Review: The discrete sampling of $\dot{x} = -L(D)x$, defined as $z(k) = x(k\delta)$, $\delta > 0$, and satisfying

$$z(k+1) = \underbrace{e^{-\delta L(D)}}_{\text{stochastic matrix (even, doubly stoch. if } D \text{ is balanced)}} z(k) \quad \text{converges} \iff x(t) \text{ does!}$$

Generalization for discrete Agreement protocol:

Discrete AP: $\quad x(t+1) = F\,x(t)$, given $x_0 \in \mathbb{R}^n$

suppose F is a primitive stochastic matrix.

Then $x(t) = F^t x(0) \longrightarrow \alpha^* \mathbb{1} \in \operatorname{span}\{\mathbb{1}\}$.

Background on Non-negative matrices:

we say: [see Horn & Johnson]

- F is irreducible if $\nexists\, P$ s.t. $PFP^{-1} = \begin{bmatrix} A & B \\ 0 & C \end{bmatrix}$ for some A, B, C matrices.
- F is primitive if it is irreducible and has only one eigenvalue of maximum modulus.

Lemma: If $A \geq 0$ then

$$A \text{ is primitive} \iff A^m > 0 \text{ for some } m \geq 1.$$

Lemma: If $A \geq 0$ and irreducible and all diagonal elements are positive $\Rightarrow A^{n-1} > 0 \Rightarrow A$ is primitive.

Thm: [Perron-Frobenius]

supp. $A \geq 0$ and irreducible. Then

- $\rho(A) \geq 0$ $\left(\rho(A) = \text{maximum modulus of eigenvalues of } A \right)$
- $\rho(A)$ is an algebraically simple eigenvalue of A.
- There exists unique positive eigenvectors x, y s.t.

$$Ax = \rho(A)x, \quad y^T A = \rho(A) y^T, \quad x^T y = 1.$$

- $$\lim_{N \to \infty} \frac{1}{N} \sum_{m=1}^{N} \left(\frac{1}{\rho(A)} \cdot A \right)^m = x y^T$$

Corollary: suppose F is a primitive stochastic matrix.

Then
- $\rho(F) = 1$ is a simple eigenvalue of F (w/ max. modulus)
- $\lim_{k \to \infty} F^k = \mathbb{1} y^T$ where $y^T F = y$, $y^T \mathbb{1} = 1$.
- if F is doubly stochastic, then $(y = \frac{\mathbb{1}}{n})$ i.e.

$$\lim_{k \to \infty} F^k = \frac{\mathbb{1}\mathbb{1}^T}{n}.$$

Therefore : Discrete AP :

$$X(t+1) = F\,X(t) \quad , \quad X(0) = \begin{bmatrix} x_1(0) \\ \vdots \\ x_n(0) \end{bmatrix}$$

F : primitive , stochastic and respects the sparsity of the underlying graph. (non-example $e^{-SL(D)}$)

Examples of primitive doubly stochastic F :

suppose $G = (V, E)$ is connected .

1) Normalized Laplacian of G :

$$F = D(G)^{-1} A(G)$$

where G has at least one self-loop .

2) discretized weighted Laplacian of G :

$$L_w(G) = D(G)\, W\, D(G)^T \quad , \quad W = \operatorname{diag}(\overbrace{w_1, \dots, w_m}^{\#\text{ of edges}})$$

$$\hookrightarrow \quad F_w = I - \Delta L_w(G)$$

where $\Delta < \dfrac{1}{\rho(L_w(G))}$.

Recall that

$\rho(L_w(G)) \geq \max_i d_w(v_i)$

so

$F_w \geq 0$.

more specifically, if we choose W s.t.

$$v_i \longrightarrow v_j \qquad w_{e_{ij}} = \frac{1}{\max\{d(v_i), d(v_j)\}}$$

Then, F_W is primitive doubly stochastic for any $\Delta \in (0,1)$.

3) metropolis weights:

$$F_{ij} = \begin{cases} \dfrac{1}{(1+\max\{d_i, d_j\})} & i,j \in E \\ 1 - \sum_{k \in N_i} F_{ik} & i = j \\ 0 & \text{o.w.} \end{cases}$$

So far, we talk about consensus on <u>scalar states</u>!

Before we move on, convince yourself how we can generalize this to <u>vector states</u> or even <u>matrix states</u>!

Answer: Do consensus element-wise!

Now, let's see some applications!

Distributed estimation:

observation matrix

measurement model: $z = H_{p \times q}\, \theta + v$

measurements

noise (independent)

$p > q$

Assume: $\text{rank}(H) = q$, i.e., we have at least q independent measurements.

Centeralized problem:

find "θ" that achieves the least square error, i.e.,

$$\hat{\theta} = \arg\min_{\theta} \left\{ J(\theta) = (z - H\theta)^T (z - H\theta) \right\}$$

J: differentiable, convex in θ $\Rightarrow$

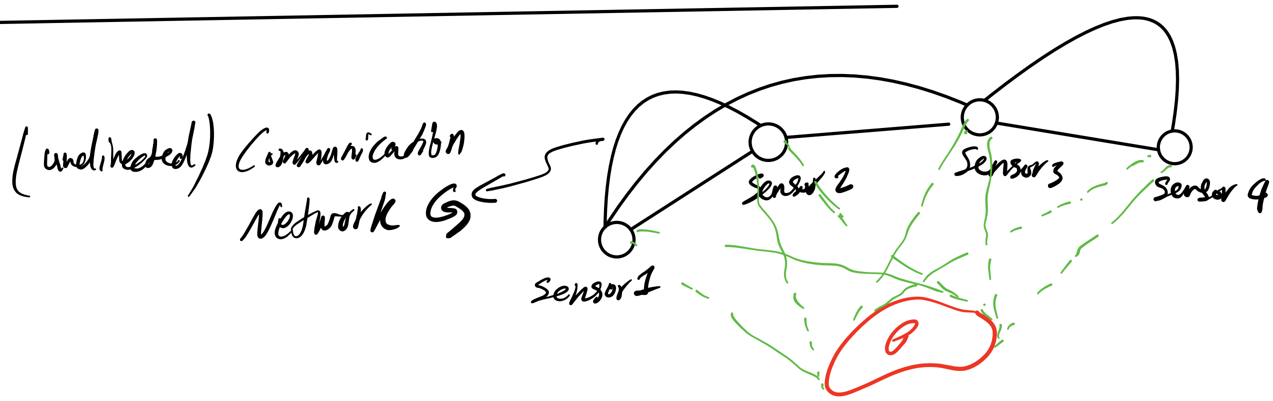
$$\hat{\theta} = (H^T H)^{-1} H^T z .$$

If $v \sim \mathcal{N}(0 I_p, \Sigma)$ then the optimal estimator minimizes

$$J(\theta) = (z - H\theta)^T \Sigma^{-1} (z - H\theta)$$

$$\Rightarrow \hat{\theta} = (H^T \Sigma^{-1} H)^{-1} H^T \Sigma^{-1} z .$$

Least square over Sensor Network



@ Sensor i:

$$z_i = H_i \theta + v_i$$

Centeralized problem:

$$\underbrace{\begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix}}_{z} = \underbrace{\begin{bmatrix} H_1 \\ H_2 \\ \vdots \\ H_n \end{bmatrix}}_{H} \theta + \underbrace{\begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}}_{v}$$

Assumption: H is full-row rank!

$$\Rightarrow \hat{\theta} = (H^T H)^{-1} H^T z = \left(\sum_{i=1}^{n} H_i^T H_i\right)^{-1} \left(\sum_{i=1}^{n} H_i^T z_i\right)$$

$$= \left(\tfrac{1}{n}\sum_{i=1}^{n} H_i^T H_i\right)^{-1} \left(\tfrac{1}{n}\sum_{i=1}^{n} H_i^T z_i\right)$$

Scalar case:

If θ is scalar and each $H_i = 1$ $\forall i$, then

$$\Rightarrow \hat{\theta} = \tfrac{1}{n} \sum_{i=1}^{n} z_i \,.$$

(we know how to achieve this using Discrete A.P.)

Vector case:

$$\hat{\theta} = \left(\frac{1}{n} \sum_{i=1}^{n} H_i^T H_i \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} H_i^T z_i \right)$$

Think about a situation where each node i, has two information: $P_i \triangleq H_i^T H_i \in \mathbb{R}^{q \times q}$ and $\tau_i = H_i^T z_i \in \mathbb{R}^q$.

→ we want to do Consensus on all P_i's and τ_i's,

because then

$$\theta = \left(\frac{1}{n} \sum_{i=1}^{n} P_i \right)^{-1} \left(\frac{1}{n} \sum_{i=1}^{n} \tau_i \right),$$

which is computable at each node.

Consensus on P_i and τ_i:

For example, we can use the "metropolis weights on G" as follows:

Recall: $F_{ij} = \begin{cases} \frac{1}{(1+\max\{d_i, d_j\})} & i,j \in E \\ 1 - \sum_{k \in N_i} F_{ik} & i = j \\ 0 & \text{o.w.} \end{cases}$ symm.

so, $X(k+1) = F\,X(k)$ generalizes to:

Distributed Estimation Algorithm

@ node i:

$$P_i(0) = H_i^T H_i \quad , \quad z_i(0) = H_i^T z_i$$

for each $k \geq 0$:

matrices

$$P_i(k+1) = P_i(k) + \sum_{j \in N_i} F_{ij} \left(P_j(k) - P_i(k) \right)$$

vectors

$$z_i(k+1) = z_i(k) + \sum_{j \in N_i} F_{ij} \left(z_j(k) - z_i(k) \right)$$

- once $P_i(k)$ becomes invertible, we also have a local estimate $\hat{\theta}_i(k) = \left[P_i(k)\right]^{-1} z_i(k)$.

$\underline{\ast}$ We now know that $\hat{\theta}_i(k) \longrightarrow \hat{\theta} = \underset{\theta}{\operatorname{argmin}}\, J(\theta)$

as long as G is connected.