Non-negative Matrices and Distributed Control

Yilin Mo

July 2, 2015

We model a network composed of m agents as a graph $G = \{V, E\}$. $V = \{1, 2, \ldots, m\}$ is the set of vertices representing the agents. $E \subseteq V \times V$ is the set of edges. $(i, j) \in E$ if and only if j can send information to i. The graph can be directed.

Define the neighbors $\mathcal{N}_i$ of agent i as the set of agents who can send information to i, i.e.,

$$\mathcal{N}_i \triangleq \{j : (i, j) \in E,\, j \neq i\}.$$

Suppose each agent has a state $x_i(t)$. The agent update the state based on the following update equation:

- *Continuous time:*

$$\frac{d}{dt}x_i(t) = a_{ii}x_i(t) + \sum_{j\in\mathcal{N}_i} a_{ij}x_j(t).$$

- *Discrete time:*

$$x_i(t+1) = a_{ii}x_i(t) + \sum_{j\in\mathcal{N}_i} a_{ij}x_j(t).$$

We can write everything in matrix form:

- *Continuous time:*

$$\frac{d}{dt}x(t) = Ax(t).$$

- *Discrete time:*

$$x(t+1) = Ax(t).$$

Question: is the system stable? Can we know the answer in a distributed fashion?

1 Some Definitions

Let $\mathbb{R}_+^{n\times m}$ and $\mathbb{R}_{++}^{n\times m}$ be convex cones defined as

$$\mathbb{R}_+^{n\times m} \triangleq \{M \in \mathbb{R}^{n\times m} : M_{ij} \geq 0, \forall i, j\}$$
$$\mathbb{R}_{++}^{n\times m} \triangleq \{M \in \mathbb{R}^{n\times m} : M_{ij} > 0, \forall i, j\}$$

Hence, we can define

$$X \geq Y \Leftrightarrow X - Y \in \mathbb{R}_{+}^{n\times m},$$
$$X > Y \Leftrightarrow X - Y \in \mathbb{R}_{++}^{n\times m}$$

Moreover, we define $Q \triangleq \mathbb{R}_{+}^{n} \backslash \{0\}$.

A matrix A is called positive if $A > 0$. It is called non-negative if $A \geq 0$. It is called a Metzler matrix if all the off-diagonal entries are non-negative, i.e., $A = B - sI$, where B is non-negative.

Some observations:

- If $A \geq 0$ and $x \geq 0$, then $Ax \geq 0$.
- On the contrary, if for all $x \geq 0$ and $Ax \geq 0$, then $A \geq 0$.
- Similarly, if $A > 0$ and $x \in Q$, then $Ax > 0$.
- On the contrary, if for all $x \in Q$ and $Ax > 0$, then $A > 0$.
- If A is Metzler, then $exp(At)$ is non-negative.

A matrix A is called Hurwitz if all its eigenvalues have strictly negative real part. A is called stable if all its eigenvalues satisfy $|\lambda| < 1$.

A non-negative matrix is called primitive if there exists an k, such that A^k is positive.

A non-negative matrix is called irreducible if for any i, j, there exists an k, such that $\left(A^k\right)_{ij}$ is positive.

In general, a matrix A is called irreducible if $|A|$ is irreducible.

Define $G(A) = (V, E)$ aw the graph associated with A, where $V = \{1, \ldots, n\}$ and $(i, j) \in E$ if and only if $a_{ij} \neq 0$.

Let the period of a vertex i to be the greatest common divisor of the lengths of all cycles starting from i.

Some observations:

- If A is irreducible then $A + I$ is primitive.
- $(A^k)_{ij} > 0$ if and only if there exists a path of length k from j to i.
- A is irreducible is equivalent to $G(A)$ to be strongly connected.
- If $G(A)$ is strongly connected, then all the vertices have the same period.
- A is primitive if A is irreducible and $G(A)$ has period 1 (aperiodic).

2 Important Properties of Non-negative matrices and Metzler matrices

Theorem 1 (Perron Frobenius Theorem)**.** *Let A be an irreducible matrix, then the following propositions hold:*

1. *Let the spectral radius of A to be $\rho(A)$, then there exists an eigenvalue λ of A, such that $\lambda = \rho(A)$.*

2. *λ has geometric and algebraic multiplicity of 1.*

3. *The left and right eigenvectors of λ is strictly positive. Any other eigenvector has negative entries.*

4. *If A is primitive, then all the other eigenvalues satisfy $|\lambda| < \rho(A)$.*

5. *$\rho(A)$ satisfies:*
$$\min_i \sum_j a_{ij} \leq \rho(A) \leq \max_i \sum_i a_{ij}.$$

Proof. First let us define the following function $L : Q \to \mathbb{R}_+$:

$$L(x) \triangleq \max\{s : sx \leq Ax\}.$$

Clearly $L(\alpha x) = L(x)$ for any $\alpha > 0$. Define $P = (I + A)^k$, where k is large enough such that P is positive. Hence, if $sx \leq Ax$,

$$P(sx) \leq PAx = APx,$$

which implies that
$$L(Px) \geq L(x).$$
Furthermore, if $L(x)x \neq Ax$, then $L(Px) > L(x)$.

Now define:
$$\lambda \triangleq \max\{L(x) : \|x\|_2 = 1, \, x \in Q\}.$$

Suppose λ is achieve at v. Then $\lambda v = Av$. (otherwise $L(Pv) \geq L(v)$.) Hence, λ is an eigenvalue of A with a positive eigenvector v.

Applying the same procedure to A^T, since the spectral radius of A is the same as A^T, we can find a strictly positive left eigenvector of A. Let us denote it as w.

Now let $\mu \neq \lambda$ be an eigenvalue of A with eigenvector y. Then

$$w^T Ay = \lambda w^T y = \mu w^T y.$$

Hence, $w^T y = 0$, which implies that y must have negative entries. Furthermore,

$$|\mu||y| = |Ay| \leq A|y|.$$

Hence, $|\mu| \leq L(|y|) \leq \lambda$, which finishes the proof of item 1.

To prove item 2, one can consider

$$\frac{d}{d\lambda} \det(\lambda I - A)\Big|_{\lambda = \rho(A)},$$

and prove that it is strictly positive. The detail is omitted. Please check the reference.

If A is primitive, then A^k is positive. Clearly the eigenvalues of A^k is the k-th power of the eigenvalues of A. Hence, without loss of generality, we can assume that A is positive and $\rho(A) = 1$ to prove item 4. Let y be an eigenvalue of A with corresponding eigenvalue μ, where $|\mu| = 1$, then

$$z = A|y| - |y| \geq 0.$$

Suppose that $z \neq 0$, then

$$Az > 0,$$

which implies that there exists an $\varepsilon > 0$, such that

$$Az \geq \varepsilon A|y|,$$

which is equivalent to

$$\frac{A}{1+\varepsilon} A|z| \geq A|z|.$$

Thus, for all k,

$$\left(\frac{A}{1+\varepsilon}\right)^k A|z| \geq A|z|,$$

which contradicts with the fact that $\rho(A) = 1$. As a result, $z = 0$. Thus,

$$|y| = A|y|, \text{ and } y = Ay.$$

Hence, y is either all non-negative or all non-positive, which implies that y is just a scalar multiplication of v.

Now to prove item 5 we have

$$L(\mathbf{1}) = \min_i \sum_j a_{ij} \leq \lambda,$$

and

$$A\mathbf{1} \leq \left(\max_i \sum_j a_{ij}\right) \mathbf{1}.$$

Hence,

$$w^T A\mathbf{1} = \lambda w^T \mathbf{1} \leq \left(\max_i \sum_j a_{ij}\right) w^T \mathbf{1},$$

which implies that $\lambda \leq \max_i \sum_j a_{ij}$. □

For a general A matrix, to prove it is stable, we need to consider a Lyapunov function of the following form:

$$V(x) = x^T P x,$$

where P is positive definite and $A^T P A - P$ is negative definite. Since there is no guarantee that P is diagonal (or comply with the network topology), this criterion cannot be easily distributed.

However if A is non-negative and irreducible, then we have

Theorem 2. *If A is non-negative and irreducible, then A is stable if and only if there exists a positive $w \in \mathbb{R}^n$ and $0 < \delta < 1$, such that*

$$w^T A < \delta w^T. \tag{1}$$

The corresponding Lyapunov function is given by

$$V(z) = w^T |z|.$$

Proof. "if": (1) is equivalent to

$$V(Az) < \delta V(z).$$

"only if": If A is stable, then we can choose w as the left eigenvector associated with $\lambda = \rho(A)$. □

We can generalize this result to continuous time and consider Metzler matrix. Assuming that A is a Metzler matrix with $A = B - sI$, where B is irreducible. Hence, A is Hurwitz if and only if $\rho(B) < s$, which is equivalent to the existence of a positive w, such that

$$w^T B < s w^T \Leftrightarrow w^T A < 0.$$

To see this, let v be the right eigenvector associated with $\rho(B)$, then

$$w^T B v = \rho(B) w^T v < s w^T v,$$

which implies that $\rho(B) < s$. Thus, we have the following theorem:

Theorem 3. *If A is Meltzer and irreducible, then A is Hurwitz if and only if there exists a positive $w \in \mathbb{R}^n$, such that*

$$w^T A < 0. \tag{2}$$

The corresponding Lyapunov function is given by

$$V(z) = w^T |z|.$$

Eq (1) and (2) can be verified in a distributed fashion.