Distributed Averaging

- a simple distributed algorithm
- A set of nodes (agents) : N nodes
- communication constraints captured by the edges (directed) in a graph $G(V,E)$. We assume G is strongly connected.
- each node possesses a real value $x_i(0)$, at node "i". Each node does:

$$x_i(t+1) = \sum_{j \in \{i\} \cup \partial_i} w_{ij}\, x_j(t) \qquad t = 1, 2, \dots$$

$w_{ij} > 0$ if $(i,j) \in E$ and "0" otherwise.

$$\sum_{j \in \{i\} \cup \partial_i} w_{ij} = 1$$

In matrix form: $X(t+1) = W\, X(t)$

weight matrix $W = [w_{ij}]$

- W is $\underbrace{\text{row-stochastic}}_{\sum_j w_{ij} = 1}$, nonnegative, and $\underbrace{\text{irreducible}}_{G \text{ is strongly connected.}}$

So from Perron-Frobenius Theorem, eigenvalues of W are s.t.

$$\lambda_1 = |\lambda_1| \geq |\lambda_2| \geq \cdots \geq |\lambda_N|$$

It's easy to check that $\mathbb{1}$ is an eigenvector of W with

$\frac{1}{\sqrt{N}}[1 \ldots 1]^T$ beign the corresponding eigenvector. It follows from Gershgorin circle theorem that all the eigenvalues of a stochastic matrix have to be ≤ 1.

Gershgorin's circle theorem: Consider a matrix $A = [a_{ij}]_{1 \leq i \leq n,\ 1 \leq j \leq n}$. Let $R_i = \sum_{j \neq i} |a_{ij}|$ and $D_i = D(a_{ii}, R_i)$ be a closed disk at center a_{ii} with radius R_i. Then all the eigenvalues of A are confined within $D_1 \cup \ldots \cup D_n$.

Hence, $\lambda_1 = 1$ and $v_1 = \frac{1}{\sqrt{N}} \mathbb{1}_N$.

If we further ensure W is primitive, then

$$1 = \lambda_1 \underbrace{>}_{\text{strict inequality}} |\lambda_2| \geq |\lambda_3| \cdots \geq |\lambda_N|$$

Primitive matrix: means $\exists t_0$ s.t. $\forall t \geq t_0$, $W_{ij}^t > 0$ for all i, j.

- One way to ensure W is primitive is to add self-loops, i.e., $w_{ii} > 0$ for at least one i.

- Spectral Decomposition $W = V \Lambda \underbrace{V^{-1}}_{U}$ where

$$\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_N)$$

$V = [v_1 \ldots v_N]$, $U = [u_1 \ldots u_N]$

eigenvectors

By iterating,

$$x(t) = W^t x(0)$$

$$= (V \Lambda U)^t x(0)$$

$$= V \Lambda^t U \; x(0)$$

$$= \sum_{i=1}^{N} \lambda_i^t v_i u_i^T x(0)$$

$$= v_1 u_1^T x(0) + \sum_{i=2}^{N} \underbrace{\lambda_i^t}_{\to 0} v_i u_i^T x(0)$$

$\lambda_i^t \to 0$ as $t \to \infty$ since $|\lambda_i| < 1$

$$\to \frac{1}{\sqrt{N}} \mathbb{1}_N u_1^T x(0) \quad \text{as } t \to \infty$$

$\hookrightarrow$ every node reaches consensus

$$x_i(\infty) = u_1^T x(0) \quad \forall i$$

If we further assume W is symmetric, then

$$U = V^{-1} = V^T \implies u_1 = v_1 = \frac{1}{\sqrt{N}} \mathbb{1}_N$$

$$\implies x_i(\infty) = \frac{1}{\sqrt{N}} \mathbb{1}^T x(0) = x_{average}$$

- Speed of convergence is determined by $\lambda_2(W)$

Q: which one has a faster convergence? complete graph or ring graph?