Homework 9
Colorado CSCI 5454

Sai Siddhi Akhilesh Appala

November 30, 2022

People I studied with for this homework: Gautam

Problem 1

Given information:
$(\frac{1}{\sqrt{2}}, 0, ..., 0), (0, \frac{1}{\sqrt{2}}, 0, ..., 0)$...have placed m points in just m dimensions, and we need to change them to the d dimensional space and estimate the big-O notation.
Solution:
lemma 1 from notes :

When given a collection of points $1, ..., x_n \in R^m$ and let $\epsilon, \delta \in (0, 1]$ and when $d \geq \frac{8}{\epsilon^2} ln(\frac{n}{\delta})$. Then for the Gaussian projection that projects the points of dimension m into d, with probability at least $1 - \delta$, for all i, j,

$$(1 - \epsilon)||x_i - x_j|| \leq ||y_i - y_j|| \leq (1 + \epsilon)||x_i - x_j||$$

- if we consider the same for our case, As given number of points and m dimensions in the space i.e m, i.e $n = m$.
- If we project that on to a dimension d which is given by the above formula, where $\epsilon = 0.1$.
- If we consider the error as $\delta = 0.0001$, points will be lying between $1 - \epsilon$ and $1 + \epsilon$ as $||x_i - x_j|| = 1$ for all points in m dimensions.
- As given 0.9 and 1.1 i.e 1 - ϵ and 1 + ϵ respectively, substituting in above equation we get $\epsilon = 0.1$.
- Lets take error rate $\delta = 0.0002$ and substitute in d.

$$d \geq \frac{8}{(0.1)^2} \ln\left(\frac{m}{0.0002}\right)$$
$$d = O(ln(m))$$

From above $O(ln(m))$
we can conclude $d = O(ln(m))$.

Problem 2

Part a

lemma 1 from notes :

When given a collection of points $1, ..., x_n \in R^m$ and let $\epsilon, \delta \in (0, 1]$ and when $d \geq \frac{8}{\epsilon^2} ln(\frac{n}{\delta})$. Then for the Gaussian projection that projects the points of dimension m into d, with probability at least $1 - \delta$, for all i, j,

$$(1 - \epsilon)||x_i - x_j|| \leq ||y_i - y_j|| \leq (1 + \epsilon)||x_i - x_j||$$

- Let say vector d_1 pointing to the origin of the original dimension R^m. then the $x_j = 0$.
- If this point is transformed to required space then the resultant vector of new space will be the origin of new space with $y_j = 0$.
- Now d_1 vector pointing to origin of original dimension R^m.
- Now we say $\epsilon = 0.1$ and substituting in the above equation we get.

$$0.9||x_i|| \leq ||y_i|| \leq 1.1||x_i||$$

From above, we can say Johnson–Lindenstrauss's lemma is proved.

Part b

Given information
if $x_i.x_j = 0$ then we need to prove the below.
To be proved
-0.1 $\leq y_i.y_j \leq 0.1$
Solution
As we know y_i and y_j are obtained using Johnson–Lindenstrauss transform.

$$(1 - \epsilon)||x_i - x_j|| \leq ||y_i - y_j|| \leq (1 + \epsilon)||x_i - x_j|| \tag{1}$$

If we square on both sides we get the below equation

$$(1-\epsilon)^2||x_i - x_j||^2 \leq ||y_i - y_j||^2 \leq (1+\epsilon)^2||x_i - x_j||^2 \tag{2}$$

if we assume that -x_j then we get the transformed one as $-y_j$ then Johnson–Lindenstrauss becomes

$$(1-\epsilon)||x_i + x_j|| \leq ||y_i + y_j|| \leq (1+\epsilon)||x_i + x_j|| \tag{3}$$

If we square on both sides we get the below equation

$$(1-\epsilon)^2||x_i + x_j||^2 \leq ||y_i + y_j||^2 \leq (1+\epsilon)^2||x_i + x_j||^2 \tag{4}$$

If we subtract the equations with positive x_j and -x_j then we get the below equation

$$(1-\epsilon)^2||x_i+x_j||^2-(1+\epsilon)^2||x_i-x_j||^2 \leq ||y_i+y_j||^2-||y_i-y_j||^2 \leq (1+\epsilon)^2||x_i+x_j||^2-(1-\epsilon)^2||x_i-x_j||^2 \tag{5}$$

$$||x_i - x_j||^2 = ||x_i||^2 + ||x_j||^2 - 2x_i.x_j \tag{6}$$

$$\because ||x_i|| = 1 and x_i.x_j = 0 \tag{7}$$

$$||x_i - x_j||^2 = 2 \tag{8}$$

$$||x_i + x_j||^2 = 2 - -with other one \tag{9}$$

$$||y_i - y_j||^2 = ||y_i||^2 + ||y_j||^2 - 2y_i.y_j$$
$$||y_i + y_j||^2 = ||y_i||^2 + ||y_j||^2 + 2y_i.y_j$$

subtracting above 2

$$||y_i + y_j||^2 - ||y_i - y_j||^2 = 4y_i.y_j \tag{10}$$

By substituting the above equations, we tend to get the

$$-4 * 0.1 \leq 4y_i.y_j \leq 4 * 0.1$$

$$-0.1 \leq y_i.y_j \leq 0.1$$

Part c

- Consider these vectors which have n dimensions, which lie in the unit sphere and are orthogonal (1,0,0....)(0,1,0,0....), (0,0,1........) and so on till (0,0,0,.......,1)
- Now from above, these vectors lie on the same unit sphere and are orthogonal.
- Lets consider x_i and y_i be i^{th} dimension of vector x and y and our vectors have one dimension then it will work like below
- If $x_i = 0$ then $y_i = 1$ or if $x_i = 1$ then $y_i = 0$ or $x_i = 0$ and $y_i = 0$.
- For any case which is satisfying the $x_i.y_i = 0$ for all i. So we can say that vectors are orthogonal.

Part d

We can combine the above 3 parts to prove this.

- The goal is to get $n = 2^{\Omega(k)}$ points of dimension k.
- By using **Part C** we create n points in dimensional space and get all the required vectors.
- Next step is to use **Part a** to achieve the projections to transform n points to k dimensions. then we get $0.9 \leq ||y_i|| \leq 1.1$.
- Now by using **Part B** as we know if $||x_i|| = 1$ are orthogonal then projections ($y_i's$ are orthogonal.
- , Therefore, we can conclude that $n = 2^{\Omega(k)}$ points in K dimensional space are the unit sphere and are orthogonal.

Problem 3

We know that $A = UDV^T$.
A $= n * d$.
U = Dimensions of $n * r$
D = dimension of $r * r$
V = Dimension of $d * r$.

$$\implies \sum_{f=0}^{f=r} U_{i,f} * D_{f,f} * V_{j,f}$$

$$\implies \sum_{f=0}^{f=r} U_{i,f} * \sigma_f * V_{j,f}$$

$$\implies \sum_{f=0}^{f=r} \sigma_f * U_{i,f} * V_{j,f}$$

Consider the matrix $H = \sum_{l=1}^{r} \sigma_l u_l v_l^T$.
Now Lets consider the (i,j)th term of H.

$$\begin{aligned} H_{i,j} &= \sum_{l=1}^{r} (\sigma_l u_l v_l^T)_{(i,j)} \\ &= \sum_{l=1}^{r} \sigma_l u_{l_{(i,1)}} {v_l^T}_{(1,j)} \\ &= \sum_{l=1}^{r} \sigma_l * u_{l_{(i,1)}} * v_{l_{(j,1)}} \\ &= \sum_{l=1}^{r} \sigma_l * U_{i,l} * V_{j,l} \\ &= \sum_{f=1}^{r} \sigma_l * U_{i,f} * V_{j,f} \end{aligned} \tag{11}$$

Therefore $A_{i,j} = H_{i,j}$, which means $A = \sum_{l=1}^{r} \sigma_l u_l v_l^T$.

Problem 4

Part a

- Consider matrix A, where rows as Types of insects and columns as locations of insects.
- The ith and Jth entry represents the number from 1 and 100 describing how many types of insects live in location j.
- **First right singular vector** - v_1, which represents the idealized location feature, probably could be any one of humidity/rainfall/temperature that best represents a one-dimensional model. The next right vector would be a feature completely orthogonal to the first. If we were to consider a one-feature model, we would just consider this vector and take components along this to get the feature.
- **First left singular vector** - u_1, which represents the number of insects of different types that prefer that feature described above say example rainfall. This can help us classify insects that prefer particular weather.

- **First singular value** - σ_1, represents how important a feature is or how better it explains data.

Part b

By using the "collaborative filtering" method to predict the insect type at a location. As described above each $A(i, j)$ represents how many insects of type i live at location j. Below are the steps to predict the missing values taken from pre-read notes.

1. Given the matrix A first uses a simple method to estimate the missing entries. In every row of Matrix A replace the missing element with the average of that row, and let's call this new matrix A'.

2. Call the full matrix with all entries estimated A‘.

3. Compute the SVD $A' = UDV^T$.

4. Compute the low-rank approximation $A'_k = U_k D_k V_k{}^T$.

5. Use the value of $A'_k(i, j)$ to estimate the missing entry i,j.