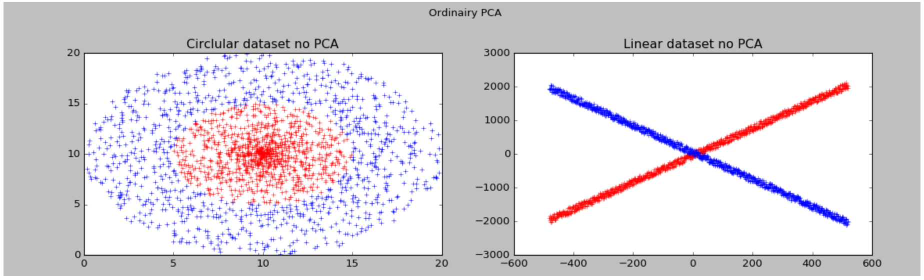
1 Data Generation

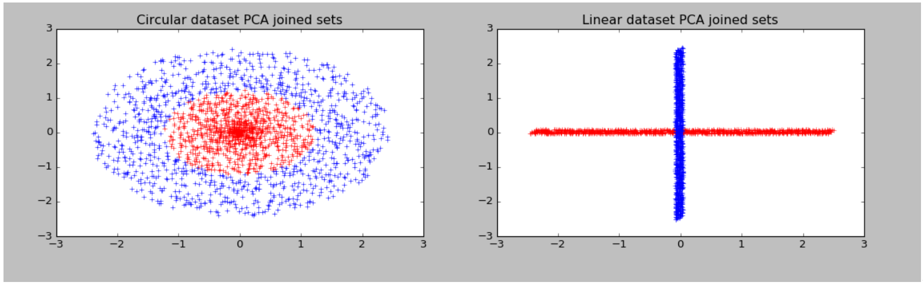
I decided to implent data generator myself (code can be found in the zip archive attached). The basic idea for circular and ring data set was generating random radius (within the range) and angle theta (within 2 π range).

The sets looked like this:



2 Applying basic PCA

Next basic Principal Components Analysis algorithm was applied. After this transformation, the data sets looked like this:



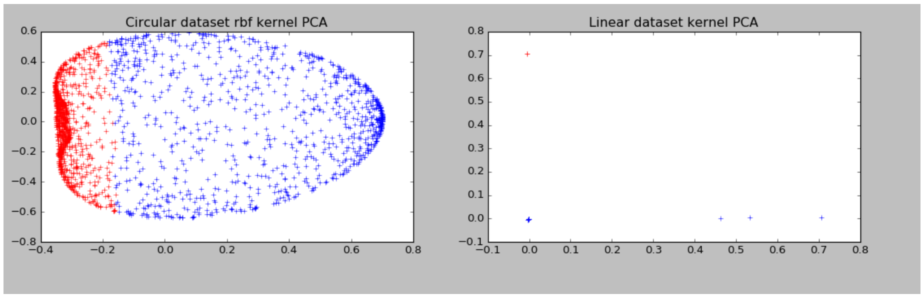
.

As the plot shows PCA managed to correctly identify most important dimensions in case of linear data, But failed to make much of a difference with non-linear data.

3 Applying PCA with kernel trick

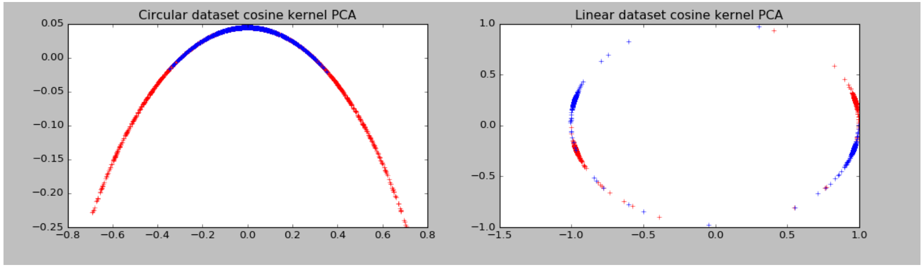
Fortunately there is a way for PCA to perform well on non-linear data: kernel tricks. Kernel tricks is a method of replacing the basic PCA'a dot product with another function. The problem was the kernel tricks that I used require a parameter called gamma. It had to be found by trial and error (otherwise I would need to implement i method like SVM to evalute how well did PCA with specific value of gamma).

After extensive search for correct gamma parameter I settled down with value 0.1, which gave this results:



Obviously the set on the left is easily separable by a vertival line ($x = -0.2$). It also seems that the linear set was sort of separated, but most probably red crosses are overlapping with blue ones.

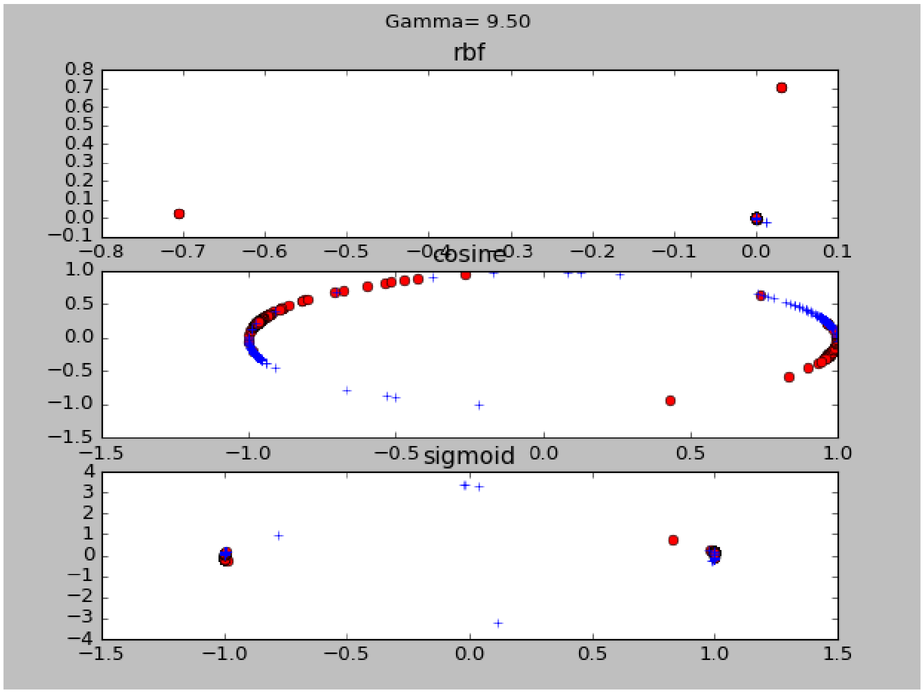
Cosine kernel was also pretty successful for the circular data set, but didn't work well for linear dataset:

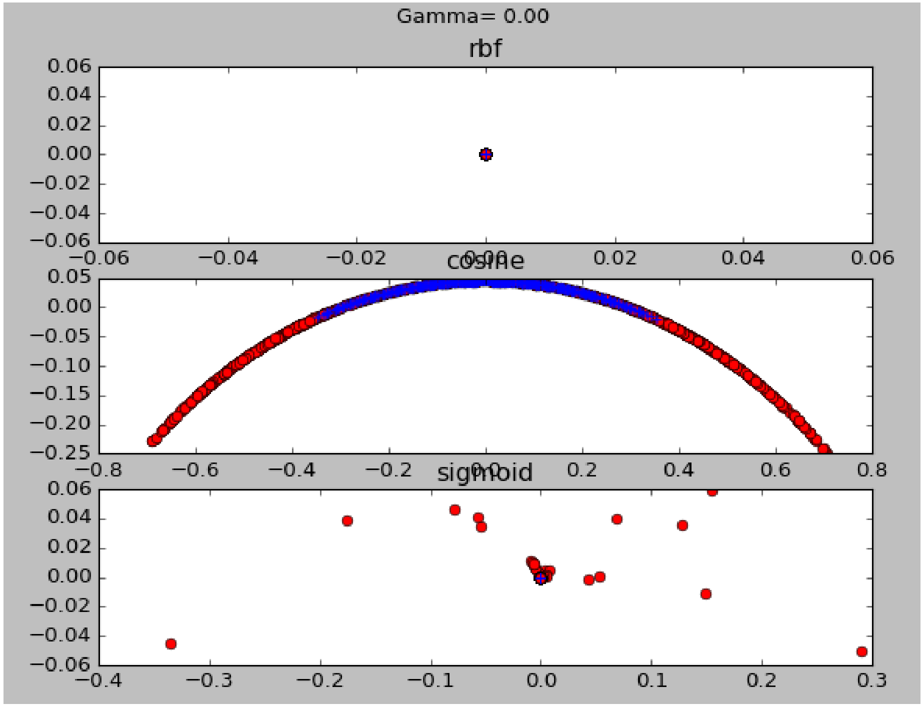


Unfortunately despite testing many values for sigmoid kernel I wasn't able to find any value, with would separate data sets properly.

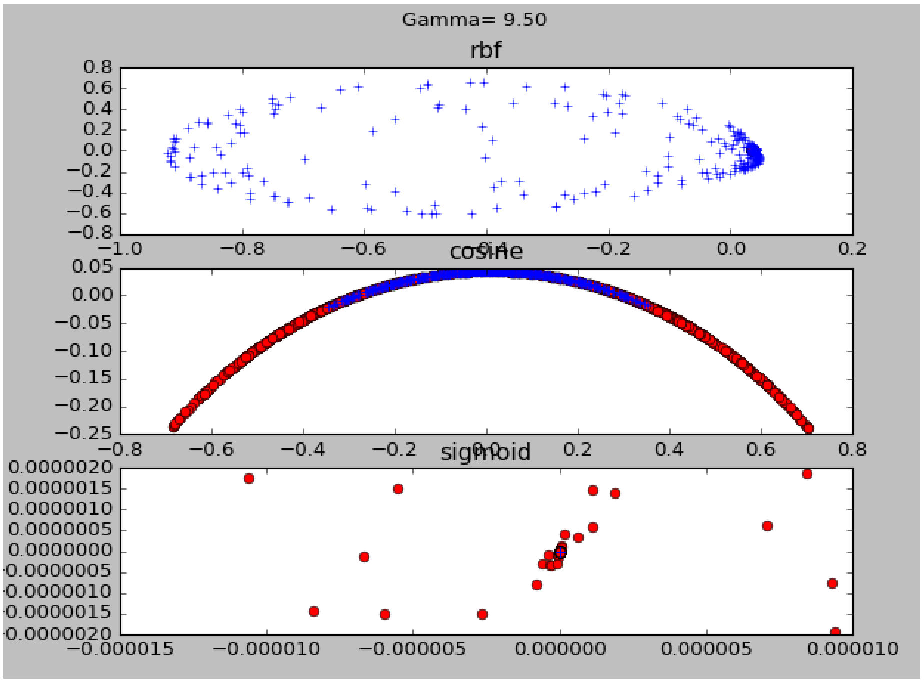
These are plots for some of the values I tried:

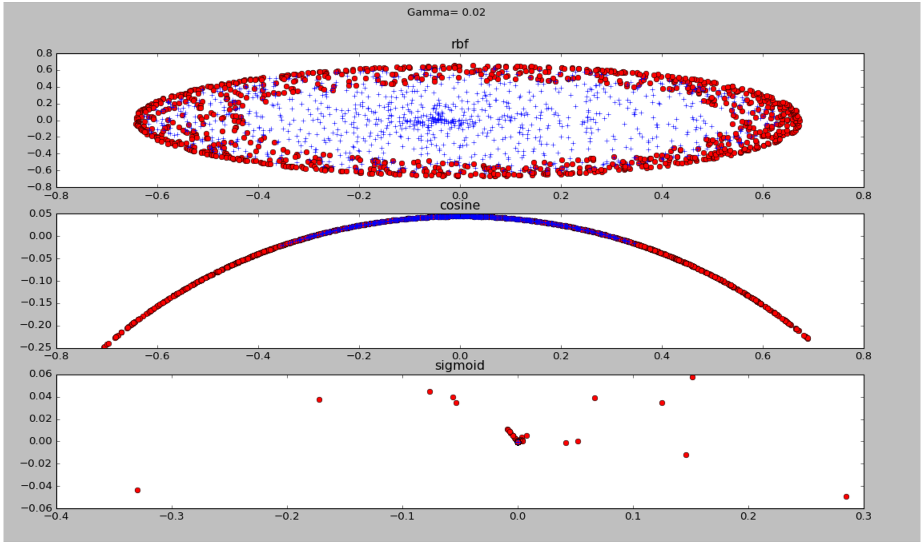
3.1 Linear:


Gamma= 9.50
rbf
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0.0
−0.1
−0.8
−0.7
−0.6
−0.5
−0.4
−0.3
−0.2
−0.1
0.0
0.1
cosine
1.0
0.5
0.0
−0.5
−1.0
−1.5
−1.5
−1.0
−0.5
0.0
0.5
1.0
sigmoid
4
3
2
1
0
−1
−2
−3
−4
−1.5
−1.0
−0.5
0.0
0.5
1.0
1.5


Gamma= 0.00
rbf
0.06
0.04
0.02
0.00
−0.02
−0.04
−0.06
−0.06
−0.04
−0.02
0.00
0.02
0.04
0.06
cosine
0.05
0.00
−0.05
−0.10
−0.15
−0.20
−0.25
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8
sigmoid
0.06
0.04
0.02
0.00
−0.02
−0.04
−0.06
−0.4
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3

3.2 Circular:


Gamma= 9.50
rbf
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−1.0
−0.8
−0.6
−0.4
−0.2
0.0
0.2
cosine
0.05
0.00
−0.05
−0.10
−0.15
−0.20
−0.25
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8
sigmoid
0.0000020
0.0000015
0.0000010
0.0000005
0.0000000
0.0000005
0.0000010
0.0000015
0.0000020
−0.000015
−0.000010
−0.000005
0.000000
0.000005
0.000010


Gamma= 0.02
rbf
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
−0.8
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8
cosine
0.05
0.00
−0.05
−0.10
−0.15
−0.20
−0.25
−0.8
−0.6
−0.4
−0.2
0.0
0.2
0.4
0.6
0.8
sigmoid
0.06
0.04
0.02
0.00
−0.02
−0.04
−0.06
−0.4
−0.3
−0.2
−0.1
0.0
0.1
0.2
0.3