#### CSci 5715, Fall 20: Homework 4

#### Table of Participation

|  |  |  |
| --- | --- | --- |
| Question ID | Answer drafted by | Answer reviewed by |
| 1 |  |  |
| 2 |  |  |
| 3 |  |  |
| 4 |  |  |
| 5 |  |  |

**Question 1**. Figure 1 shows the occurrences of three types of events represented by red circle, blue triangle, and yellow square. It further shows two different types of spatial partitioning (A, B). Assume that events co-occur if the distance between their closest point is at most 1.5 units.

A picture containing scatter chart

Description automatically generated

**Q1a).** Draw the neighborhood graph for the events shown in Figure 1.



**Q1b).** Fill out Table 1 to show the values of Pearson’s Correlation (for partition A, and partition B), Support (partition A, and partition B), Ripley’s Cross-K and Participation Index. For calculating Pearson’s Correlation refer to **Appendix A**.

Table 1. Results of co-occurrence measures

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Pearson’s Correlation**  **(Partition A)** | **Pearson’s Correlation**  **(Partition B)** | **Support**  **(Partition A**) | **Support (Partition B)** | **Ripley’s Cross-K** | **Participation Index** |
| **red-blue** | **-1** | **1** | **0** | **0.5** | **0.25** | **0.5** |
|  | **1** | **-1** | **0.5** | **0** | **0.25** | **0.5** |
|  | **-1** | **-1** | **0** | **0** | **0** | **0** |

**Question 2**

Consider 9 cells, namely, A, B, …, I, in a raster dataset as shown in Figure 2(a). The attribute values of these cells are shown in Figure 2(b). Assume that the neighborhood of a cell consists of the cells **sharing an edge** with the cell. For example, B and D are the neighbors of A.

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| |  |  |  | | --- | --- | --- | | A | B | C | | D | E | F | | G | H | I | | |  |  |  | | --- | --- | --- | | 42 | 40 | 36 | | 38 | 36 | 31 | | 36 | 33 | 45 | |
| Figure 2(a): Location of Cells | Figure 2(b): Attribute Values of Cells |

**Q2a)** When computing the Moran’s I Index for the dataset in Figure 2, what is the size of the matrix W?If W is defined as: wij = 1 if cells i and j are neighbors and wij = 0 if cells i and j are not neighbors. Show the matrix W for the dataset in Figure 2.

9 cells, so the W matrix is 9x9.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F | G | H | I |
| A | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 | 0 |
| B | 1 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 0 |
| C | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | 0 |
| D | 1 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 0 |
| E | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 |
| F | 0 | 0 | 1 | 0 | 1 | 0 | 0 | 0 | 1 |
| G | 0 | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 |
| H | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 | 1 |
| I | 0 | 0 | 0 | 0 | 0 | 1 | 0 | 1 | 0 |

**Q2b)** Compute the Moran’s I for the dataset shown in Figure 2(b).

First, normalize each row of the W matrix.

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | A | B | C | D | E | F | G | H | I |
| A | 0 | 1/2 | 0 | 1/2 | 0 | 0 | 0 | 0 | 0 |
| B | 1/3 | 0 | 1/3 | 0 | 1/3 | 0 | 0 | 0 | 0 |
| C | 0 | 1/2 | 0 | 0 | 0 | 1/2 | 0 | 0 | 0 |
| D | 1/3 | 0 | 0 | 0 | 1/3 | 0 | 1/3 | 0 | 0 |
| E | 0 | 1/4 | 0 | 1/4 | 0 | 1/4 | 0 | 1/4 | 0 |
| F | 0 | 0 | 1/3 | 0 | 1/3 | 0 | 0 | 0 | 1/3 |
| G | 0 | 0 | 0 | 1/2 | 0 | 0 | 0 | 1/2 | 0 |
| H | 0 | 0 | 0 | 0 | 1/3 | 0 | 1/3 | 0 | 1/3 |
| I | 0 | 0 | 0 | 0 | 0 | 1/2 | 0 | 1/2 | 0 |

x\_bar (mean) = 37.44

z-matrix: z = x\_i - x\_bar

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cell | A | B | C | D | E | F | G | H | I |
| x\_i | 42 | 40 | 36 | 38 | 36 | 31 | 36 | 33 | 45 |
| z | 4.56 | 2.56 | -1.44 | 0.56 | -1.44 | -6.44 | -1.44 | -4.44 | 7.56 |

Moran’s I = -0.26825 (from Excel)

**Q2c)** According to the textbook when calculating Moran’s I index , we represent Z score at each location as , where is the independent variable at each location and   is its mean. If we denote , where , and is the standard deviation of . Prove and are equal.

sigma (std. dev.) = 4.362

z’-matrix: z = (x\_i - x\_bar)/sigma

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| Cell | A | B | C | D | E | F | G | H | I |
| x\_i | 42 | 40 | 36 | 38 | 36 | 31 | 36 | 33 | 45 |
| z’ | 1.04 | 0.59 | -0.33 | 0.13 | -0.33 | -1.48 | 0.33 | -1.02 | 1.73 |

I’ = -0.26825 (from Excel)

The result of multiplying the z matrix by a scalar (1/sigma) does not affect the results of matrix multiplication.

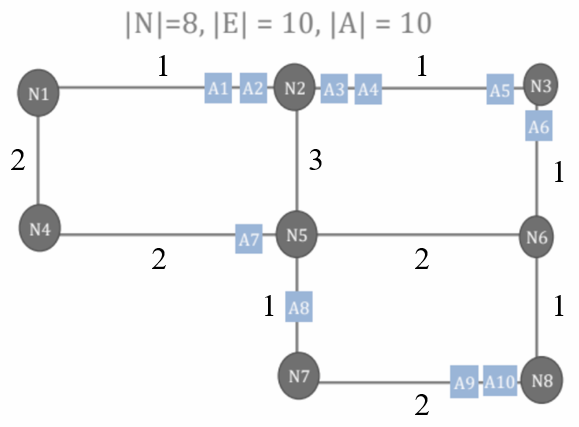
For example, if a matrix A is multiplied by a scalar c, then the result of its product with another matrix B is not otherwise affected.

Thus, because the scalar constants on top and bottom, will be equal .

**Question 3**

The undirected graph shown in Figure 3 has 8 nodes and 10 edges. Each edge is associated with a weight and a number of activities shown as blue square (Ai). A path is a sequence of edges and the weight and the number of activities of a path is the total weight and number of activities of all the edges contained in the path. For example, Path N1N2N3 has a weight 1 + 1 = 2 and a number of activities 2+ 3 = 5. The density of a path is defined as the ratio between the number of activities and the weight of a path. For example, Path N1N2N3 has a density of 5/2 = 2.5.

An interest measure named “density ratio” [1] is applied on network space for linear hotspot detection. It is defined as the ratio between the density of activities within a path and the density of activities outside the path. For example, the density ratio of Path N1N2N3 is .

**Table

Description automatically generated**

**Figure 3. An undirected graph.**

**Note: the weights may not be proportional to the**

**length of the edges**

**Q3a)** List the density ratios of all the **shortest paths** starting from Node N1. Fill your answer inside Table 3. You need to give both the shortest path and the density ratio.

**Note: If there are multiple shortest path, pick the path with higher number of activities.**

**Table 3: Shortest Paths and their density ratios**

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Start** | **End** | **Shortest Path** | **Weight** | **#Activities** | **Inside Density** | **Outside Density** | **Density ratio** |
| N1 | N2 | N1-N2 | 1 | 2 | 2 | 0.53 | 3.75 |
| N1 | N3 | N1-N2-N3 | 2 | 5 | 2.5 | 0.36 | 7 |
| N1 | N4 | N1-N4 | 2 | 0 | 0 | 0.71 | 0 |
| N1 | N5 | N1-N2-N5 | 4 | 2 | 0.5 | 0.67 | 0.75 |
| N1 | N6 | N1-N2-N3-N6 | 3 | 6 | 2 | 0.31 | 6.5 |
| N1 | N7 | N1-N2-N5-N7 | 5 | 3 | 0.6 | 0.64 | 0.94 |
| N1 | N8 | N1-N2-N3-N6-N8 | 4 | 6 | 1.5 | 0.33 | 4.5 |

[1] Janeja, V.P. and Atluri, V., 2005, March. LS 3: A linear semantic scan statistic technique for detecting anomalous windows. In Proceedings of the 2005 ACM symposium on Applied computing (pp. 493-497). ACM.

**Q3b)** A simple path is defined as a path that has no repeating nodes (i.e., no cycle). For example, Path N1N2N3N6N5N7 is a simple path from N1 to ~~N5~~N7. List all the simple paths from N1 to N5 that have a higher density ratio than the shortest path from N1 to N5 (i.e., N1N2N5). Which of them has the highest density ratio?

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Simple Path | Weight | # Activities | Inside Density | Outside Density | Density Ratio |
| N1-N2-N5 | 4 | 2 | 0.5 | 0.67 | 0.75 |
| N1-N4-N5 | 4 | 1 | 0.25 | 0.75 | 0.33 |
| N1-N2-N3-N6-N5 | 5 | 6 | 1.2 | 0.36 | 3.3 |
| N1-N2-N3-N6-N8-N7-N5 | 7 | 9 | 1.29 | 0.11 | 11.6 |

The longer simple paths (with weights greater than 4) both have higher density ratios than the shortest path. The last simple path in the table above, N1-N2-N3-N6-N8-N7-N5, has the highest density ratio of 11.57.

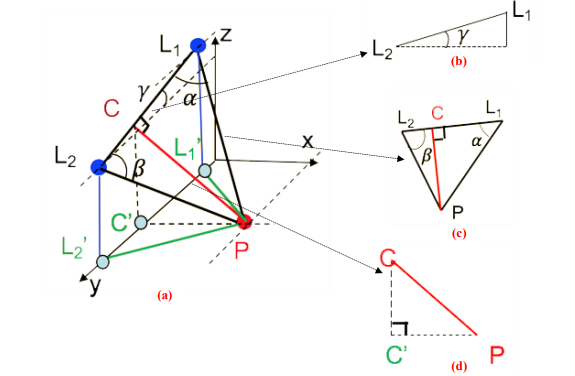
**Question 4**.

A ski athlete wants to estimate his position P through 3D triangulation. The athlete uses the following knowledge to estimate the position. **1.** Positional knowledge of two consecutive chairlift stations (L1, L2), and **2.** A theodolite [1] to estimate the angle () between each of the stations. 3D triangulation uses two 2D triangulation to compute the position. The following figures are from slide #27 on positioning, that shows the formulation of two 2D triangulations. (**Note:** sin-1(0.78) = 0.89 radians, or 51.26 degrees. cos(0.89) = 0.63).

There are two construction(s).

**1.** As shown in Figure 4 c. **PC** is the line perpendicular to line L1L2.

**2.** As shown in Figure 4 d. (L1’, L2’, C’) are the xy plane projections of (L1, L2, C).



**Figure 4.** 3D Triangulation using two 2D triangulation.

Using the following numerical values estimate the position of the athlete.

|  |  |  |  |
| --- | --- | --- | --- |
| **Coordinates of L1** (x, y, z) | **Coordinates of L2** (x, y, z) | **Angle (L1P, L1L2)** | **Angle (L2P, L1L2)** |
| (0, 12, 100) | (0, 72, 25) | 60° | 60° |

PL2L1 is an equilateral triangle

gamma=tan^-1(75/60)=51.34 degrees

L1L2=sqrt((72-12)^2+(25-100)^2)=96.05

PC=sqrt(3)/2\*L1L2=83.18

Coordinates of C=(0,12+60/2,25+75/2)=(0, 42, 62.5)

CC’=62.5

PC’=sqrt((83.18)^2-(62.5)^2)= 54.89=Px

L1’C’=42-12=30

Py=L1’y+PC\*cot(alpha)\*cos(gamma)= 12+83.18\*cot(60)\*cos(51.34)=42.0005

Coordinates of P=(54.89,42.0005,0)

**Question 5**.

Three sound recorders in city streets recorded sound of a gunshot at 4:06:29 PM, 4:06:30 PM, and 4:06:31 PM, respectively. According to the flash light of the gunshot detected by recorders, this gunshot occurred at 4:06:28.500 PM. The locations of these three recorders are (0, 0), (0, 400 meters), and (990.1690 meters, 0), respectively. The speed of sound is 340 meters / second. (Assume that all the recorders and the gunshot are on a two-dimensional plane).

**Q5a)** Determine the exact position of the gunshot. Justify the answer by showing the intermediate calculations.

(x-0)^2+(y-0)^2=(0.5\*340)^2= 28900

(x-0)^2+(y-400)^2=(1.5\*340)^2=260100

(x-990.169)^2+y^2=(2.5\*340)^2=722,500

(y-400)^2-y^2=231200

-800\*y+400^2=231200

y=-89

x^2=28900-89^2=+-144.84

(x-990.169)^2=714579

Therefore, x=144.84, y=-89

**Q5b)** Now suppose we do NOT have the third sound recorder at (990.1690 meters, 0). Can we still determine the exact position of the gun shot? Justify your answer.

We cannot determine the position of the gun shot. There will be two possibilities: (144.84,-89) and (-144.84,-89)

**Q5c)** Now suppose we do NOT have the third sound recorder at (990.1690 meters, 0). And, we do NOT know the time when the gunshot happened. Give all the possible locations where this gunshot happened.

All solutions to: sqrt((x-0)^2+(y-400)^2)-sqrt((x-0)^2+(y-0)^2)=340

**Appendix A:**

A picture containing diagram

Description automatically generated