

COMS 4995: Network and Crowd HW1

Zhuangyu Ren(zr2209)

September 17, 2019

Exercise 1

1. The best lower bound is 8. This means that the longest path is the diameter, and all other people are on this path.

Graph is like: x-x-x-x-x-x-x-Jane

The best upper bound is 16. This means that there are more than (or equal to) two different paths with Jane Number equal to 8.

Graph is like: x-x-x-x-x-x-x-Jane-x-x-x-x-x-x-x

2. The best lower bound is 8. This means that the all the paths are from the maximum Jane Number, but no path from Jane to any one. So Jane can not get to anyone else. The graph is not connected, diameter is 8.

The best upper bound is 16. This means that there are more than (or equal to) two different paths with Jane Number equal to 8, and all these paths have edges in the direction from Jane to the source. It can be no more than 16 because each semi-diameter should be equal to or less than 8, as more edges will cause a larger Jane Number.

3. The maximum Jane Number can increase by one, as we connect the new node to the previous longest path.

The The maximum Jane Number can decrease by 6, as we connect the new node to Jane and all the other nodes in graph. So now all the other nodes can get to Jane through the new node, the new Jane Number is 2.

The best lower bound is 2. As verified above, the Jane number can be 2. This means that the longest path is the diameter.

The best upper bound is 17. This means that besides the new Jane Number path, which length is 9, there exists some paths with length equal to 8. And the two paths have no intersection.

Exercise 2

1.

1.3

$$\begin{aligned} S_{1000} &= \sum_{i=1}^{1000} (X_i + 1) \\ Pr[(X_1 + \dots + X_{1000}) > 10] &= Pr[(X_1 + 1 + \dots + X_{1000} + 1) > 1010] \\ Pr[(X_1 + 1 + \dots + X_{1000} + 1) > 1010] &\leq \frac{E[(X_1 + 1 + \dots + X_{1000} + 1)]}{1010} \\ \text{By independence, } E[X_i + 1] &= 1 \\ \frac{E[(X_1 + 1 + \dots + X_{1000} + 1)]}{1010} &= \frac{100}{101} \end{aligned}$$

1.4

Suppose z is a random vector drawn from a d -dimensional multivariate Gaussian distribution with mean 0 and covariance I .

So $x = \mu + \sqrt{\Sigma}z$

$$y = v + Q(\mu + \sqrt{\Sigma}z) = (Q\sqrt{\Sigma})z + (v + Q\mu)$$

So $y \sim N(v, Q\Sigma Q^T)$

1.5

i Disprove.

Suppose playing a dice, A means the roll is 1, B means the roll is 1 or 3, C means the roll is even.

A and B are not independent, but under the condition C, A and B are independent:

$$Pr(AB|C) = 0, Pr(A|C) = 0, Pr(B|C) = 0, Pr(AB|C) = Pr(A|C) * Pr(B|C)$$

So under condition C, A and B are independent.

But $Pr(A) = 1/6$, $Pr(B) = 1/3$, $Pr(AB) = 1/6 \neq Pr(A) * Pr(B)$, A and B are not independent.

ii Disprove.

Suppose playing a dice, A means the first roll is 3, B means the second roll is 4, C means the sum of two rolls is 7.

A and B are independent, but under the condition C, A and B are no more independent:

$$Pr(AB|C) = 1/6, Pr(A|C) = 1/6, Pr(B|C) = 1/6, Pr(AB|C) \neq Pr(A|C) * Pr(B|C)$$

So under condition C, A and B are not independent.

2.1

$x \cdot v = 0$ means that $\sum_{i=1}^d x_i v_i = 0$

There are d variables (x_1, \dots, x_d) and only one equation, so S_v is a set with dimension $d-1$.

2.2

Projection of w onto $\text{span}\{v\}$ (a line) has coordinate $\|w\|_2 * \cos\theta$,

where $\cos\theta = \frac{w^T v}{\|w\|_2 \|v\|_2}$

So the Euclidean distance between w and the closest point to w in S_v is $\frac{w^T v}{\|v\|_2}$

2.3

Because the matrix is symmetric positive semi-definite, so A can be represented as $A = Q^T \Lambda Q$, where the columns of Q are the eigenvectors of A and the diagonal entries of diagonal matrix Λ are the eigenvalues of A .

$$A = Q^T \Lambda Q = Q^T \Lambda^{\frac{1}{2}} \Lambda^{\frac{1}{2}} Q$$

Λ is a diagonal matrix, so $\Lambda^T = \Lambda$

$$A = (\Lambda^{\frac{1}{2}} Q)^T (\Lambda^{\frac{1}{2}} Q)$$

Let v_i be the vectors in $\Lambda^{\frac{1}{2}} Q$,

$$A = \sum_{i=1}^d v_i v_i^T$$

3.1

i $\frac{\partial F_b}{\partial A} = v v^T$

ii $\frac{\partial F_b}{\partial v} = Av + A^T v = 2Av$

iii If A is positive semi definite, then $\forall v \in \mathbb{R}^d, v^T A v \geq 0$

If we want $\frac{\partial F_b}{\partial v} = 0$, $v = 0$

If A has one negative eigenvalue λ , corresponding eigenvector is v , then $Av = \lambda v$, which means that $v^T Av = v^T \lambda v = \lambda v^T v = \lambda \|v\|^2$

$\lambda < 0$, the first part is always < 0 ; with the norm of v increase, F_b will decrease. So there does not exist a minimum value.

iv $\|v\| = 1$ means that $v^T v = 1$

Let $g(x) = v^T v - 1$, using the method of Lagrange multipliers:

$$\mathcal{L}(v, \lambda) = v^T A v + b^T b + \lambda g(x) = v^T A v + b^T b + \lambda(v^T v - 1)$$

Let $\frac{\partial \mathcal{L}}{\partial v} = 2Av + 2\lambda v = 0$, so $Av = -\lambda v$, v is an eigenvector of A , $-\lambda$ is an eigen value of A .

Take it back to \mathcal{L} , $\mathcal{L} = -\lambda + b^T b$

So we want the eigen value $-\lambda$ to be minimum, v is the eigen vector corresponds to the minimum eigen value of A .

4.1

i Print the matrix as below

```
1 print(data[ 'M' ] )
```

```
[ [ 92  99   1   8  15  67  74  51 ]
  [ 98  80   7  14  16  73  55  57 ]
  [  4  81  88  20  22  54  56  63 ]
  [ 85  87  19  21   3  60  62  69 ]
  [ 86  93  25   2   9  61  68  75 ]
  [ 17  24  76  83  90  42  49  26 ]
  [ 23   5  82  89  91  48  30  32 ]
  [ 79   6  13  95  97  29  31  38 ]
  [ 10  12  94  96  78  35  37  44 ]
  [ 11  18 100  77  84  36  43  50 ] ]
```

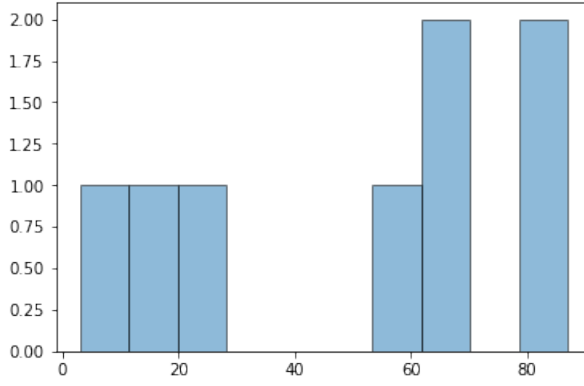
ii Dimensions of M: (10, 8)

iii The 4-th row: [85 87 19 21 3 60 62 69]

The 5-th column: [15 16 22 3 9 90 91 97 78 84]

iv the mean value of the 5th column of M: 50.5

v The histogram of the 4th row of M is:

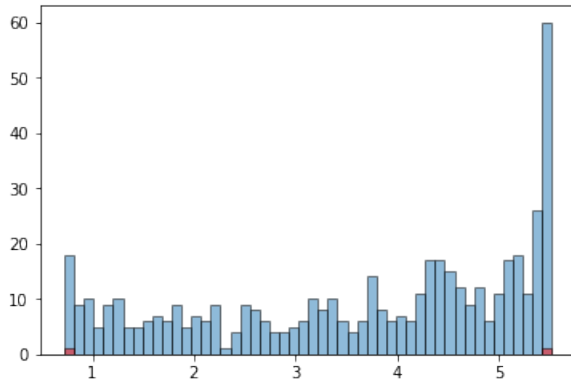


- vi The top three eigenvalues of the matrix $M^T M$ are: 1022.06770617, 263.62801936, 102.8303462

4.2

- iv $\lambda_{max} = 5.52617159, \lambda(min) = 0.72382841$

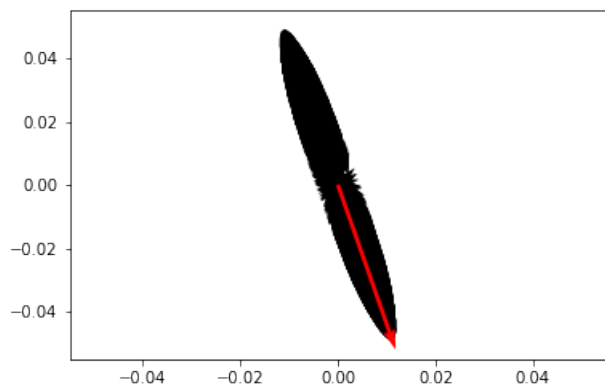
- vi A histogram of values of $\|\tilde{r}\|$
 λ_{max} and λ_{min} are shown in red



- vii The λ_{max} and λ_{min} represents the max and min value of the distorted $\|\tilde{r}\|$

- viii $vmax = [0.35078, -1]$

- ix The histogram is:



x vmax points to where changes most, the principle components.