Assignment 4

Neev Shaw

September 25, 2025

Exercise 4

Part (a)

We are only using 10% of the space in one dimension, so the fraction of the available observations will be used to make a prediction is $0.1 = \frac{1}{10}$.

Part (b)

We are only using 10% of the space in each dimension, so the fraction of the available observations will be used to make a prediction is $0.1 \cdot 0.1 = (0.1)^2 = \frac{1}{100}$.

Part (c)

Again, we are only using 10% of the space in each dimension, so the fraction of the available observations will be used to make a prediction is $(0.1)^{100} = \frac{1}{10^{100}}$.

Part (d)

As we can see, as the number of features and dimensions increase in the KNN model, a smaller fraction of the total space is used to make a prediction - and this fraction decreases exponentially. Therefore with a large p, there are very few training observations that are near X which is a big drawback.

Part (e)

We need the prediction hypervolume to be 10% of the total hypervolume which is 1. If the side length of the prediction hypercube is s, then we see that $s^p = 0.1$ so $s = \sqrt[p]{0.1}$

For $p = 1$, $s = 0.1$

For $p = 2$, $s \approx 0.32$

For $p = 100$, $s \approx 0.98$

This shows that as p increases, s approaches 1, meaning that to even search 10% of the overall you would have to search almost (but not quite) 100% of the feature space for each dimension.