Network Science
IST-CMU PhD course
Fall 2016
Instructor: João Xavier, `jxavier@isr.ist.utl.pt`
TA: João Martins, `joaoa@andrew.cmu.edu`

Important: the homework is due the October 12. Send a scanned pdf file with your answers (or typed in LaTeX, if you prefer) to the TA's email.

Homework 1

1. *Applications of consensus.* Consensus algorithms compute, in a distributed manner, the arithmetic mean of a dataset $\{\theta_i \in \mathbf{R} : i = 1, \ldots, n\}$ where θ_i is the data point held by agent i. Each agent initializes $x_i(0) = \theta_i$ and the consensus' iterates $x_i(t)$, $t = 1, 2, \ldots$, converge to $\overline{\theta} = (\theta_1 + \cdots + \theta_n)/n$ as $t \to \infty$, for all i.

We can compute other quantities from the dataset by properly initializing and processing the consensus' iterates. For example, suppose that the θ_i's are positive and we want to determine the geometric mean $g = (\theta_1 \theta_2 \cdots \theta_n)^{1/n}$. If agent i initializes $x_i(0) = \log \theta_i$ then the local estimates $g_i(t) := \exp(x_i(t))$ converge to g as $t \to \infty$, for all i.

(a) Assume that the agents measure a parameter of interest θ in additive noise: $x_i = \theta + v_i$ where v_i denotes zero-mean gaussian noise with variance σ_i^2. The observation noise v_i is independent across agents. The maximum-likelihood estimate of θ, given the network measurements $x_1, x_2, \ldots, x_n$, is

$$\widehat{\theta}_{\mathrm{ML}} = \frac{\frac{1}{\sigma_1^2} x_1 + \cdots + \frac{1}{\sigma_n^2} x_n}{\frac{1}{\sigma_1^2} + \cdots + \frac{1}{\sigma_n^2}}.$$

Each agent i only knows its measurement x_i and its noise power σ_i^2.

Show how to use a consensus algorithm—or several in parallel, if needed—to obtain $\widehat{\theta}_{\mathrm{ML}}$ at each agent. State clearly how each consensus is initialized and what processing is done to form local estimates $\theta_i(t)$ that converge to $\widehat{\theta}_{\mathrm{ML}}$ as $t \to \infty$, for all i.

(b) Assume now that the parameter of interest is $\theta \in \mathbf{R}^p$ and the observation model is $x_i = H_i \theta + v_i$ where $H_i \in \mathbf{R}^{m_i \times p}$ and $v_i \in \mathbf{R}^{m_i}$ denotes zero-mean gaussian noise with covariance $\sigma^2 I$. The observation noise v_i is independent across agents. The maximum-likelihood estimate of θ, given the network measurements $x_1, x_2, \ldots, x_n$, is

$$\widehat{\theta}_{\mathrm{ML}} = \left(H_1^T H_1 + \cdots + H_n^T H_n\right)^{-1} \left(H_1^T x_1 + \cdots + H_n^T x_n\right).$$

Each agent i only knows its measurement x_i and its measurement matrix H_i.

Show how to use a consensus algorithm—or several in parallel, if needed—to obtain $\widehat{\theta}_{\mathrm{ML}}$ at each agent. State clearly how each consensus is initialized and what processing is done to form local estimates $\theta_i(t)$ that converge to $\widehat{\theta}_{\mathrm{ML}}$ as $t \to \infty$, for all i.

2. *Primative matrix.* Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a connected undirected graph. Let $W = I - \alpha L$ where L is the laplacian of $\mathcal{G}$ and $0 < \alpha < 1/\mu_n$, where $0 \leq \mu_1 \leq \cdots \leq \mu_n$ are the eigenvalues of L. Show that W is a primitive matrix.

3. *Dynamic consensus.* Each agent i measures a time-varying parameter $\theta_i(t)$ for $t = 1, 2, \ldots$, and the n agents want to track the time-varying average

$$\overline{\theta}(t) := \frac{1}{n} \sum_{i=1}^{n} \theta_i(t).$$

Assume that all parameter increments are bounded:

$$|\underbrace{\theta_i(t+1) - \theta_i(t)}_{=:\delta_i(t+1)}| \leq B \tag{1}$$

for all i and t.

Let the undirected graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ represent the communication system, where nodes correspond to agents and edges correspond to channels. Let $W \in \mathbf{R}^{n \times n}$ be a given symmetric matrix that conforms to the graph ($W_{ij} = 0$ whenever $i \not\sim j$) and secures consensus: $W\mathbf{1} = \mathbf{1}$ and $\rho\,(W - J) < 1$ with $J = (1/n)\mathbf{1}\mathbf{1}^T$. The agents run the distributed algorithm

$$x(t+1) = Wx(t) + \delta(t+1) \tag{2}$$

where $x(t) = (x_1(t), \ldots, x_n(t))$ is the network state vector, $x_i(t)$ is the estimate of $\overline{\theta}(t)$ at agent i and time t) and $\delta(t) = (\delta_1(t), \ldots, \delta_n(t))$ collects the agents' increments, *cf.* (1). In the update (2), the agents fuse their previous estimate with neighbors and add their observed increment.

The goal of this problem is to show that the estimation errors

$$e_i(t) = x_i(t) - \overline{\theta}(t) \tag{3}$$

remain bounded for all t. Note that, if the agents don't cooperate—each agent takes its measurement $\theta_i(t)$ as the estimate $\overline{\theta}(t)$—the error can grow unbounded.

(a) Show that J and $I - J$ are symmetric and idempotent[1].

(b) Show that both J and $I - J$ commute with W: $JW = WJ = J$ and $(I - J)W = W(I - J)$.

(c) For $x \in \mathbf{R}^n$, consider the orthogonal decomposition

$$x = \overline{x}\mathbf{1} + \widetilde{x}, \tag{4}$$

where $\overline{x} = \mathbf{1}^T x/n$. Note that $\overline{x}\mathbf{1} = Jx$ and $\widetilde{x} = (I - J)x$.
Let $y = Wx$ be the result of acting with W on x and decompose y as in (4): $y = \overline{y}\mathbf{1} + \widetilde{y}$.
Show that

$$\overline{y} = \overline{x} \quad \text{and} \quad \widetilde{y} = \widetilde{W}\widetilde{x},$$

where $\widetilde{W} = (I - J)W(I - J)$. (This means that the action of W on x preserves the component $\overline{x}\mathbf{1}$ and changes the component $\widetilde{x}$ to $\widetilde{W}\widetilde{x}$.)

[1] A matrix X is idempotent if $X^2 = X$.

(d) Show that $\rho(W - J) = \rho\left(\widetilde{W}\right)$, where $\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}$ is the spectral radius of matrix A.

(e) Let $e(t) = x(t) - J\theta(t)$ and note that the ith entry of $e(t)$ corresponds to (3). Show that the recursion $e(t+1) = We(t) + (I - J)\delta(t+1)$ holds.

(f) Show that $\|\overline{e}(t+1)\mathbf{1}\| = \|\overline{e}(t)\mathbf{1}\|$ and $\|\widetilde{e}(t+1)\| \leq \rho\left(\widetilde{W}\right)\|\widetilde{e}(t)\| + \sqrt{n}B$.

(g) Conclude that

$$\|e(t)\| \leq \sqrt{\|\overline{e}(0)\mathbf{1}\|^2 + \left(\|\widetilde{e}(0)\| + \frac{\sqrt{n}B}{1 - \rho\left(\widetilde{W}\right)}\right)^2}$$

for all $t \geq 1$.

4. *Ongoing observations.* Agents take repeated measurements of a static parameter: $y_i(t) = \theta + v_i(t)$, where $y_i(t)$ is the observation at agent i and time $t = 1, 2, \ldots$; θ is the parameter of interest; and $v_i(t)$ denotes zero-mean bounded noise ($|v_i(t)| \leq B$ for all t and i). The noise is assumed independent across agents and time. A reasonable estimate of θ at time t, given *all* the network measurements from time 1 to time t, is

$$\widehat{\theta}(t) = \frac{1}{nt}\sum_{s=1}^{t}\sum_{i=1}^{n} y_i(s). \tag{5}$$

The agents want to track $\widehat{\theta}(t)$.

As in problem 2, assume that $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ represents the communication system and W is a symmetric matrix that conforms to the graph and secures consensus. The agents run the distributed algorithm

$$x(t+1) = W\left(\frac{t}{t+1}x(t) + \frac{1}{t+1}y(t+1)\right) \tag{6}$$

where $x(t) = (x_1(t), \ldots, x_n(t))$ is the network state vector ($x_i(t)$ is the estimate of $\widehat{\theta}(t)$ at agent i and time t) and $y(t) = (y_1(t), \ldots, y_n(t))$ collects the network measurements at time t. The algorithm is initialized as $x(0) = y(0)$.

The goal of this problem is to show that the errors

$$e_i(t) = x_i(t) - \widehat{\theta}(t) \tag{7}$$

converge to zero as $t \to \infty$. This means that, asymptotically, the estimate at each agent is as good as the estimate of a central node that sees all network measurements, instantaneously—collaboration empowers everyone (as in real life).

(a) Let $a(t) = (1/t)\sum_{s=1}^{t} y(s)$ denote the running average of the network measurements. Show that

$$a(t+1) = \frac{t}{t+1}a(t) + \frac{1}{t+1}y(t+1)$$

for $t \geq 1$.

(b) Recall the properties and notation introduced in parts (a)–(e) of problem 3. Let

$$e(t) = x(t) - \overline{a}(t)\mathbf{1}$$

for $t \geq 1$ (note that $\overline{a}(t)\mathbf{1} = Ja(t)$ as defined in part (c) of problem 2). Show that the ith entry of $e_i(t)$ corresponds to (7). (Thus, the goal of this problem is to show that the error $e(t)$ converges to zero as $t \to \infty$. We will prove that both components $\overline{e}(t)\mathbf{1}$ and $\widetilde{e}(t)$ converge to zero.)

(c) Show that

$$e(t+1) = \frac{t}{t+1} W e(t) + \frac{1}{t+1} (W - J)\, y(t+1)$$

for $t \geq 1$.

(d) Show that

$$\overline{e}(t+1) = \frac{t}{t+1} \overline{e}(t)$$

for $t \geq 1$ and conclude that $\overline{e}(t)$ converges to zero as $t \to \infty$.

(e) Show that

$$\widetilde{e}(t+1) = \frac{t}{t+1} \widetilde{W} \widetilde{e}(t) + \frac{1}{t+1} \widetilde{W} y(t+1) \tag{8}$$

for $t \geq 1$.

(f) Consider a recursion

$$z(t+1) = \alpha(t) z(t) + \beta(t+1), \quad t = 0, 1, \ldots,$$

where the sequence $\alpha(t) \in \mathbf{R}$ converges to α with $|\alpha| < 1$ and the sequence $\beta(t) \in \mathbf{R}$ converges to zero. The sequence $z(t)$ can be interpreted as the output of a first-order time-varying filter that is fed by a decaying input $\beta(t)$; the time-variant filter is asymptotically stable (the magnitude of the gains $\alpha(t)$ get sub-unit).

Show that $z(t)$ converges to zero. Hint: start by analyzing the case of fixed gains $\alpha(t) \equiv \alpha$ and a bounded input $|\beta(t)| \leq \epsilon$.

(g) Use part (f) and equation (8) to conclude that $\widetilde{e}(t)$ converges to zero.