Information sheet to be provided with Final Exam
Network Science 2024

There may be material on the exam which is not included here. There may be material here which is not needed for the exam.

Adjacency matrix: A_{ij} is the number of links from node j to node i

Simple graphs: A graph is *simple* if it is undirected, unweighted, and does not have multiedges or self-loops.

Complete graph: A simple graph where each distinct pair of nodes is linked.

Graph Laplacian (also called the *Laplacian matrix*): $\mathbf{L} = \mathbf{D} - \mathbf{A}$ where $\mathbf{D}$ is the *diagonal degree matrix* for the graph, $D_{ii} = k_i$, k_i is the degree of node i, and $L_{ij} = \delta_{ij}k_j - A_{ij}$.

Centralities and similarities

- *eigenvector centrality*: $x_i = \alpha \sum_{j=1}^{N} A_{ij}x_j$, where α is a proportionality constant.
- *Katz centrality*: $x_i = \alpha \sum_{j=1}^{N} (A_{ij}x_j) + 1$
- *PageRank centrality*: $x_i = \sum_{j=1}^{N} \left[(1-m)\frac{A_{ij}}{\max(k_j^{out},1)}x_j + m\frac{x_j}{N}\right]$, where m is constant such that $0 < m < 1$.
- *cosine similarity*: $\sigma_{ij} = \frac{n_{ij}}{\sqrt{k_i k_j}}$, where n_{ij} is the number of common neighbours of node i and j.
- *Jaccard similarity*: $\sigma_{ij} = \frac{n_{ij}}{k_i + k_j - n_{ij}}$

Matrix resolvent: $R(\mathbf{M};\mu)$ is the *resolvent* for a square N x N matrix, $\mathbf{M}$. The resolvent is defined as $R = (\mu\mathbf{I} - \mathbf{M})^{-1}$ for μ where $\mu \neq \lambda_i,\ i = 1, 2, ..., N$. Here λ_i is the i^{th} eigenvalue of $\mathbf{M}$

- $\rho(\mathbf{M})$ is the *spectral radius* of $\mathbf{M}$: $\rho(\mathbf{M}) = \max\{|\lambda_1|, |\lambda_2|, ..., |\lambda_N|\}$
- If $|\mu| > \rho(\mathbf{M})$, then $R(\mathbf{M};\mu) = \sum_{l=0}^{\infty} \frac{\mathbf{M}^l}{\mu^{l+1}}$

Perron-Frobenius theorem: We have applied the Perron-Frobenius (P-F) theorem to three different classes of real square matrices:

1. *Positive matrices* where each element of the matrix is positive, $\mathbf{B} > 0$. Then, the theorem tells us that there is a real positive eigenvalue λ where:
   - $\lambda = \rho(\mathbf{B}) > 0$, and all other eigenvalues are smaller in magnitude.
   - This eigenvalue is simple, all elements of the corresponding eigenvector have the same sign, and there are no other eigenvectors where all elements have the same sign. Note that a simple eigenvalue has algebraic multiplicity equal to 1.
2. *Irreducible matrices* Let $B_{ij} > 0$ if there is a link in a graph from node i to node j with $B_{ij} = 0$ otherwise. Then $\mathbf{B}$ is irreducible if and only if the corresponding graph is strongly connected (i.e. every node is reachable from every other node). For irreducible matrices, there is a real, positive eigenvalue λ where:
   - $\lambda = \rho(\mathbf{B}) > 0$, and this eigenvalue is simple.
   - All elements of the corresponding eigenvector have the same sign, and there are no other eigenvectors where all elements have the same sign
     - There may be other eigenvalues equal in magnitude to λ
3. *Non-negative matrices* where each element of the matrix is non-negative: $\mathbf{B} \geq 0$. There is a real, non-negative eigenvalue λ where:
   - $\lambda = \rho(\mathbf{B}) \geq 0$, and there may be other eigenvalues equal in value or equal in magnitude

- All non-zero elements of the corresponding eigenvector will have the same sign, and there may be other eigenvectors with the same property
- Note: This version of the P-F theorem is considerably weaker than the other 2

Markov's inequality: Let X be a random variable that assumes only non-negative values. Then, for all $a > 0$, $P(X \geq a) \leq \frac{\langle X \rangle}{a}$.

Chebyshev's inequality: Let X be a random variable with finite expected value, μ, and finite non-zero variance, σ^2. Then, for any $a > 0$, $P(|X - \mu| \geq a) \leq \frac{\sigma^2}{a^2}$.

The **configuration model** requires the specification of an N-element *degree sequence*: $\{k_1, k_2, ..., k_N\}$ with $k_i > 0$ for all i. The probability of two stubs being linked is $\frac{1}{K-1}$,
$K = \sum_{i=1}^{N} k_i$ is the **total degree**.

Average degree: $\overline{k} = \frac{1}{N} \sum_{i=1}^{N} k_i$.

Second moment of the degree distribution: $\overline{k^2} = \frac{1}{N} \sum_{i=1}^{N} k_i^2 = \sum_{k=1}^{k_{max}} p_k k^2$, where p_k is the probability that a node has degree k.

Preferential attachment: $\rho_i(G_a(t))$: probability that node i in graph $G_a(t)$ receives a link at time t. For linear preferential attachment: $\rho_i(G_a(t)) = \frac{k_i(G_a(t))}{K(t)}$

Graph diffusion equation: $\frac{d\langle \mathbf{n} \rangle}{dt} = -\alpha \mathbf{L} \langle \mathbf{n} \rangle$.

Fick's law on graphs: $\langle j_{ab} \rangle = -\alpha(\langle n_a \rangle - \langle n_b \rangle)$; n_a is the number of particles on node a, j_{ab} the net flux of particles per unit time from node b to node a.

Orthogonal diagonalization: A square real matrix, $\mathbf{M}$, is orthogonally diagonalizable if and only if $\mathbf{M}$ is symmetric. Then, $\mathbf{M} = \mathbf{V \Lambda V}^T$.

Rayleigh quotient: For a symmetric matrix $\mathbf{M}$, $r(\mathbf{M}, \mathbf{x}) = \frac{\mathbf{x}^T \mathbf{M} \mathbf{x}}{\mathbf{x}^T \mathbf{x}}$ is maximized when $\mathbf{x} = \mathbf{v}_1$ in which case $r = \lambda_1$. ($\mathbf{M}\mathbf{v}_1 = \lambda_1 \mathbf{v}_1$, $\lambda_1 \geq \lambda_2 \geq ...$)

Gershgorin's theorem: Let $\mathbf{B} \in \mathbb{C}^{N \times N}$ and suppose that $\mathbf{X}^{-1}\mathbf{B}\mathbf{X} = \mathbf{H} + \mathbf{F}$, where $\mathbf{H}$ is diagonal and $\mathbf{F}$ has zeros on its main diagonal. Then the eigenvalues of $\mathbf{B}$ lie on the union of the discs $\Delta_1, \Delta_2, \ldots, \Delta_N$, where $\Delta_i = \{l \in \mathbb{C} : |l - H_{ii}| \leq \sum_{j=1}^{N} |F_{ij}|\}$.

Random walks on graphs: $T_{ij} = \frac{A_{ij}}{k_i}$; $\mathbf{T}$ is the *transition matrix*, and T_{ij} is the probability that a walker takes a step from node i to node j on a simple graph.

Network-SI model: $\frac{d\langle x_i \rangle}{dt} = \beta \sum_{j=1}^{N} A_{ij} \langle (1 - x_i) x_j \rangle$, where $x_i(t)$ is a random variable that indicates the state of node i at time t.

Degree-based approximation: $\frac{d\phi_k}{dt} = k\beta (1 - \phi_k) \sum_{k'=1}^{k'_{max}} \theta(k, k')\phi_{k'-1}$, where ϕ_k is the probability that a node with degree k is infectious and $\theta(k, k')$ the function giving the probability of a link on a node with degree k being connected to a node with degree k'.

Second-moment equation: $\frac{d\langle s_i x_j \rangle}{dt} = \beta \sum_{l=1}^{N} (A_{jl} \langle s_i s_j x_l \rangle - A_{il} \langle s_i x_j x_l \rangle)$, where $s_i = 1 - x_i$.

Modularity: The modularity of a set of nodes, S_a, is $M_a = \frac{1}{2L} \sum_{i \in S_a} \sum_{j \in S_a} \left(A_{ij} - \frac{k_i k_j}{2L} \right)$

The *modularity matrix* $\mathbf{B}$ is defined using, $B_{ij} = A_{ij} - \frac{k_i k_j}{2L}$.

Spectral modularity maximization: Find $\tilde{\mathbf{s}}$ such that $\tilde{\mathbf{s}}^T \mathbf{B} \tilde{\mathbf{s}}$ is maximized with $|\tilde{\mathbf{s}}|^2 = N$.

Cut size: For a partition that breaks a simple connected graph into two disjoint groups of nodes, the cut size, c, is the number of links crossing from one group to another, and $c = \frac{1}{4}\mathbf{s}^T \mathbf{L} \mathbf{s}$, where each element of $\mathbf{s}$ is set to ± 1.

Laplacian graph partitioning: Find $\tilde{\mathbf{s}}$ such that $\tilde{c} = \tilde{\mathbf{s}}^T \mathbf{L} \tilde{\mathbf{s}}$ is minimized with $|\tilde{\mathbf{s}}|^2 = N$.

Weighted graphs: $\mathbf{W}$ is the *weight matrix*; $W_{ij} \neq 0$ if there is a link from j to i and is 0 otherwise.

Weighted normalized Laplacian: $\hat{\mathbf{L}} = \hat{\mathbf{D}}^{-1/2}(\hat{\mathbf{D}} - \mathbf{W})\hat{\mathbf{D}}^{-1/2}$. $\hat{\mathbf{D}}$ is a diagonal matrix with, $\hat{D}_{ii} = \sum_{j=1}^{N} W_{ij}$.