

Assignment 2

CSCI-6390

Yiqing(Alice) Guo

September 29, 2017

1 Part 1

linear kernel project point:

$[[-38292.4570163 \ 2101.47455625]$

$[-38413.30362699 \ 2978.28952548]$

$[-12268.52727985 \ -28345.6849681 \]$

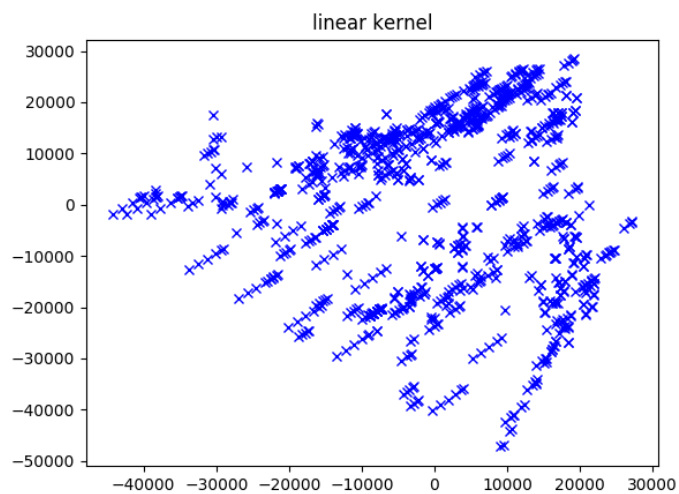
...,

$[\ 18548.88051646 \ -8111.42717031]$

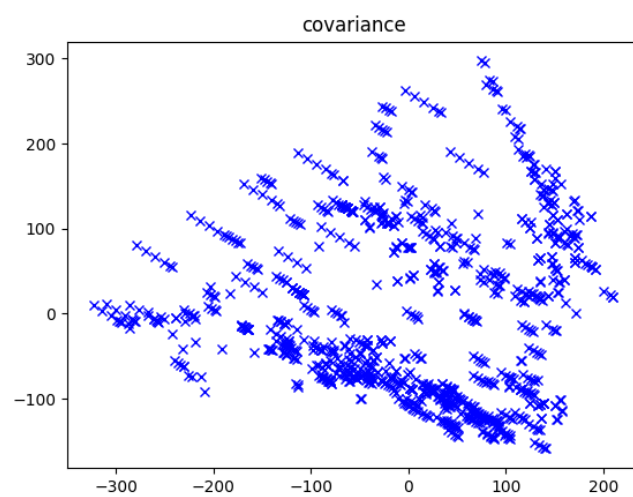
$[\ 17020.38035273 \ -15121.00127116]$

$[\ 3476.71591872 \ -7708.10369414]]$

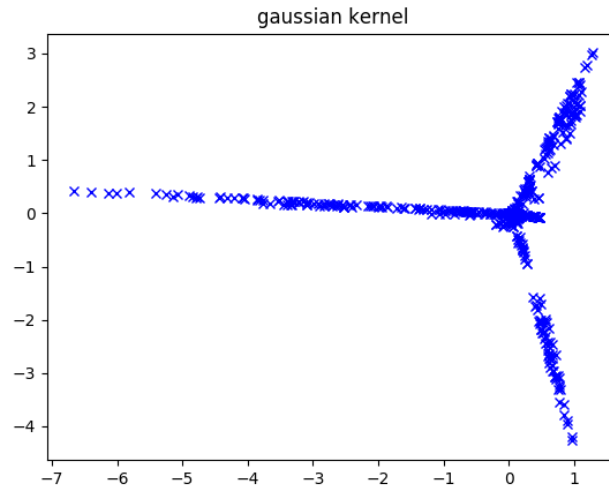
5 dimensions are required to capture 95% of the total variance



PCA:



Gaussian kernel(with $\sigma_2 = 2000$):



How do the eigenvalues and the projection compare with that obtained via Kernel PCA with linear kernel?

The linear PCA and PCA have the same shape of plot but different derivation and scale.

2 Part 2

$N = 100,000$, $d = 10$

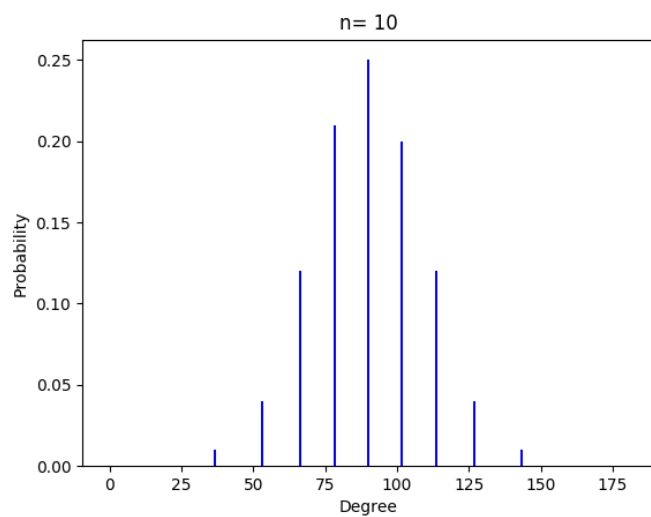
Max: 180.0

Min: 0.0

Range value: 180.0

Mean: 89.9519671

Variance: 373.806371612



$N = 100,000, d = 100$

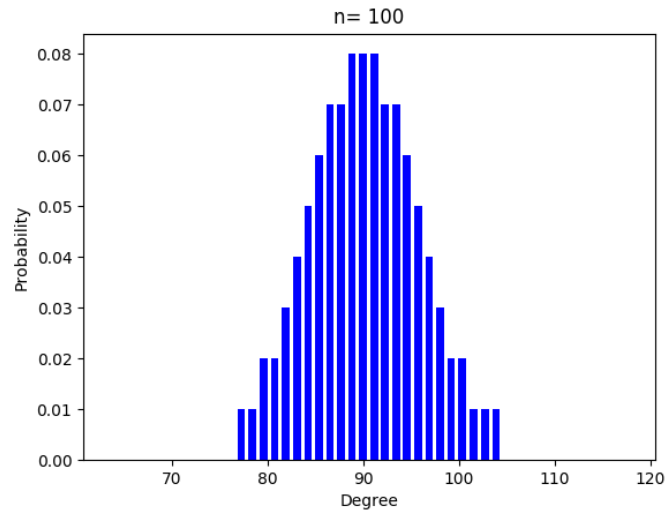
Max: 117.39

Min: 63.9

Range value: 53.49

Mean: 89.9812138

Variance: 33.2748129087



$N = 100,000$, $d = 1000$

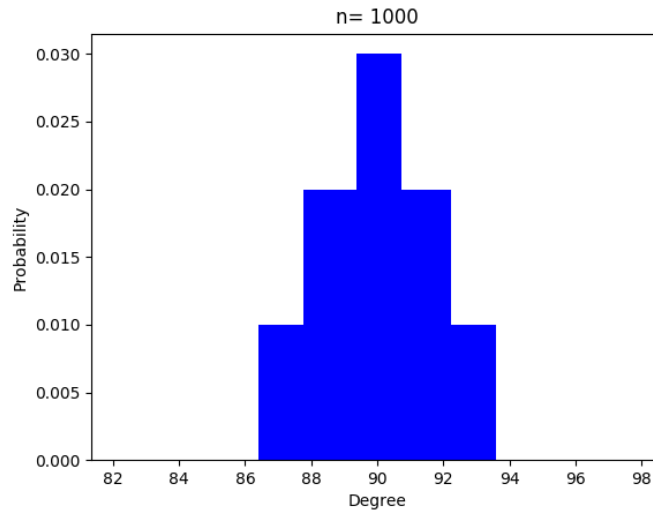
Max: 97.47

Min: 82.53

Range value: 14.94

Mean: 90.0038613

Variance: 3.26839886336



Extra Question:

As $d \rightarrow \infty$: we will have $\cos(\theta) = \frac{v_1 v_2}{d} = \frac{v_1 v_2}{\infty}$

we will have $\cos(\theta) = 0$ for $d \rightarrow \infty$ which is 90° , we have the EPMF trend as all points are close to the 90° degree. we will certainly have mean = 90.