5. Introduction to Clustering

Exercises due Oct 20, 2021 17:29 IST Completed

Clustering Measures

1/1 point (graded)

Which of the following could be valid measures of clustering? Pick the ones that will make sense more often than not. That is, choose only the ones that do not just apply under a few contrived settings. For this problem, think of clustering as a notion that captures how close a set of nodes are with respect to other nodes in the graph.

- [ ] Nodes with the same degree.
- [ ] Nodes with node identification values (node index) close to each other.
- [x] Nodes with edges connecting each other more often than their edges connecting outside of them.
- [ ] Nodes that are "key" to the connectivity of the graph in the following sense: in the absence of them (along with their edges), the graph is no longer connected.

✔

Solution:

1. Nodes with the same degree – No, two nodes with the same degree need not be close to each other.
2. Nodes with node identification values (node index) close to each other – No, this is not a good measure of clustering. It is dependent on the index we give to every node.
3. Nodes with edges connecting each other more often than their edges connecting outside of them – Yes, this captures some measure of clustering.
4. Nodes that are "key" to the connectivity of the graph in the following sense: in the absence of them (along with their edges), the graph is no longer connected – No, this notion captures nodes that are critical to the graph. Such nodes need not be close to each other in the graphical sense (such as in 3.)

Submit You have used 2 of 2 attempts

Answers are displayed within the problem

Graph Laplacian

We now introduce the **graph Laplacian matrix** of a graph. Throughout the remainder of this lecture, we assume only that the graph is *simple and undirected*. Everything discussed here also works for undirected, *weighted*, simple graphs.

The **graph Laplacian matrix** L is defined as

$$L = D - A,$$

where A is the adjacency matrix and D is the degree matrix. The **degree matrix** for an undirected, unweighted graph is a matrix whose off-diagonal elements are equal to 0 and whose diagonal elements are given by

$$D_{ii} = \sum_j A_{ij}.$$

In other words, the degree matrix of an undirected, unweighted, simple graph is simply a diagonal matrix whose diagonal entries are the degrees of the nodes. In the case of a weighted, undirected, simple graph the definition is the same but the interpretation no longer concerns the degree of the nodes, but rather the sum weight of the edges emanating from the nodes.

edges emanating from the nodes.

Graph Laplacian – Example

4/4 points (graded)

Consider the adjacency matrix

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

1. How many components does the graph have?

2 ✔ **Answer:** 2

2. What are the diagonal values of the Laplacian matrix of the graph?

First diagonal element: 1 ✔ **Answer:** 1

Second: 1 ✔ **Answer:** 1

Third: 0 ✔ **Answer:** 0

Solution:

1. The first and second node are connected to each other, while the third node is not connected at all. Thus there are two components: $\{\{1,2\},\{3\}\}$.

2. The diagonal elements of A are zero (as it is simple), so the diagonal elements of L come from D. The first diagonal element is the degree of the first node, etc.

Submit You have used 1 of 2 attempts

Answers are displayed within the problem

Symmetry

1/1 point (graded)

Let L be the Laplacian of an undirected simple graph (weighted or unweighted). **True** or **False**? The Laplacian is symmetric.

- (●) True
- ( ) False

✔

Solution:

The adjacency matrix and the degree matrix are both symmetric – hence the Laplacian is also symmetric.

Submit You have used 1 of 1 attempt

Answers are displayed within the problem

Discussion

Hide Discussion

Topic: Module 3: Network Analysis:Spectral Clustering / 5. Introduction to Clustering

Add a Post

All Posts

@STAFF: Clustering Measures Formulation

question posted 2 months ago by **kozaronek**

This question is extremely vaguely formulated.

Can you please reformulate the question to make it less ambiguous?

Here are some formulations that are not quite clear to me.

1) By nodes with the same degree do you mean same in the sense that they have equal degree or that they are in a similar "degree range" ?

2) Is there some kind of logic applied to the identification of nodes or not?

4) These 'key' nodes make me think of gatekeepers but I'm unsure whether they could be clustered, since gatekeepers in any organization might not have anything in common (are not close to one another) other than being of the same type/cluster, namely they are all "gatekeepers".

I think this sort of question belongs in a report where it can be thoroughly discussed and not in a True False setting with 2 tries.

This post is visible to everyone.

Add a Response

4 responses

_alxk
2 months ago

100%

Agreed this question can be made clearer...

posted 2 months ago by **simon-templar**

This part of knowledge check is poorly reviewed, that's for sure

posted 2 months ago by **mikhail_potapov**

Add a comment

StBishop
2 months ago

I'm sure very few guessed it right xD

In my case I was applying similarity logic as per video - and that didnt help =)

Add a comment

michael_zhang_2021
2 months ago

I ever got same confusion when I worked on this question. One idea to help you find right option is, don't be confused with data points categorization and clustering. Categorization means split data points to different categories based on point attributes, while clustering means measure similarity distance by hyperspace defined in data points dimension. For network graph problem, the similarity distance is more easier, as it's just about path distance between nodes(data points). If think like this way, I think you can find right answer.

Add a comment

marika_catapano
2 months ago

100%

Add a comment

Showing all responses

Add a response:

Preview

Submit

< Previous Next >

edX

About
Affiliates
edX for Business
Open edX
Careers

News

Legal

Terms of Service & Honor Code
Privacy Policy
Accessibility Policy
Trademark Policy
Sitemap

Connect

Blog
Contact Us
Help Center
Media Kit
Donate