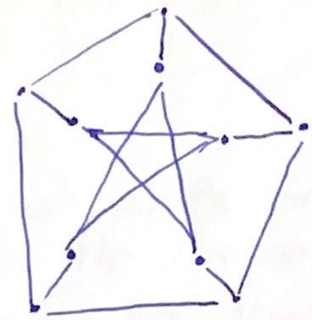
MATH517 Midterm Exam

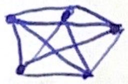
Nathan Bick

1. Using Kuratowski's theorem, prove that this network is not planar:



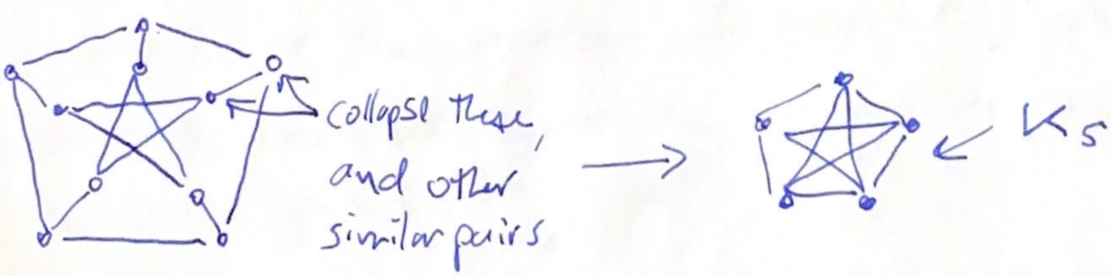
A planar network is defined as one that can be drawn on a plane without any edges crossing. This means that there is at least one such arrangement, meaning some depictions may have crossing edges.

Kuratowski's theorem states that any nonplanar graph contains one of two subgraphs, K_5 or UG.

K_5 is a Star Shaped Network  or pentagram.

The book indicates that there may be "extra" nodes along edges of the depiction of K_5 or UG in the theorem.

We can clearly see the K_5 network is contained within, but requires removal of 5 nodes to get to $K5$.



2. In section 5.3.1, we gave one possible definition of the trophic level x_i of a species in a directed food web as the mean of the trophic level of the species' prey, plus one.

(a) Show that x_i, when defined in this way, satisfies

$$x_i = 1 + \frac{1}{k_i^{in}} \sum_j A_{ij} x_j .$$

(b) The expression does not work for autotrophs — species with no prey — because the corresponding vector element is undefined. Such species are usually given a trophic level of one. Suggest a modification of the calculation that will correctly assign trophic levels to these species, and hence to all species. Thus, show that x_i can be calculated as the i^{th} element of a vector

$$\vec{x} = (D-A)^{-1} D \cdot \vec{1}$$

and specify how D is defined.

(a) First, we interpret the intuitive meaning of the expression.

A_{ij} is the incidence matrix. In the directed network of the food web, the A_{ij} entry is nonzero if i eats j.

x_j is the trophic level of species j. k_i^{in} is the sum of edges into i, or the number of prey for j species.

This formula, therefore, gives the average trophic level of the prey species of j, plus 1.

Nonprey species trophic levels are zero'd out in the sum

~~and the denom~~

(b) If the value $k_i^{in} = 0$, which happens in the case of an autotroph, the previous definition has issues. If we first translate the formula into matrix algebra, we get: $\vec{x} = D^{-1} A \vec{x} + \vec{1}$. In this direct translation we have $D_{ii} = k_i^{in}$, so D is a diagonal matrix whose elements are the in-degrees, but this has the same issue. Therefore, we must modify such that $D_{ii} = \max(k_i^{in}, 1)$, giving us

$\vec{x} = D^{-1} A \vec{x} + \vec{1}$ defined for autotrophs.

Then we can do algebra to get the formula.

$$D\vec{x} = A\vec{x} + D\vec{1}$$

$$\Leftrightarrow D\vec{x} - A\vec{x} = D\vec{1} \Leftrightarrow (D-A)\vec{x} = D\vec{1}$$

$$\Rightarrow \vec{x} = (D-A)^{-1} D\vec{1}.$$

3. As we saw in Section 7.1.3, the Katz centrality in vector form satisfies the equation $\vec{x} = \alpha A \vec{x} + \vec{1}$.

(a) Show that the Katz centrality can also be written in series form $\vec{x} = \vec{1} + \alpha A \vec{1} + \alpha^2 A^2 \vec{1} + \cdots$

(b) Hence, argue that in the limit where α is small but non-zero, the Katz centrality is equivalent to the degree centrality.

(c) Conversely, in the limit $\alpha \to \frac{1}{\kappa_1}$, where κ_1 is the largest eigenvalue of the adjacency matrix A, argue that $\vec{x}$ becomes proportional to the leading eigenvector, which is the eigenvector centrality.

(a). The Katz centrality, given as $\vec{x} = \alpha A \vec{x} + \vec{1}$, can be rewritten as $\vec{x} = (I - \alpha A)^{-1} \vec{1}$.

If we rewrite as $\vec{x} = \dfrac{\vec{1}}{I - \alpha A}$, we realize that this has the form of the sum of a geometric series of the form $\frac{a}{1-r}$. Therefore, by the definition of a geometric series, we see

$$\vec{x} = \frac{\vec{1}}{1 - \alpha A} = \vec{1} + \alpha A \vec{1} + \alpha^2 A^2 \vec{1} + \cdots = \sum_{k=1}^{\infty} \vec{1} (\alpha A)^k .$$

(b) In the limit $\lim\limits_{\alpha \to 0,\text{ but not } = 0} \left[\vec{1} + \alpha A \vec{1} + \alpha^2 A^2 \vec{1} + \cdots \right]$.

we see that the terms with higher degree "dissapear" much more quickly.

Degree centrality is given as $\vec{x} = A D^{-1} \vec{x}$, but is defined as the value of the degree of the node for each node.

In the limit, we get [illegible]

(c) Now, in considering the limit $\alpha \to \frac{1}{\kappa_1}$, first we note that the eigenvector centrality is defined as $x_i = \frac{1}{\kappa_1} \sum_{i=1} A_{ij} x_j$, which is equivalent to $A\vec{x} = \kappa_1 \vec{x}$. We know we can use κ_1 because of the Perron-Frobenius Theorem, which states that for a non-negative matrix, like the adjacency matrix, there is only one non-negative eigenvector and it has the leading eigenvalue, which is κ_1.

So we see in the Katz centrality definition, which is $x_i = \alpha \sum_j A_{ij} x_j + \beta$, then in the limit

$$\lim_{\alpha \to \frac{1}{\kappa_1}} \left(\alpha \sum_j A_{ij} x_j + \beta\right) = \frac{1}{\kappa_1}\left(\sum_j A_{ij} x_j\right) + \beta$$

which is nearly the same definition as the eigenvector centrality, with an additional constant term. Therefore, Katz is proportional to the eigenvector centrality in this limit.

4. Suppose a directed network takes the form of a tree with all edges pointing inward towards a central node.
What is the PageRank centrality of the central node in terms of the single parameter α appearing in the definition of PageRank and the distances d_i from each node i to the central node.

PageRank is defined by $x_i = \alpha \sum_j A_{ij} \frac{x_j}{k_j^{out}} + \beta$, which is intuitively Katz Centrality with the addition of taking into account the node out-degree of neighbors.
This can be rewritten as $\vec{x} = (I - \alpha A D^{-1})^{-1} \vec{1}$, $D_{ii} = \max(k_i^{out}, 1)$

By the definition of a tree (only one parent per node), the $k_j^{out} = 1$ for all j, so $D = I$. So our equation becomes

$$\vec{x} = (I - \alpha A)^{-1} \vec{1} \Leftrightarrow \alpha \sum_j A_{ij} x_j + \beta = x_i$$

We also note that in a directed network such as this, the adjacency matrix only presents 1 if there is a link from j to i, and 0 otherwise.

If node i, which is not the center, is at the "end" of the tree, $\alpha \sum_j A_{ij} x_j = 0$, so $x_i = \beta$ is the PageRank centrality.
Its parent is then $\alpha\beta + \beta$ PageRank centrality, at least.
The next parent of i's parent would have ~~some~~ an additional power of α, and the expression of β depends on its number of children.
We see that the power of α is related to the distance from the center

5.

		Women D	Women I	Women R	Total
Men	D	0.25	0.04	0.03	0.32
Men	I	0.06	0.15	0.05	0.26
Men	R	0.06	0.05	0.30	0.41
Total		0.37	0.24	0.38	

Calculate the modularity of the network with respect to political persuasion.

As is mentioned in the textbook, when we have data in this format, it is useful to use the following definition of Modularity: $Q = \sum_r (e_r - a_r^2)$. We know that

$e_r = \frac{1}{2m} \sum_{ij} A_{ij} \delta_{g_i,r} \delta_{g_j,r}$, which is the fraction of edges that join nodes of type r,

and

$a_r = \frac{1}{2m} \sum_i k_i \delta_{g_i,r}$, which is the fraction of ends of edges attached to nodes of type r.

For D: $e_D = 0.25$, $a_D = \frac{0.32 + 0.37}{2} = 0.34$

For I: $e_I = 0.15$, $a_I = \frac{0.26 + 0.24}{2} = 0.25$

For R: $e_R = 0.30$, $a_R = \frac{0.41 + 0.38}{2} = 0.395$

Then
$$\left.\begin{aligned} Q_D &= 0.25 - 0.34^2 = 0.134 \\ Q_I &= 0.15 - 0.25^2 = 0.086 \\ Q_R &= 0.30 - 0.395^2 = 0.144 \end{aligned}\right\} \Rightarrow Q = 0.364$$