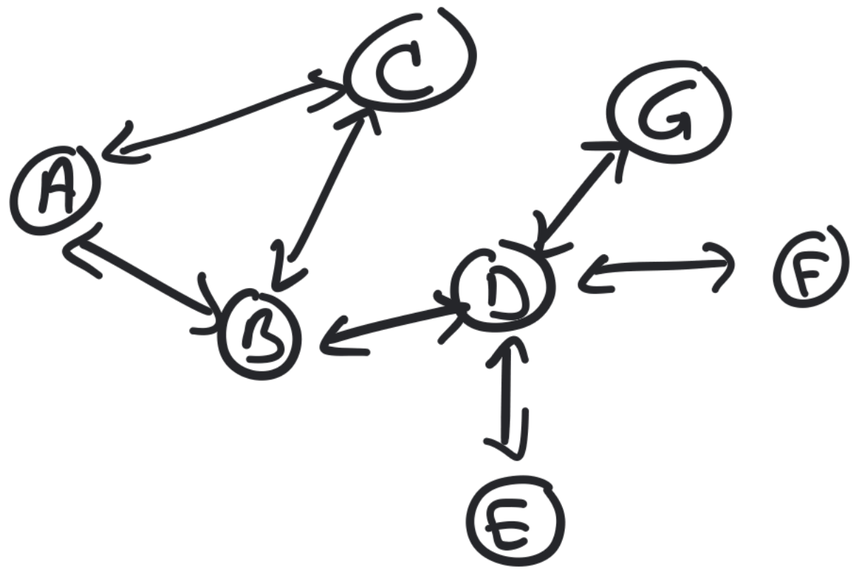
Problems 21

Network Centrality



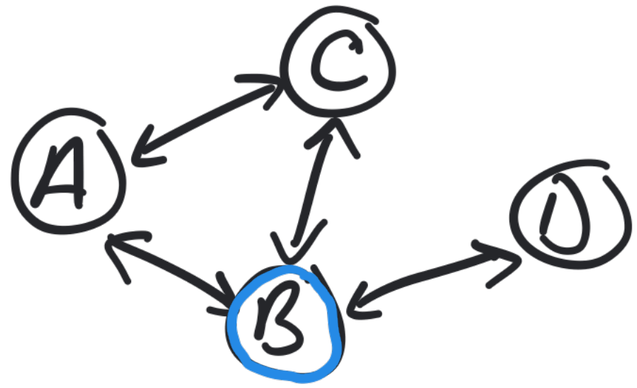
Which protein is the most "important" or "central" in this network?

- Rather than counting links, we should weight links based on importance of the proteins they connect.

- to calculate "importance" of node i, we need the "importance" of all nodes $j \neq i$ to weight the edges.

Adjancey matrix

$$\underline{A} = \begin{array}{c} \\ A \\ B \\ C \\ D \end{array} \begin{array}{c} \begin{array}{cccc} A & B & C & D \end{array} \\ \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \end{array}$$



$$a_{ij} = \begin{cases} 0 \\ 1 \text{ if } i \leftrightarrow j \end{cases}$$

- only include direct links
- no self connections ($A \leftrightarrow A$)

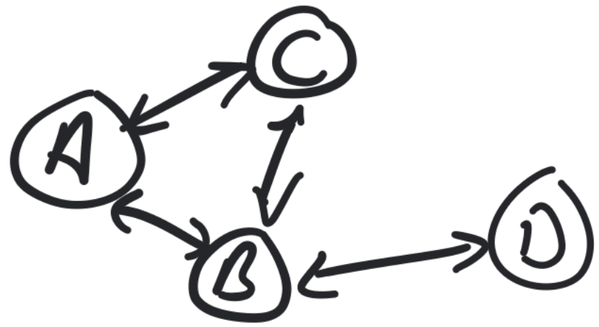
The eigenvector associated w/ the largest eigenvalue contains the "importance" or "centrality" of the nodes.

For $\underline{A}$: $\lambda_{max} \approx 2.2$

$$\underline{v}_{max} = \begin{pmatrix} 0.52 \\ 0.61 \\ 0.52 \\ 0.28 \end{pmatrix}$$

"Eigencentrality"

Dynamic Interpretation of Centrality



- Randomly walking along links.
- Stop at random times & note which node in at.
- Eigencentrality is proportional to the probability of stopping @ node i.

Largest application is PageRank (invented by Google).

- find pages w/ term.
- Calculate centrality based on web links.
- present results in order of eigencentrality.
- PageRank includes "damping factor".