**Part I: Principal Components Analysis**

The number of dimensions required to capture alpha=0.975 is: r = 6.

The three largest eigen values are: 10528.138672, 1162.571045, and 210.181961.

The three PCs are:

报纸上的文字

描述已自动生成

The mean squared error is mse = 603.407955.

The reduced dimensionality dataset A on the first three PCs is:

[[-38.393417 -35.34714 12.594683 ]

[-38.401295 -35.06396 7.200858 ]

[-48.385708 -33.792137 -2.8894534]

...

[174.53008 53.118015 -4.529205 ]

[324.41605 47.573956 18.068136 ]

[334.34076 46.71564 -9.731801 ]]

The 3D-plot is:

地图上有字

描述已自动生成

**Part II: Diagonals in High Dimensions (50 points)**

If d=10, min=0.000001, max=180.000000, value range=180.000000, mean=90.039772, and variance=375.509888.

If d=100, min=62.612892, max=121.332253, value range=58.719360, mean=90.012573, and variance=33.171886.

If d=1000, min=82.761490, max=97.932129, value range=15.170639, mean=90.011032, and variance=3.273134.

The plots of 3 PMFs are as follows:

手机屏幕截图

描述已自动生成

手机屏幕截图

描述已自动生成

手机屏幕截图

描述已自动生成

The formulas for what should happen to angle between two half-diagonals and as  is:

The formula conforms to what we observe in the above pmf plots. This is reasonable because as d grows very large, the chances that the -1 and 1 cancel out grows larger and larger as well; therefore, the dot products of is likely to be much smaller than d, which renders as for most of the angles.

**Reference**

<https://www.geeksforgeeks.org/matplotlib-pyplot-title-in-python/>

https://stackoverflow.com/questions/17901218/numpy-argsort-what-is-it-doing