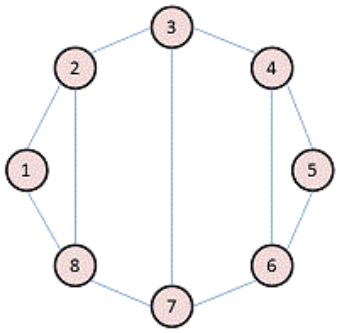
Communities

Question 1:

For the following graph:

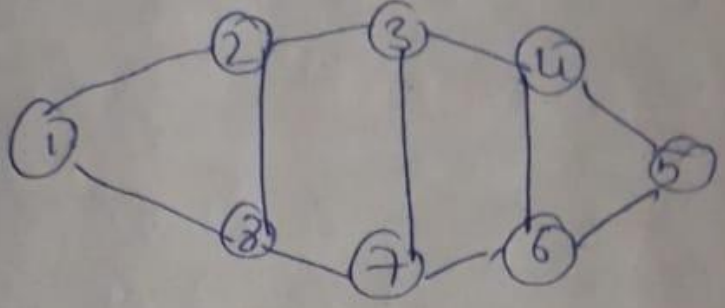


Write the adjacency matrix A, the degree matrix D, and the Laplacian matrix L. For each, find the sum of all entries and the number of nonzero entries.

sol:

Assignment -5

8-1: Communities



Adjacy matrix A

	1	2	3	4	5	6	7	8
1	0	1	0	0	0	0	0	1
2	1	0	1	0	0	0	0	1
3	0	1	0	1	0	0	1	0
4	0	0	1	0	1	1	0	0
5	0	0	0	1	0	1	0	0
6	0	0	0	1	1	0	1	0
7	0	0	0	1	0	0	1	1
8	1	1	0	0	0	0	1	0

Degree matrix, D

	1	2	3	4	5	6	7	8
1	2	0	0	0	0	0	0	0
2	0	3	0	0	0	0	0	0
3	0	0	3	0	0	0	0	0
4	0	0	0	3	0	0	0	0
5	0	0	0	0	2	0	0	0
6	0	0	0	0	0	3	0	0
7	0	0	0	0	0	0	3	0
8	0	0	0	0	0	0	0	3

Laplacian Matrix $L = D - A$

	1	2	3	4	5	6	7	8
1	2	-1	0	0	0	0	0	-1
2	-1	3	-1	0	0	0	0	-1
3	0	-1	3	-1	0	0	-1	0
4	0	0	-1	3	-1	-1	0	0
5	0	0	0	-1	2	-1	0	0
6	0	0	0	-1	-1	3	-1	0
7	0	0	-1	0	0	-1	3	-1
8	-1	-1	0	0	0	0	-1	3

A has 22 non-zero entries

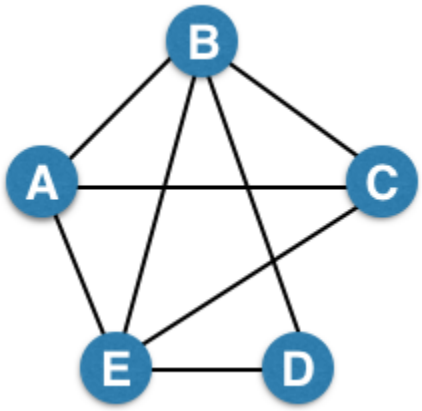
D has 8 non-zero entries

L has 30 non-zero entries.

The sum of the entries of A is 22

Question 2:

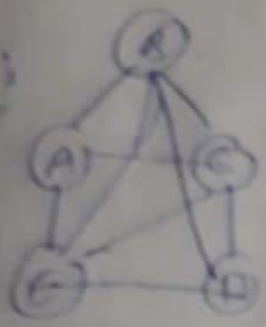
Consider the following undirected graph (i.e., edges may be considered bidirectional):



Run the "trawling" algorithm for finding dense communities on this graph and find all complete bipartite subgraphs of types $K_{3,2}$ and $K_{2,2}$. Note: In the case of $K_{2,2}$, we consider {{W, X}, {Y, Z}} and {{Y, Z}, {W, X}} to be identical.

sol:

Q2)



itemsets

A = {B, E, C}

B = {A, E, D, C}

C = {B, A, E}

D = {A, E}

E = {A, B, C, D}

K_3, 2 and K_2, 2

K_3, 2 min support threshold = 3

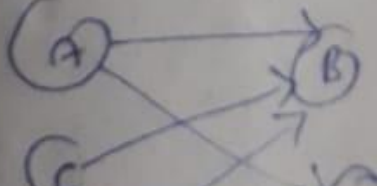
item	support
A	3
B	4
C	3
D	2
E	4

item	Support
B,E	3
B,C	2
E,C	2
A,E	2
A,D	2
A,C	2
E,D	2
D,C	2
B,A	2
B,D	1
C,D	2

→

item	Support
B,E	3

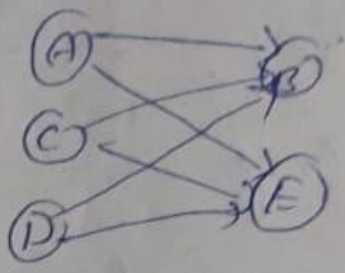
1 bipartite graph

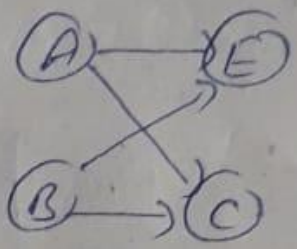


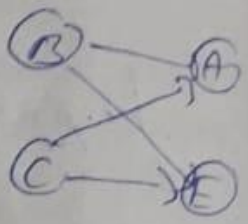
$K_{2,2}$ min support = 2

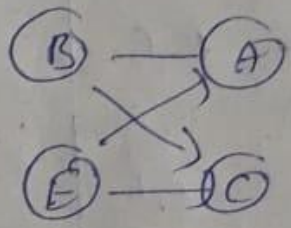
7 - bipartite graphs

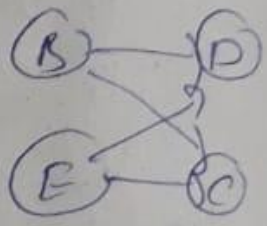
item	Support
B, E	3
B, C	2
E, C	2
A, E	2
A, D	2
A, C	2
D, C	2
B, A	2
C, D	2

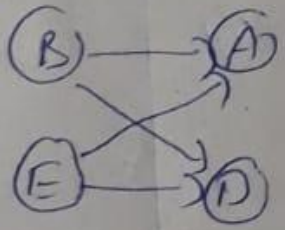


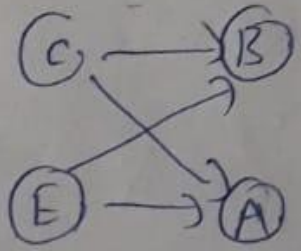






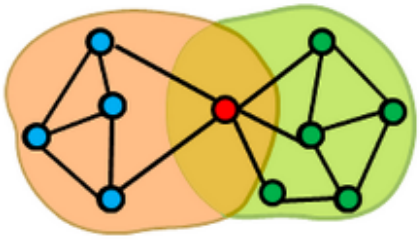




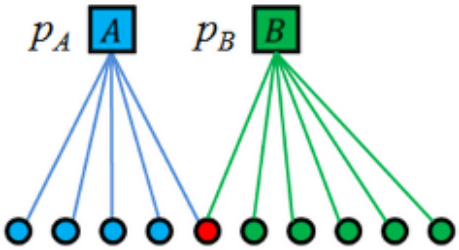


Question 3:

We fit AGM to the network on the left, and found the parameters on the right:



Network



Learned AGM parameters

Find the optimal values for p_A and p_B.

Sol:

Pa = Number of edges in the network / Total possible number of edges = 7/5c2 = 7/10.

Pb = Number of edges in the network / Total possible number of edges = 9/6c2 = 9/15.