Introduction to Machine Learning (67577)

Exercise I – Mathematical Background

Hadar Sharvit – 208287599

***PART I***

**Question 1 – Linear Algebra**

1. Since we have
2. In the first direction - let be two non-zero vectors with angle between them. This means that , i.e.

In the second direction – if , and since the norms it must be that

1. Denoting where are orthonormal matrices and is a diagonal matrix with values we can write

Since is orthogonal then and we get

This is useful because it only takes two calculation of transpose in order to find

* To find the eigenvalues we calculate :

Therefore, we can write

* Now for the eigenvectors:

For we have

So, we can conclude and after normalization (

For we have

So, we can conclude and after normalization:

This gives us

* Finally, since then :

Let’s check everything:

1. , where and are the eigenvectors and eigenvalues of respectively.

Iteratively expanding :

Using the hint for :

Using ’s *EVD* we have:

Where is a diagonal matrix in which are the eigenvalues of . From here, as we’ve seen in class

This can be substituted to in summation form as followed

Since then so we can write

this is because so and the sign depends on the value of

**Question 2 – Multivariable Calculus**

Which can be written as where represents the column.

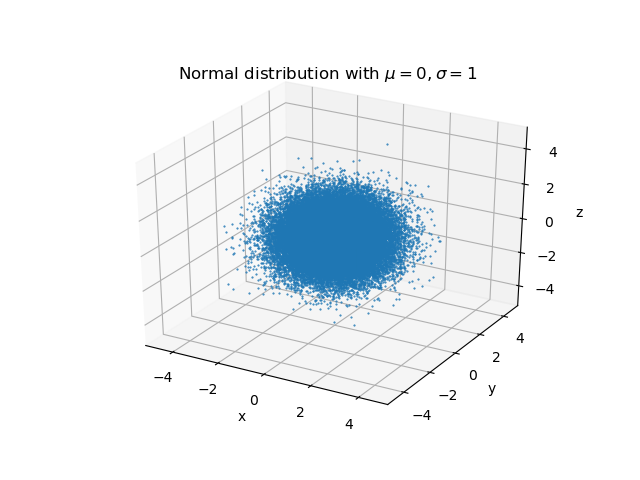
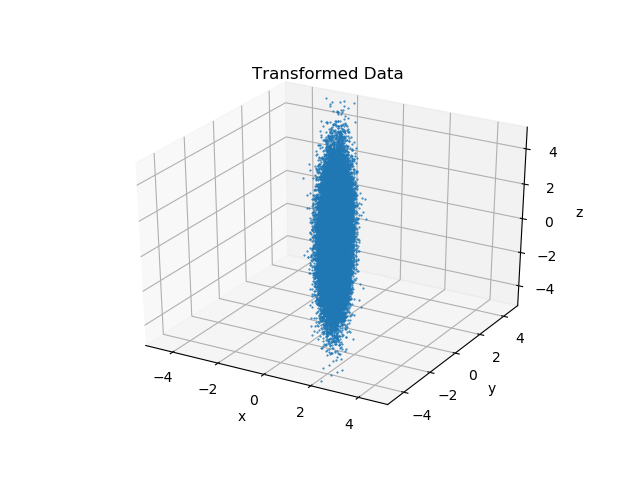
From here we can calculate

The is added since only exists where , and otherwise . From here

We get in matrix notation

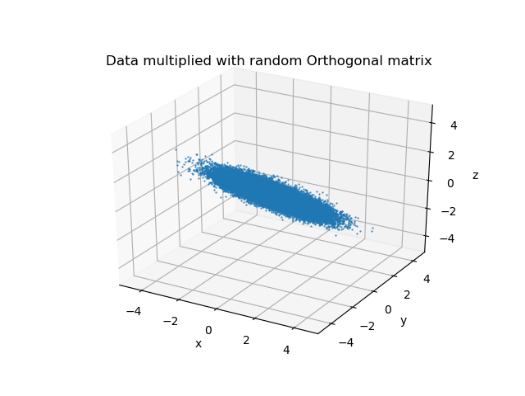
1. ,

***PART II***

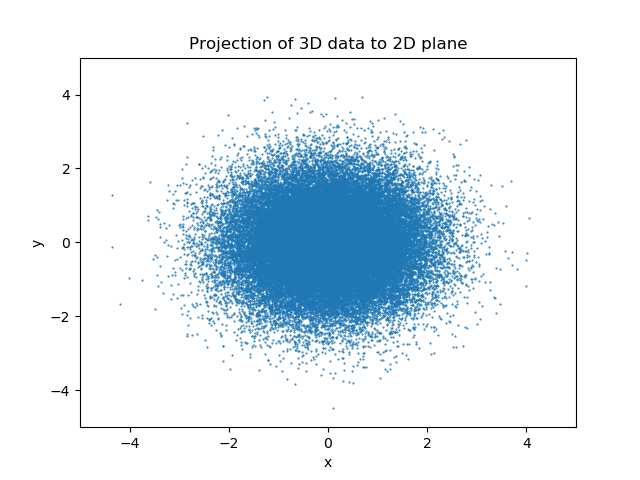
1. 
2. 

The covariance matrix is

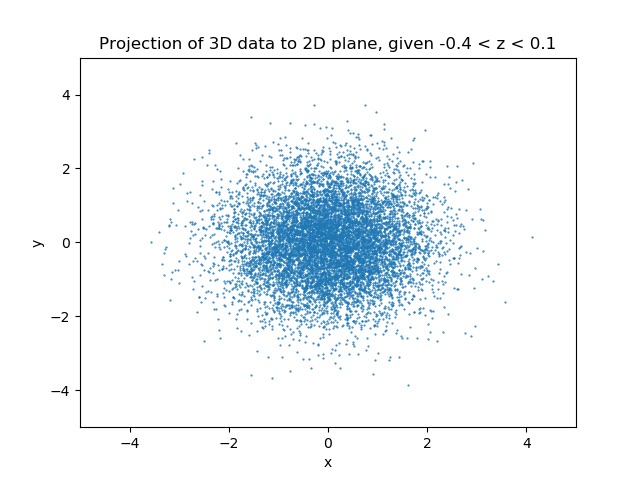
Which is approximately (up to values that are close to zero outside the diagonal line). this was expected because the elements on the diagonal are the variance of the new data, and the variance of the new data (call it is .

1. 

After multiplication with a random unitary matrix, the covariance matrix isn’t diagonal anymore, and contains random values outside the diagonal line.

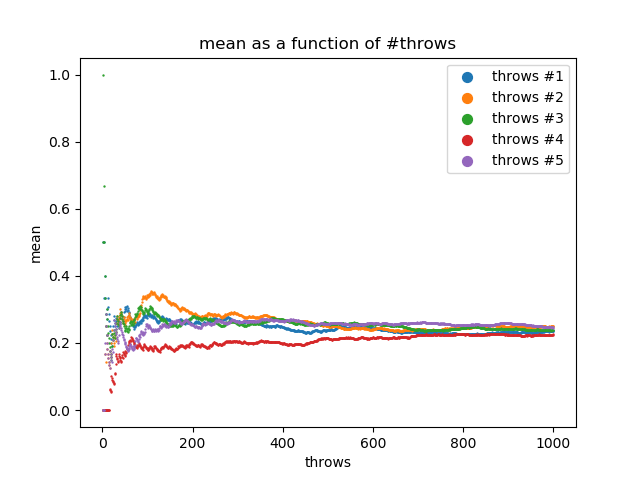
1. 

We can see that the projection yields a 2D gaussian distribution, as expected.

1. 

We can see that the projection remains a 2D gaussian (with less data points, due to the condition )

1. we expect that the as a function of will approach constant value . this is due to the fact that a coin toss follows a binomial distribution, and in our case , therefore



B&C. the results for :

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |

For large values of : The *Percentage* plot has small values

For medium values of : The *Percentage* plot has significant values where is small, and vice versa

For small values of : The *Percentage* plot has significant values . This is due to the fact values very close to the expectancy are scarce.