Average Consensus

Yilin Mo

July 2, 2015

1 Average Consensus Algorithm

We model a network composed of n agents as a graph $G = \{V, E\}$. $V = \{1, 2, \dots, n\}$ 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 sensor i and j can communicate directly with each other. We will always assume that G is undirected, i.e. $(i, j) \in E$ if and only if $(j, i) \in E$. We further assume that there is no self loop, i.e., $(i, i) \notin E$. The neighborhood of sensor i is defined as

$$\mathcal{N}(i) \triangleq \{j \in V : (i, j) \in E\}. \tag{1}$$

A path $p = (v_0, v_1)(v_1, v_2) \dots (v_{l-1}, v_l)$ is a sequence of edges, such that each $(v_k, v_{k+1}) \in E$.

A graph is called connected if for any pair $i, j \in V$, there always exists a path that connects i and j.

Suppose that each agent has an initial state $x_i(0)$. At each iteration, sensor i will communicate with all its neighbors and update its state according to the following update equation

$$x_i(k+1) = p_{ii} x_i(k) + \sum_{j \in \mathcal{N}(i)} p_{ij} x_j(k). \tag{2}$$

Let us define the vector $x(k) \triangleq [x_1(k), \dots, x_N(k)]' \in \mathbb{R}^n$ and matrix $P \triangleq [p_{ij}] \in \mathbb{R}^{n \times n}$. Now we can rewrite (2) in its matrix form as

$$x_{k+1} = P x_k. \tag{3}$$

Let us define the average vector to be

$$x_{ave} \triangleq \frac{\mathbf{1}' x(0)}{N} \mathbf{1}, \tag{4}$$

where $\mathbf{1} \in \mathbb{R}^n$ is a vector whose elements are all equal to 1. Also let us define the error vector $y(k)$ to be

$$y(k) \triangleq x(k) - x_{ave}. \tag{5}$$

The goal of average consensus is to guarantee that $y(k) \to 0$ as $k \to \infty$ through the update equation (3).

Let us arrange the eigenvalues of P in decreasing order as $\lambda_1(P) \geq \lambda_2(P) \ldots \geq \lambda_n(P)$.

Theorem 1. *The following conditions are necessary and sufficient in order to achieve average consensus from any initial condition $x(0)$:*

1. $\lambda_1(P) = 1$ *and* $|\lambda_i(P)| < 1$ *for all* $i = 2, \ldots, N$.
2. $P\mathbf{1} = \mathbf{1}$, *i.e.* $\mathbf{1}$ *is an eigenvector of* P.
3. $\mathbf{1}^T P = \mathbf{1}^T$, *i.e.* $\mathbf{1}$ *is also a left-eigenvector of* P.

Proof. First, suppose condition 1-3 hold. Hence, P can be written as

$$P = J + Q.$$

where $J = \mathbf{1}\mathbf{1}^T/n$ and Q is stable and

$$JB = BJ = 0.$$

Hence

$$\lim_{k\to\infty} P^k = J + \lim_{k\to\infty} Q^k = J.$$

On the other hand, suppose that $y(k) \to 0$ for any initial condition $x(0)$. As a result

$$\lim_{k\to\infty} P^k = J.$$

However,

$$P^k = \sum_{i=1}^{n} \lambda_i^k w_i v_i^T, \tag{6}$$

where w_i, v_i are the right and left eigenvectors of P.

(we assume P is diagonalizable. The proof can be revise to consider P has a Jordan form.)

Hence, condition 1-3 hold. □

Define the convergence rate ρ as

$$\rho \triangleq \lim_{k\to\infty} \sup_{y(0)\neq 0} \sqrt[k]{\frac{\|y(k)\|_2}{\|y(0)\|_2}}$$

By (6), $\rho = \max(\lambda_2(P), -\lambda_n(P))$.

2 Fast Convergence via Convex Optimization

We want to solve the following problem:

$$\begin{aligned} \underset{P}{\text{minimize}} \quad & \rho \\ \text{subject to} \quad & \mathbf{1}^T P = \mathbf{1}^T \\ & P\mathbf{1} = \mathbf{1} \\ & a_{ij} = 0 \text{ if } (i,j) \notin E \text{ and } i \neq j. \end{aligned}$$

In general, this problem is very difficult for arbitrary P, since ρ is not a convex function of P.

In general, the largest eigenvalue is not a convex function. For example,

$$\rho\left(\begin{bmatrix} 0 & \alpha \\ \beta & 0 \end{bmatrix}\right) = \sqrt{\alpha\beta}.\ \rho\left(\begin{bmatrix} 0 & (\alpha+\beta)/2 \\ (\alpha+\beta)/2 & 0 \end{bmatrix}\right) = (\alpha+\beta)/2.$$

However, if P is assumed to be symmetric, then the problem is a convex optimization problem and can be solved efficiently.

3 Laplacian based Consensus

The degree of sensor i is defined as

$$d_i \triangleq |\mathcal{N}(i)|. \tag{7}$$

A graph is called d-regular graph if all the vertices have the same degree d,i.e. $d_{min} = d_{max} = d$.

Now we can define the Laplacian matrix L of graph G as

$$L \triangleq D - A, \tag{8}$$

where $D = diag(d_1, \dots, d_n)$ is the degree matrix. A is the adjacency matrix, $a_{ij} = 1$ if and only if $(i,j) \in E$.

Theorem 2. *L is positive semidefinite. Furthermore, L has an eigenvalue 0 and the corresponding eigenvector* $\mathbf{1}$. *As a result, arrange the eigenvalues of L in the ascending order:*

$$0 = \lambda_1(L) \leq \lambda_2(L) \leq \cdots \leq \lambda_n(L). \tag{9}$$

Furthermore the graph G is connected if and only if $\lambda_2(L) > 0$ is strictly positive.

Proof. Assume $v = [v_1, \dots, v_n]^T \in \mathbb{R}^n$, then

$$v^T L v = \frac{1}{2} \sum_{(i,j)\in E} (v_i - v_j)^2 \geq 0.$$

Hence, L is positive semidefinite and has a eigenvalue 0 and the corresponding eigenvector $\mathbf{1}$.

If the graph is connected, then $v^T L v = 0$ implies $v = \mathbf{1}$. Hence, $\lambda_2(L) > 0$. If the graph is disconnected, then we can construct a $v \neq \mathbf{1}$, such that $v^T L v = 0$. Hence, $\lambda_2(L) = 0$. □

We now have the following corollary:

Corollary 1. *There exists an P satisfies condition 1-3 if and only if G is connected.*

Proof. If G is not connected, then clearly consensus cannot be achieved.

On the other hand, if G is connected, then we can choose $P = I - \alpha L$, where $\alpha < 2/\lambda_n(L)$. □

Since $\rho = \max(\lambda_2(P), -\lambda_n(P)$, if we consider the P of the form $I - \alpha L$, then the optimal α is given by

$$\alpha^* = \frac{2}{\lambda_2(L) + \lambda_n(L)}$$

and

$$\rho^* = \frac{\lambda_n(L) - \lambda_2(L)}{\lambda_n(L) + \lambda_2(L)}.$$

4 Laplacian from some graph

4.1 Complete Graph

$$L = nJ + nI$$

Hence, L has eigenvalue 0 with multiplicity 1 and eigenvalue n with multiplicity $n - 1$.

4.2 Complete Bipartite graph $K_{a,b}$

$$L = \begin{bmatrix} bI_a & -\mathbf{1}_{a\times b} \\ -\mathbf{1}_{b\times a} & aI_b \end{bmatrix}$$

The eigenvalues are

$$0, a, b, a + b$$

with multiplicities

$$1, b - 1, a - 1, 1.$$

The corresponding eigenvectors are

$$\mathbf{1}_{a+b}, \begin{bmatrix} v \\ \mathbf{0} \end{bmatrix}, \begin{bmatrix} \mathbf{0} \\ w \end{bmatrix}, \begin{bmatrix} b\mathbf{1}_a \\ -a\mathbf{1}_b \end{bmatrix},$$

where $\mathbf{1}_a v^T = 0$ and $\mathbf{1}_b w^T = 0$.

4.3 Cayley Graph

Let $H = (V, *)$ be a group and $S = S^{-1}$ be a symmetric set. We can define a graph $G = (V, E)$, such that

$$(x, y) \in E \Leftrightarrow x^{-1}y \in E.$$

Cayley graph is d-regular with $d = |S|$. For example,

- $H = (\mathbb{Z}, +)$ and $S = \{-1, 1\}$ is an infinite line.
- $H = (\mathbb{Z}_n, +)$ and $S = \{1, n-1\}$ is a cyclic graph.

We consider the Cayley graph generated by $H = (\mathbb{Z}_n, +)$ and S.

Theorem 3. *Define* $\omega = \exp(2j\pi/n)$.

$$\lambda_k(L) = |S| - \sum_{s \in S} \omega^{ks},$$

with eigenvector

$$\begin{bmatrix} 1 & \omega^k & \dots & \omega^{(n-1)k} \end{bmatrix}^T.$$

In general, one can consider

$$f_k(i) = \omega^{ki}.$$

Hence, for any $x, y \in \mathbb{Z}_n$, $f_k(x)f_k(y) = f_k(x+y)$. Such an f_k is called a character of the graph G.

The above theorem can be generalized to

Theorem 4. *For any Cayley graph of group H and symmetric set S. Define vector*

$$\left[\chi(1), \chi(2), \dots, \chi(n)\right]^T,$$

where χ is a character of G. Then v is an eigenvector with eigenvalue:

$$|S| - \sum_{s \in S} \chi(s),$$

The nice thing about Cayley graph is that we can construct an infinitely many d-regular graphs, called Ramanujan Graphs, which satisfy

$$\lambda_2(L) \geq d - 2\sqrt{d-1}, \ \lambda_n(L) \leq d + 2\sqrt{d-1}.$$

Hence, the convergence rate of a consensus algorithm on these graphs is given by

$$\rho^* \leq 2\frac{\sqrt{d-1}}{d},$$

which does not grow with respect to the number of node n.