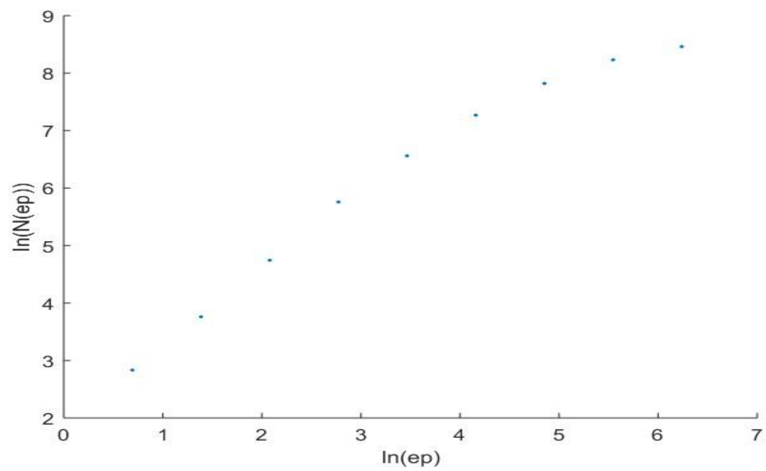
Dmitriy Tarasov
CSCI 4446

Assignment 10: Box Dimension

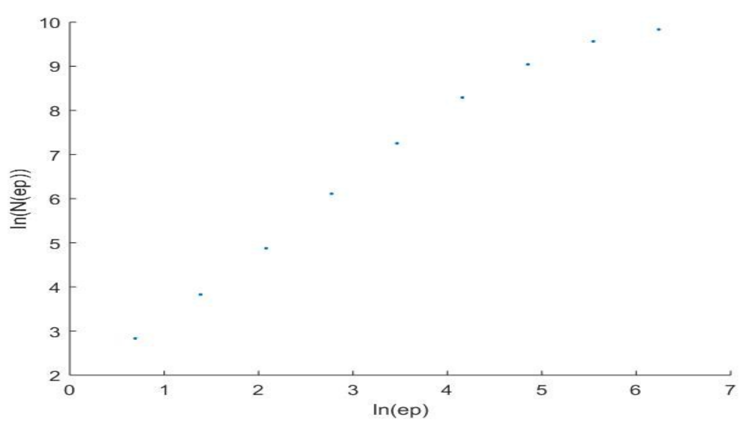
The Lorenz system was initialized with the parameters r=45,b=4, and a=16 and initial conditions of x0=(-20,-20,50). The Runge-Kutta four was run with a fixed time-step of h=0.001 with 5000 time-steps.

1) The data set of the Lorenz attractor is generated and then used to generate a single titght fitting box by finding the min and max points along each state variable. Then the parameter side was used to split each side of the box into 2^n partitions in order to save on computation. Then it counts the number of the boxes that have an element of the Lorenz attractor. The launcher iterates the number of boxes power $n \in [0,9]$ as n=10 causes memory issues.
2) a) The plot of with the ep being the number of cubes and N(ep) being the number of marked cubes for the box counting dimension. The plot of the $\ln(n(\epsilon))$ vs $\ln(1/\epsilon)$ is the following.

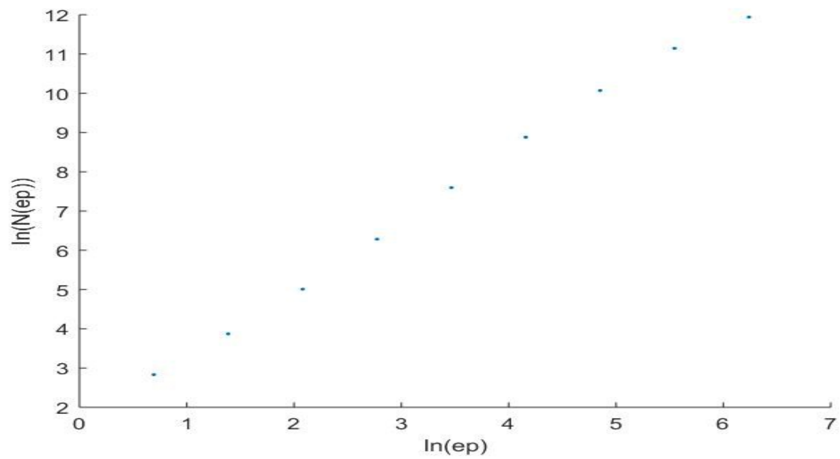


The result is expected due to the balance of having too few and too many partitions. Having too few partitions would result in each large cube being marked and thus The result of having too many partitions is also expected as it would hit the limit of data which in this case was 5000 data points would cause more cubes to be generated that wouldn't have any data due to the cubes being smaller than the spacing of the data.

b) The result of increasing the number of data points in the Lorenz trajectory to 20,000 is the following.

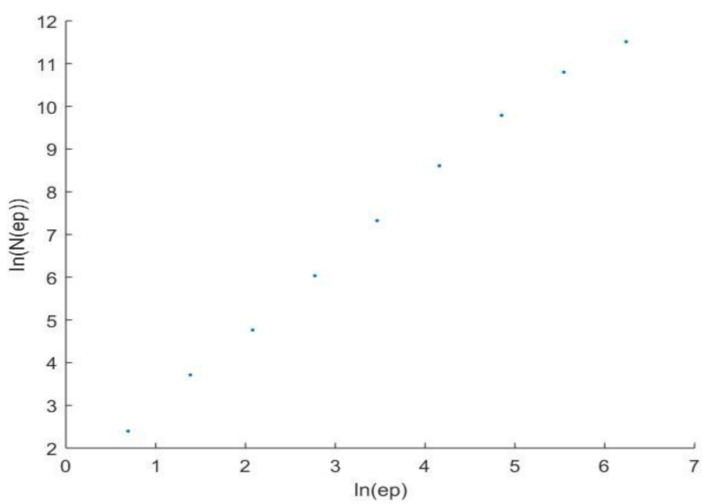


Considering that the results are similar means that there isn't a pathology caused by the data. Considering that Eckmann had a theoretical limit of $2*\log(N)/\log(r/d)$ where r is the resolution, d is the diameter of attractor, and N is amount of data. Though there is a theoretical issues as the box dimension is only at the limit of number of cubes going to infinity. Thus the second result is more reliable. The box dimension is measured by taking the slope of the 4th and 5th entry. The box dimension of the system was $c_p = 1.6496$. This is an issue as the Lorenz attractor should have a dimension between a two and three. Taking 1,500,000 data points shows that the amount of data previously was actually insufficient as the upper asymptote is reached.



Which has the capacity dimension of $c_p = 1.9339$.

c) considering that previously the embedded system had the embedding parameters of T=110 and m=7. The plot of the embedded $\ln(n(\epsilon))$ vs $\ln(1/\epsilon)$ for is the following



which has the box dimension of $c_p = 1.8577$ at the 4th and 5th indexes. The amount of embedded data falls into the same problem of having insufficient data after the 150,000 time series in embedded in the three state spaces.

3) Considering the memory issue with the boxes for an n by n by n array could be reduced by only storing the boxes that actually had a point of the Lorenz system on the region. The memory requirement would be reduced from $\Theta(P^3)$ where P is the partition of a side to $\Theta(H)$ where H is the number of data points in the system which would be a lot better for large number of cubes.