Distributed Autonomous Systems M

Dante Piotto

spring semester 2024

Contents

Chapter 1

Introduction and scenarios

1.1 Distributed Autonomous System

Each agent $i \in \{1, \dots, N\}$ has

- local physical and/or cyber state x_i
- computational and sensing capabilities
- communication capability: exchange messages with "neighbours"

1.2 Scenarios and applications of distributed systems

- Averaging: distributed estimation, opinion dynamics
- Distrbuted control in cooperative robotics
- Distributed optimization
  - distributed machine learning
  - distributed decision-making in cooperative robotics
  - distributed optimal control in energy systems and cooperative robotics

1.3 Measurement filtering in wireless sensor networks

Consider a network of N sensors with local sensing, computation and communication. Agent $i, i \in \{1, \dots, N\}$, takes a local measurement from the environment (temperature, pressure, etc.). Let $x_{i0} \in \mathbb{R}$ be the scalar local measurement Agents are interested in agreeing on the average of the measurements,

$$x_{\text{avg}} = \frac{1}{N} \sum_{i=1}^{N} x_{i0}$$

to have a better estimate of the environment quantity

Consider the following "distributed algorithm" based on "local" linear averaging, for each $i \in \{1, \dots, N\}$

$$\begin{aligned} x_i^0 &= x_{i0} \\ x_i^{k+1} &= \text{average}(x_i^k, \{x_j^k, j \text{ "neighbour" of } i\}), \qquad k \in \mathbb{N} \end{aligned}$$

generalizing coefficients of the update:

$$\begin{aligned} x_i^0 &= x_{i0} \\ x_i^{k+1} &= \sum_{j=1}^{N} a_{ij} x_j^k \qquad k \in \mathbb{N} \end{aligned}$$

Remark. $a_{ij} \geq 0$ and $\sum_{j=1}^{N} a_{ij} = 1$

Remark. $a_{ij} = 0$, for some $j \in \{1, \dots, N\}$, i.e. $a_{ij} = 0$ if i does not have access to the estimate of j

1.4 Parameter Estimation in Wireless Sensor Networks

Consider a network of N sensors with local sensing, computation and communication aiming at estimating a common parameter $\theta^* \in \mathbb{R}$ Each sensor i measures

$$y_i = B_i \theta^* + v_i$$

with $y_i \in \mathbb{R}^{m_1}, B_i$ known matrix and v_i a random measurement noise. Assume $v_1, \dots, v_N$ independent and Gaussian, with zero mean and covariance $E[v_i v_i^T] = \Sigma_i$. Assume $\sum_{i=1}^{N} m_i \geq m$ and $\begin{bmatrix} B_1 \\ \vdots \\ B_N \end{bmatrix}$ full rank Compute a least-squares estimate

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{N} (y_i - B_i\theta)^T \Sigma_i^{-1} (y_i - B_i\theta)$$

The optimal solution is

$$\begin{aligned} \hat{\theta} &= \left(\sum_{i=1}^{N} B_i^T \Sigma_i^{-1} B_i \right)^{-1} \sum_{i=1}^{N} B_i^T \Sigma_i^{-1} y_i \\ &= \left(\frac{1}{N} \sum_{i=1}^{N} B_i^T \Sigma_i^{-1} B_i \right)^{-1} \frac{1}{N} \sum_{i=1}^{N} B_i^T \Sigma_i^{-1} y_i \end{aligned}$$

The optimal solution can be obtained by computing two averages $\frac{1}{N} \sum_{i=1}^{N} \beta_i$ and $\frac{1}{N} \sum_{i=1}^{N} \beta_i$

1.4.1 Opinion Dynamics in Social Influence Networks

Group of N individuals, with x_i^k being the opinion of individual i at time k. Opinions are updated according to

$$x_i^{k+1} = \sum_{j=1}^{N} a_{ij} x_j^k$$

1.5 Main questions in averaging algorithms

- Do node estimates converge? Do they converge to a common value ("reach consensus")?
- Do they reach consensus to the average ("average consensus")?
- How can we model communication in general networks?
- Can we answer the above questions for general networks and communication protocols?
- What assumptions do we need on the communication network?

1.6 Distributed control in cooperative robotics

Team of N (mobile) robots aiming to execute complex tasks

Basic tasks

- rendevous, containment
- formation, flocking
- coverage

Complex tasks

- pickup and delivery
- surveillance and patrolling
- exploration
- satellite constellation

1.6.1 Main questions in cooperative robotics

- Do robot states asymptotically converge?
- Do the asympototic staes satisfy the global, desired task?
- How can we model communication in (general) robotic networks?
- What assumptions do we need on the communication network?
- Can we answer the above questions for general networks and communication protocols?

1.6.2 Distributed optimal control

$$\begin{aligned}
\min_{\substack{x_1,\dots,x_N \\ u_1,\dots,u_N}} & \sum_{i=1}^{N} \Big(\sum_{\tau=0}^{T-1} \ell_i(z_{i,\tau}, u_{i,\tau}) + m_i(z_{i,T}) \Big) \\
\text{subj to} & \sum_{i=1}^{N} H_i z_{i,\tau} \leq h, && \tau \in [0,T] \\
& z_{i,\tau+1} = A_i z_{i,\tau} + B_i u_{i,\tau} && \forall i, \tau \in [0,T] \\
& z_{i,\tau} \in Z_i, \quad u_{i,\tau} \in U_i, && \forall i, \tau \in [0,T]
\end{aligned}$$

Chapter 2

Preliminaries on Algebraic Graph Theory

Definition 2.1 (Digraph)
A digraph is a pair $G = (I, E)$ where $I = 1, \dots, N$ is a set of elements called *nodes* and $E \subset I \times I$ is a set of ordered node pairs called *edges*
Edge: the pair (i, j) denotes an edge from i to j
Self-loop: edge from a node to itself, i.e. (i, i)

Definition 2.2 (Undirected (di)graph)
if for any $(i, j) \in E$ then $(j, i) \in E$

Definition 2.3 (Subgraph)
(I', E') subgraph of (I, E) if $I' \subset I$ and $E' \subset E$. Spanning subgraph if $I' = I$

Definition 2.4 (In-neighbours of i)
$j \in I$ is an in-neighbour of $i \in I$ if $(j, i) \in E$

Definition 2.5 (Set of in-neighbours of i)
$\mathcal{N}_i^{\mathrm{IN}} = \{j \in \{1, \dots, N\} | (j, i) \in E\}$

Definition 2.6 (Out-neighbours of i)
$j \in I$ is an out-neighbour of $i \in I$ if $(i, j) \in E$

Definition 2.7 (Set of out-neighbours of i)
$\mathcal{N}_i^{\mathrm{IN}} = \{j \in \{1, \dots, N\} | (i, j) \in E\}$

Definition 2.8 (In-degree $\deg_i^{\mathrm{IN}}$)
number of in-neighbours, i.e. carinality of $\mathcal{N}_i^{\mathrm{IN}} (\deg_i^{\mathrm{IN}} = |\mathcal{N}_i^{\mathrm{IN}}|)$
Out-degree analogous

Definition 2.9 (Balanced digraph)
A digraph G is balanced if $\deg_i^{\mathrm{IN}} = \deg_i^{\mathrm{OUT}}$ for all $i \in \{1, \dots, N\}$

Definition 2.10 (Complete graph)
Unweighted graph such that $\forall i, j \; \exists \; (i, j), \; (j, i) \in E$

Chapter 3

Averaging Systems

3.1 Distributed algorithm

Given a network of N agents communicating according to a fixed digraph G, i.e. each agent i can receive messages only from in-neighbours in the graph, i.e. from $j \in \mathcal{N}_i^{\text{IN}}$. We start by considering a fixed graph, thus, each agent communicates with the same neighbours at each iteration $k \in \mathbb{N}$

$$x_i^{k+1} = \text{stf}_i(x_i^k, \{x_j^k\}_{j \in \mathcal{N}_i^{\text{IN}}}), \qquad i \in \{1, \dots, N\}$$

where stf_i is a function depending only on state x_i and states $x_j, j \in \mathcal{N}_i^{\text{IN}}$.

Alternative version with out-neighbours:

$$x_i^{k+1} = \text{stf}_i(\{x_j\}_{j \in \mathcal{N}_i^{\text{OUT}}})$$

3.2 Discrete-time averaging systems

Let $G^{\text{comm}} = (I, E)$ be a fixed (communication) digraph (self loops included). A linear averaging distributed algorithm can be written as:

$$x_i^{k+1} = \sum_{J \in \mathcal{N}_i^{\text{IN}}} a_{ij} x_j^k \qquad i \in \{1 \dots, N\}$$

where $x_i^k \in \mathbb{R}$ is the state of agent i at k and $a_{ij} > 0$ are positive weights.

Remark. The weights a_{ij} are defined only for $(i, j) \in E$

Wach i uses only the states of neighbours $j \in \mathcal{N}_i^{\text{IN}}$, thus distributed algorithm.

For analysis purposes, let us define weights $a_{ij} = 0$ for $(j, i) \notin E$. Thus we can rewrite the distributed algorithm as

$$x_i^{k+1} = \sum_{j=1}^{N} a_{ij} x_j^k \qquad i \in \{1, \dots, N\}$$

This is a LTI autonomous system

$$\begin{bmatrix} x_1^{k+1} \\ \vdots \\ x_N^{k+1} \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1N} \\ \vdots & \ddots & \vdots \\ a_{N1} & \cdots & a_{NN} \end{bmatrix} \begin{bmatrix} x_1^k \\ \vdots \\ x_N^k \end{bmatrix}$$

Which can be compactly written as

$$x^{k+1} 0 A x^k$$

Remark. The matrix A can be seen as the weighted adjacency matrix of the reverse digraph $G^{\text{comm,rev}}$ of the digraph G^{comm}

If instead of in-neighbours we use out-neighbours, we call the digraph a sensing digraph G^{sens}. In this case the notation becomes consistent with graph theory, so we get

$$x^{k+1} = Ax^k$$

where A can be seen as the weighted adjacency matrix of the sensing digraph G^{sens}

3.3 Stochastic matrices

The non-negative square matrix $A \in \mathbb{R}^{N \times N}$ is

- row stochastic if $A\mathbf{1} = \mathbf{1}$ (each row sums to 1)
- column stochastic if $A^\top \mathbf{1} = \mathbf{1}$ (each column sums to 1)
- doubly stochastic if both row and column stochastic.

Lemma. Let A be a row-stochastic matrix and G the associate digraph. If G is strongly connected and aperiodic, then

1. the eigenvalue $\lambda = 1$ is simple;
2. all the other eigenvalues μ satisfy $|\mu| < 1$

Remark. The condition "G contains a globally reachable node and the subgraph of globally reachable noes is aperiodic" is necessary and sufficient

Theorem 3.1 (Consensus)
Consider a (discrete-time) averaging system with associated digraph G and wieghted adjacency matrix A. Assume G is strongly connected and aperiodic, and A is row stochastic. Then

1. there exists a left eigenvector $w \in \mathbb{R}^N, w > 0$ (i.e. with positive components $w_i > 0$ for all $i = 1, \dots, N$) such that

$$\lim_{k \to \infty} x^k = \mathbf{1} \frac{w^\top x^0}{w^\top \mathbf{1}} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \frac{\sum_{i=1}^N w_i x_i^0}{\sum_{i=1}^N w_i}$$

   i.e., consensus is reached to $\dfrac{\sum_{i=1}^N w_i x_i^0}{\sum_{i=1}^N w_i}$

2. if additionally A is doubly stochastic, then

$$\lim_{k \to \infty} x^k = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} \frac{\sum_{i=1}^N x_i^0}{N}$$

   i.e., average consensus is reached

3.4 Example: Metropolis-Hastings weights

Given an undirected unweighted graph G with edge set E and degrees $d_1, \dots, d_n$

$$a_{ij} = \begin{cases} \dfrac{1}{1 + \max\{d_i, d_j\}} & \text{if } (i,j) \in E \text{ and } i \neq j \\ 1 - \displaystyle\sum_{h \in \mathcal{N}_i \setminus \{i\}} a_{ih} & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Result: the matrix A is symmetric and doubly-stochastic.

3.5 Time-varying digraphs

A time-varying digraph is a sequence of digraphs $\{G(k)\}_{k\geq 0}$.

Remark. The main definitions of in/out neighbours, in/out degree, adjacency matrix can be generalized by considering time-varying versions, i.e. $\mathcal{N}_i^{\text{IN}}(k)$, $\mathcal{N}_i^{\text{OUT}}(k)$, $\deg_i^{\text{IN}}(k)$,$\deg_i^{\text{OUT}}(k)$, $A(k)$ associated to each graph $G(k)$. Connectivity requires new definitions as assuming each $G(k)$ to be connected is too conservative.

Definition 3.1 (Jointly strongly connected digraph)
if $\bigcup_{\tau=k}^{+\infty} G(\tau)$ is strongly connecetd $\forall k \geq 0$

Definition 3.2 (Uniformly jointly strongly connected (or B-strongly connected) digraph)
if there exists $B \in \mathbb{N}$ such that $\bigcup_{\tau=k}^{k+B} G(\tau)$ is strongly connecetd $\forall k \geq 0$

Remark. The graph can be disconnected at some time k.

3.5.1 Averaging distributed algorithms over time-varying graphs

Let $\{G(k)\}_{k\geq 0}$ be a time-varying digraph (with self loops for each $G(k)$). Consider the distributed algorithm

$$x_i^{k+1} = \sum_{j\in\mathcal{N}_i^{\text{IN}}(k)} a_{ij}(k)x_j^k \qquad \forall i \in \{1,\dots,N\}$$

or the out-neighbours version

$$x_i^{k+1} = \sum_{j\in\mathcal{N}_i^{\text{OUT}}(k)} a_{ij}(k)x_j^k \qquad \forall i \in \{1,\dots,N\}$$

where $x_i^k \in \mathbb{R}$ is the state of agent i at k and $a_{ij}(k) > 0$.

For analysis purposes, let us define weights $a_{ij}(k) = 0$ for $(i,j) \notin E(k)$. Thus we can rewrite the distributed algorithm as

$$x_i^{k+1} = \sum_{j=1}^{N} a_{ij}(k)x_j^k \qquad i \in \{1,\dots,N\}$$

This is a Linear Time-Varying system

$$x^{k+1} = A(k)x^k$$

with state $x := [x_1,\dots,x_N]^\top$ and state matrix

$$A(k) := \begin{bmatrix} a_1 1k & \cdots & a_{1N}(k) \\ \vdots & \ddots & \vdots \\ a_N 1k & \cdots & a_{NN}(k) \end{bmatrix}$$

being a weighted adjacency matrix associated to the digraph $G(k)$.

3.5.2 Discrete-time consensus over time-varying graphs

Theorem 3.2
Let $\{A(k)\}_{k\geq 0}$ be a sequence of row-stochastic matrices with associated digraphs $\{G(k)\}_{k\geq 0}$. Assume

1. each digraph $G(k)$ has a self-loop at each node;
2. each non-zero edge weight $a_{ij}(k)$, including the self-loop wights $a_{ii}(k)$, is larger than a constant $\epsilon > 0$;
3. there exists $B \in \mathbb{N}$ such that, for all times $k \geq 0$, the union digraph $G(k) \cup \cdots \cup G(k+B)$ is strongly connected.

Then

1. there exists a non-negative vector $w \in \mathbb{R}^N$ such that the solution to $x^{k+1} = A(k)x^k$ converges (exponentially) to $\mathbf{1}\dfrac{w^\top x^0}{w^\top \mathbf{1}}$, i.e.
$$\lim_{k\to\infty} x^k = \mathbf{1}\left(\frac{w^\top x^0}{w^\top \mathbf{1}}\right)$$
2. if additionally each matrix in the sequence is doubly-stochastic, then
$$\lim_{k\to\infty} x^k = \mathbf{1}\frac{1}{N}\sum_{i=1}^{N} x_i^0$$
i.e., average consensus is achieved

3.6 Laplacian dynamics

Consider a network of dynamical systems with dynamics

$$\dot{x}(t) = u_i(t) \qquad i \in \{1, \dots, N\}$$

with states $x_i \in \mathbb{R}$ and inputs $u_i \in \mathbb{R}$, communicating (or interacting) according to a digraph $G = (\{1, \dots, N\}, E)$. Consider a (distributed) "proportional" feedback control

$$u_i(t) = -\sum_{j\in\mathcal{N}_i^{\mathrm{IN}}} a_{ij}(x_i(t) - x_j(t))$$

or the out-neighbour version

$$u_i(t) = -\sum_{j\in\mathcal{N}_i^{\mathrm{OUT}}} a_{ij}(x_i(t) - x_j(t))$$

For analysis purposes, let us define weights $a_{ij}(k) = 0$ for $(i, j) \notin E(k)$. Thus we can rewrite the distributed control systems as

$$\dot{x}_i(t) = -\sum_{j=1}^{N} a_{ij}(x_i(t) - x_j(t)) \qquad \forall i \in \{1, \dots, N\}$$

Defining $x := [x_1 \cdots x_N]^\top$, it can be shown that it can be rewritten as the following Linear Time Invariant continuous-time system

$$\dot{x}(t) = -Lx(t)$$

where L is the (weighted) Laplacian associated to the digraph G with (weighted) adjacency matrix A

Let

$$\dot{x}_i(t) = -\sum_{j=1}^{N} a_{ij}(x_i(t) - x_j(t)) \qquad \forall i \in \{1, \dots, N\}$$

rearranging terms

$$\dot{x}_i(t) = -\left(\sum_{j=1}^{N} a_{ij}\right) x_i(t) + \sum_{j=1}^{N} a_{ij}x_j(t) = -\deg_i^{\mathrm{OUT}} x_i(t) + (Ax(t))_i$$

where $(Ax(t))_i$ is the i-th element of $Ax(t)$. Writing the previous dynamics in a compact form

$$\dot{x}(t) = -(D^{\mathrm{OUT}} - A)x(t)$$

where we recall that D^{OUT} is the (weighted) out-degree matrix. Recalling that $L = D^{\mathrm{OUT}} - A$, it holds that

$$\dot{x}(t) = -Lx(t)$$

Remark. if the in-neighbours version is considered, then $\dot{x}(t) = -L^{\mathrm{IN}}x(t)$, where $L^{\mathrm{IN}} = D^{\mathrm{IN}} - A^T$ is the in-degree Laplacian (i.e. the Laplacian of the reverse graph of G)

3.6.1 Properties of the Laplacian matrix

It can be easily verified that

$$L\mathbf{1} = D^{\mathrm{OUT}}\mathbf{1} - A\mathbf{1} = \begin{bmatrix} \deg_1^{\mathrm{OUT}} \\ \vdots \\ \deg_i^{\mathrm{OUT}} \end{bmatrix} - \begin{bmatrix} \deg_1^{\mathrm{OUT}} \\ \vdots \\ \deg_i^{\mathrm{OUT}} \end{bmatrix} = 0$$

i.e., $\lambda = 0$ is an eigenvalue of L and $\mathbf{1}$ is an associated eigenvector.

Lemma. Given a weighted digraph with Laplacian L, then all eigenvalues of L different from zero have strictly positive real part

Lemma. Given a weighted digraph with Laplacian L, the following statements are equivalent:

1. G is weight-balanced, i.e. $D^{\mathrm{IN}} = D^{\mathrm{OUT}}$
2. $\mathbf{1}L = 0$

Theorem 3.3
A weighted digraph with Laplacian L contains a globally reachable node if and only if $\lambda = 0$ is simple.

Corollary. If a weighted digraph is strongly connected, then $\lambda = 0$ is simple

3.6.2 Consensus for Laplacian dynamics

Theorem 3.4
let L be a (weighted) Laplacian matrix with associated strongly connected (weighted) digraph G. Consider the Laplacian dynamics $\dot{x}(t) = -Lx(t),\ t \geq 0$, then

1. $$\lim_{t\to\infty} x(t) = \mathbf{1}\left(\frac{w^\top x(0)}{w^\top \mathbf{1}}\right)$$
   with $w^\top L = 0$, i.e. w is a left eigenvector for the eigenvalue $\lambda = 0$;
2. if additionally G is weight-balanced then
   $$\lim_{t\to\infty} x(t) = \mathbf{1}\frac{\sum_{i=1}^N x_i(0)}{N}$$