CS235 ASSIGNMENT
Yidi Wang
862114701

Question 1
Part 1

```
clear; clc;
edgelist = dlmread('graph.txt');
edgelist = unique(edgelist, 'rows');
A = sparse(edgelist(:,1), edgelist(:,2), 1);
```

Part 2, 3
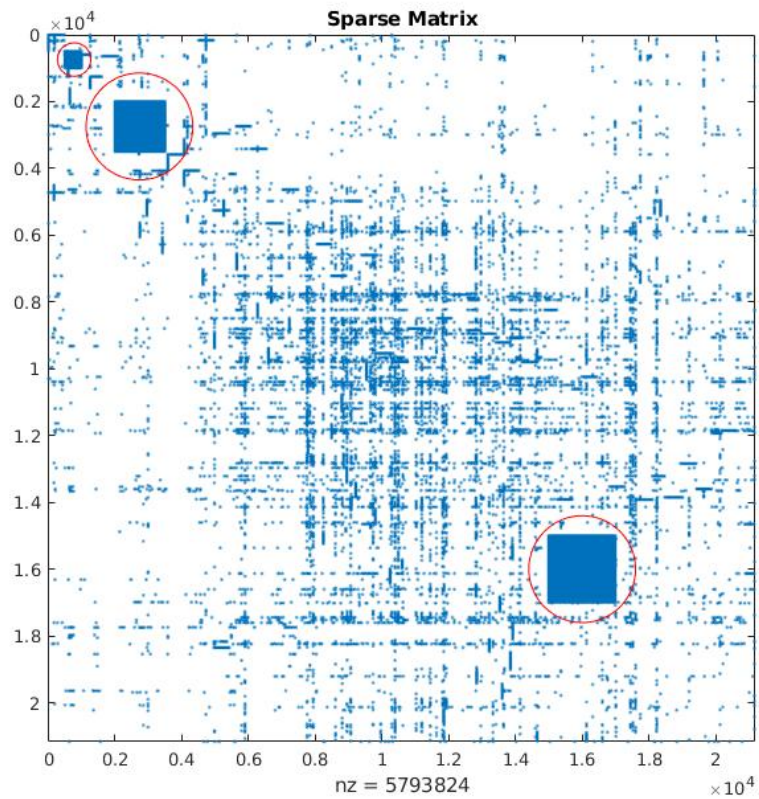Three dense blocks are annotated by red circles.

Figure 1: spy-plot

Question 2
Part 1

The number of connected nodes are calculated for each node in the edge-list of the graph.

```
max_node = max(edgelist(:));
deg_node = zeros(1, max_node);
for i = 1:max_node
    deg_node(i) = sum(rows(:) == i);
end
```
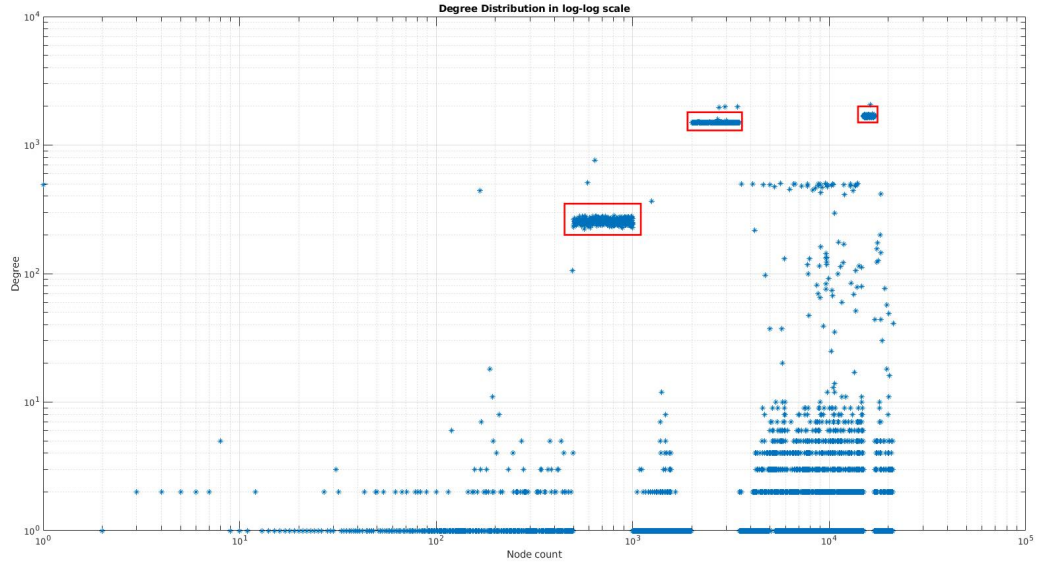
Part 2

Figure 2: Degree distribution in log-log scale

Part 3

Yes. In the clean degree distribution figure, the majority of the nodes have a small degree. While the degree becomes larger, the number of corresponding nodes become less and less. Degree and node counts are negtively correlated. However, in this distribution, node count is increasing while the degree increase. They are positively correlated.

Part 4

(a). There are three abnormal blocks. They are annotated in red circles. Above the 10^2 on the y axis in the above figure (when degree is big enough), there are three nearly horizontal lines, which means that there are many consecutive nodes having similar degrees. This will cause dense blocks in the sparse matrix. (b). The upper thick line area in Fig.2 is corresponding to the upper right dense block in Fig.1. The lower left thick line in Fig.2 is corresponding to the lowest right dense block in Fig.1. And the middle thick line in Fig.2 is corresponding to the middle block in Fig.1.

Question 3
Part 1

Part 2
(a).

```
[u,s,v] = svds(A);
ds = diag(s);
dsp = norm(ds(n+1:end)) / norm(ds);
```

I tried the above code with different n. When $n = 2$, the error is 0.1157. When $n = 3$, the error is 0.0271. So, if we want a reconstruction correctness larger than 90%, the number of singular value is 3.

$$\begin{bmatrix} 1679.9 \\ 1501.0 \\ 255.22 \\ 37.715 \\ 35.5439 \end{bmatrix}$$
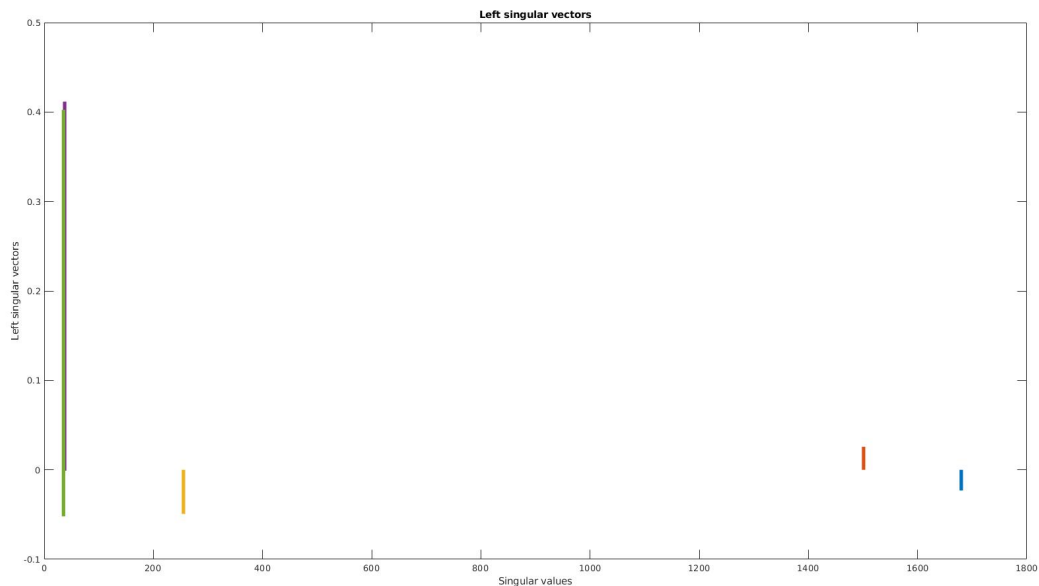
(b).

Figure 3: Left singular vectors

(c). We can identify from Fig.3 that there are three abnormal blocks in the sparse matrix. The first two left singular vectors' range are much bigger than the last three vectors. Let's explain it from the geometric meaning of SVD decomposition: S means the scale transformation of the original matrix,

3

while U means the rotation to the matrix.