

Lezione 17 — Maggio 23

Docente: Luca Schenato

Stesori: Comin, Dal Bianco, Fabris, Parmeggiani

17.1 Consensus

17.1.1 Nozioni preliminari

L'idea dell'algoritmo del consenso è quello di trovare un punto di accordo tra tutte le variabili (*agenti*) in gioco in un sistema dinamico descritto in forma compatta dall'equazione di aggiornamento vettoriale

$$\mathbf{x}(t+1) = P\mathbf{x}(t) \quad (17.1)$$

con P matrice stocastica che soddisfa le proprietà:

$$P_{ij} \geq 0 \quad \text{e} \quad \sum_j P_{ij} = 1$$

che equivale a

$$P\mathbf{1} = \mathbf{1}$$

Si vuole cioè far tendere x a un vettore di tutte componenti uguali; formalmente

$$\mathbf{x}(t) \rightarrow \alpha\mathbf{1} = \begin{bmatrix} \alpha \\ \vdots \\ \alpha \end{bmatrix}$$

e $x_i(t) \in \Re$

Si vuole innanzi tutto studiare la velocità di convergenza alla condizione di agreement (*consenso*) al variare della matrice P e quindi delle proprietà del grafo associato.

Si può anticipare qualche risultato che verrà dimostrato nel proseguo: un tipo di grafo *circolante* ha velocità di convergenza bassa e si possono creare delle traiettorie cicliche che rallentano ulteriormente la dinamica desiderata; una situazione intermedia è descritta invece da un grafo *random-geometrico* (i nodi possono comunicare solo entro un certo raggio d'azione, come per esempio i sensori wireless). Il caso migliore è quello di un grafo totalmente casuale, che ha prestazioni migliori se è tempo-variante.

Analizzando la (17.1), per la componente i -esima del vettore $\mathbf{x}$ vale

$$x_i(t+1) = \sum_j P_{ij}x_j(t)$$

Definiamo ora il grafo come:

$$G_p = (V, \epsilon)$$

dove l'insieme dei vertici (in questo caso gli agenti) è $V = \{1, \dots, N\}$ e l'insieme degli archi (*edge*) è $\epsilon \subseteq V \times V$; la coppia di vertici $(i, j) \in \epsilon \Leftrightarrow P_{ij} \neq 0$ ovvero l'elemento i - j -esimo di P è diverso da zero solo se c'è un arco tra il vertice j e il vertice i . La cardinalità di ϵ , $|\epsilon|$, si può intendere come indice di quantità di comunicazione possibile del grafo.

A questo punto si ipotizza di avere $P_{ii} > 0 \forall i$ (ciò non toglie nessuna generalità alla trattazione perchè equivale a supporre che ogni nodo possa sempre comunicare a se stesso la propria informazione); si ipotizza quindi che ogni vertice ha un *self-loop*.

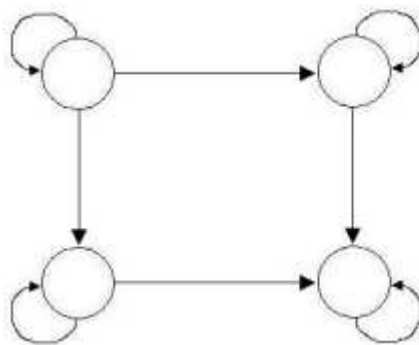


Figura 17.1.

Si definisce la matrice delle adiacenze E i cui elementi sono:

$$E_{ij} = \begin{cases} 1 & \Leftrightarrow (j, i) \in \epsilon \\ 0 & \text{altrimenti} \end{cases}$$

Si può costruire semplicemente E da P inserendovi 1 in posizione (i, j) se l'elemento $P_{ij} > 0$, 0 se $P_{ij} = 0$. Si consideri ora la matrice E^2 i cui elementi sono dati da:

$$(E^2)_{ij} = \sum_l E_{il} E_{lj}$$

Si nota come per avere un contributo su $(E^2)_{ij}$, E_{il} e E_{lj} devono essere entrambi 1. Se questa condizione è verificata significa che esiste il vertice l tale per cui il percorso da j a i , passando per l , sia di lunghezza 2.

$(E^2)_{ij}$ è quindi il numero di cammini da j a i di lunghezza 2. Allo stesso modo $(E)_{ij} = (E^1)_{ij}$ è il numero di *path* di lunghezza 1 e, genericamente, $(E^k)_{ij}$ è il numero di *path* di lunghezza k da j a i .

Si osservi che per l'ipotesi fatta di self-loop ad ogni vertice, se esiste un cammino di lunghezza $k - 1$ tra due vertici, esistono anche cammini di lunghezza k , $\forall k$.

17.1.2 Proprietà spettrali di P

Si vanno ora a studiare autovalori e autovettori della matrice stocastica P ; per fare questo si userà il risultato del Teorema di *Perron-Frobenious*, che fornisce delle caratterizzazioni a matrici non negative in generale (ovvero ogni singolo elemento è non negativo). Le matrici P e E appartengono a questa classe di matrici. Prima di enunciare per intero il teorema se ne anticipano i risultati per dare un'idea circa l'influenza che essi hanno sulla dinamica di convergenza alla condizione di agreement in un problema di consensus.

Il Teorema di Perron-Frobenious dimostra che per le matrici non negative esiste un autovalore dominante λ_0 di modulo maggiore rispetto a tutti gli altri e che il suo miniblocco di Jordan ha dimensione 1. E può essere riscritta come:

$$E = T^{-1} \left[\begin{array}{c|c} \lambda_0 & 0 \\ \hline 0 & \Lambda \end{array} \right] T$$

con Λ sottomatrice di Jordan dimensione $(N-1) \times (N-1)$ corrispondente agli altri autovalori che hanno modulo minore di λ_0 .

Per la proprietà delle matrici di Jordan,

$$E^k = T^{-1} \left[\begin{array}{c|c} \lambda_0^k & 0 \\ \hline 0 & \Lambda^k \end{array} \right] T$$

Le matrici T^{-1} e T si possono riscrivere mettendo in evidenza rispettivamente la prima colonna e la prima riga:

$$T^{-1} = \left[\begin{array}{c|c} v_0 & V \end{array} \right] \quad T = \left[\begin{array}{c} w_0^T \\ \hline W \end{array} \right]$$

da cui

$$TT^{-1} = \left[\begin{array}{c} w_0^T \\ \hline W \end{array} \right] \left[\begin{array}{c|c} v_0 & V \end{array} \right] = \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & I \end{array} \right]$$

Con questa notazione si ha che

$$E^k = \left[\begin{array}{c|c} v_0 & V \end{array} \right] \left[\begin{array}{c|c} \lambda_0^k & 0 \\ \hline 0 & \Lambda^k \end{array} \right] \left[\begin{array}{c} w_0^T \\ \hline W \end{array} \right] = \lambda_0^k v_0 w_0^T + V \Lambda^k W \cong \lambda_0^k v_0 w_0^T$$

per $k \rightarrow \infty$, visto che λ_0 è l'autovalore di modulo maggiore.

Da notare che essendo $(E^k)_{ij}$ l'elemento (i, j) di $\lambda_0^k v_0 w_0^T$ per k sufficientemente grande, λ_0 deve essere necessariamente reale e positivo, non potendo il numero di path di dimensione k essere negativo.

Si consideri ora un grafo completo, in cui cioè

$$E = \mathbf{1}\mathbf{1}^T$$

$$E^k = \overbrace{\mathbf{1}\mathbf{1}^T}^1 \overbrace{\mathbf{1}\mathbf{1}^T}^2 \dots \overbrace{\mathbf{1}\mathbf{1}^T}^k = N^k \frac{\mathbf{1}\mathbf{1}^T}{N}$$

visto che $\mathbf{1}^T \mathbf{1} = N$. In questo caso risulta $\lambda_0 = N$, mentre per un grafo generico $\lambda_0 \leq N$.

Teorema (di Perron-Frobenius): Se una matrice $A > 0$ allora esiste un vettore $\mathbf{v}_0 > 0$ e un autovalore $\lambda_0 > 0$ tali che:

1. $A\mathbf{v}_0 = \lambda_0\mathbf{v}_0$
2. $\Delta_A(s) = (s - \lambda_0)^{\mu_0}(s - \lambda_1)^{\mu_1} \dots (s - \lambda_k)^{\mu_k}$
 - (a) $\mu_0 = 1$
 - (b) $|\lambda_i| < \lambda_0$

Anche intuitivamente si capisce che l'importante è il valore del modulo del secondo autovalore più grande (λ_1 se si usa un ordinamento decrescente).

Da un punto di vista grafico è come se gli autovalori fossero allocati come in Fig.17.2.

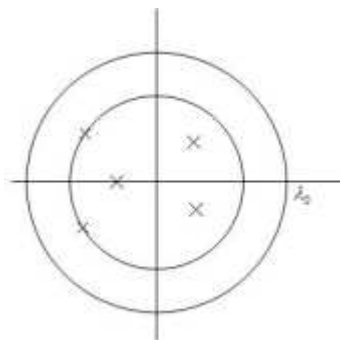


Figura 17.2.

Si vedranno ora un paio di applicazioni del Teorema alle matrici stocastiche.

Caso 1: $P > 0$

Visto che $P\mathbf{1} = \mathbf{1}$ per definizione di matrice stocastica, $\mathbf{1}$ è autovettore di P associato all'autovalore destro 1. Si dimostra ora che λ_0 è proprio 1.

Dimostrazione: Per il teorema di Perron-Frobenius $\exists \lambda_0, \mathbf{v}_0$ perchè $P > 0$, quindi

$$P\mathbf{v}_0 = \lambda_0\mathbf{v}_0 \quad (17.2)$$

La componente i -esima di (17.2) risulta

$$\sum_j P_{ij}v_j = \lambda_0 v_i \quad (17.3)$$

Si definiscono ora $M = \max(v_j)$ e $m = \min(v_j)$. Sostituendo tali valori nell'equazione (17.3) si ottengono un limite superiore e inferiore per $\lambda_0 v_i$:

$$m \leq \lambda_0 v_i \leq M$$

($\sum_j P_{ij} m = m \sum_j P_{ij} = m$ perchè P matrice stocastica; lo stesso vale naturalmente per M). Visto che $\exists v_h = m$, $v_k = M$ e che $m \leq \lambda_0 v_i \leq M \forall i$, si valuta tale disuguaglianza per $i = h$ e $i = k$. Nel primo caso

$$m \leq \lambda_0 v_h = \lambda_0 m$$

da cui visto che $m > 0$, $\lambda_0 \geq 1$.

Nel secondo caso

$$\lambda_0 M = \lambda_0 v_k \leq M$$

da cui visto che $M > 0$, si ottiene $\lambda_0 \leq 1$. Combinando i due risultati si trova $\lambda_0 = 1$.

Si è così dimostrato che per le matrici stocastiche il più grande autovalore reale la cui esistenza è garantita dal teorema di Perron-Frobenius.

Si analizza ora il problema iniziale di consensus (17.1) per $P > 0$ e sfruttando quanto appena dimostrato. Lo stato ha evoluzione dinamica data da:

$$\begin{aligned} \mathbf{x}(t) &= P^t \mathbf{x}_0 \\ P &= \left[\begin{array}{c|c} v_0 & V \end{array} \right] \left[\begin{array}{c|c} 1 & 0 \\ \hline 0 & \Lambda \end{array} \right] \left[\begin{array}{c} w_0^T \\ \hline W \end{array} \right] \\ P^t &= \lambda_0^t \mathbf{v}_0 \mathbf{w}_0^T + V \Lambda^t W \end{aligned}$$

Il secondo addendo della terza equazione tende a 0 perchè ha tutti autovalori con modulo minore di 1, quindi

$$\begin{aligned} P^t &\xrightarrow{t \rightarrow \infty} \mathbf{v}_0 \mathbf{w}_0^T = \mathbf{1} \mathbf{w}_0^T \\ \mathbf{x}(t) &= P^t \mathbf{x}_0 \xrightarrow{t \rightarrow \infty} \mathbf{1} \mathbf{w}_0^T x(0) = \begin{bmatrix} \alpha \\ \vdots \\ \alpha \end{bmatrix} \end{aligned}$$

Si arriva cioè alla condizione di agreement in cui $\alpha = \sum_i w_i x_i(0)$ ovvero il valore finale di consenso è combinazione lineare delle condizioni iniziali.

Si può inoltre dimostrare che $\mathbf{w}_0^T P = \mathbf{w}_0^T$, ovvero che $\mathbf{w}_0$ è autovettore sinistro di P . Il vettore $\mathbf{w}_0$ dà quindi il peso alle condizioni iniziali sul valore di agreement.

Si può concludere che per matrici stocastiche $P > 0$ si arriva sempre a un agreement dato dalla combinazione lineare delle condizioni iniziali e i pesi di queste sul valore finale del consenso sono dettati dagli elementi dell'autovettore sinistro di P , $\mathbf{w}_0$.

Caso 2: $\exists k | P^k > 0$

In queste condizioni non è detto che $P > 0$; succede però che visto che P è stocastica, anche P^k lo è ($P\mathbf{1} = \mathbf{1} \Rightarrow P^k\mathbf{1} = \mathbf{1}$).

Se gli autovalori di P sono

$$\lambda(P) = \{\lambda_0, \lambda_1, \dots, \lambda_h\}$$

gli autovalori di P^k risultano

$$\lambda(P^k) = \{\lambda_0^k, \lambda_1^k, \dots, \lambda_h^k\}$$

Tra questi, per il Teorema di Perron-Frobenius, ne esiste sicuramente uno dominante, ad esempio sia esso λ_0^k ; visto che P^k è stocastica e $P^k > 0$, per quanto appena visto, $\lambda_0^k = 1$, quindi $\lambda_0 = 1$.

Gli altri autovalori λ_i^k saranno in modulo minori di 1, quindi $|\lambda_i| < 1$.

Si può concludere che, data una matrice stocastica P e $\exists k | P^k > 0$, si ritorna all'analisi effettuata al caso 1 e valgono tutte le osservazioni già fatte.

Il problema ora è verificare se esiste k che soddisfa la condizione $P^k > 0$. La risposta si trova analizzando il grafo associato a P : condizione necessaria e sufficiente è che il grafo sia fortemente connesso.

Definizione: Il grafo G_P associato alla matrice stocastica P è detto fortemente connesso se e solo se fissati due vertici i e j , esiste sempre un percorso che va da j a i .

Lemma Se P è una matrice stocastica, $(P^k)_{ij} > 0 \Leftrightarrow \exists$ un percorso in G_P di lunghezza k che va da j a i .

Dimostrazione La dimostrazione si fa per induzione:

- per $k = 1$ è ovvio: $P_{ij} > 0 \Leftrightarrow \exists$ un arco da j a i .
- (si dimostra solo la sufficienza del Lemma ($\Leftarrow$)).
Si supponga che il risultato sia vero per $k - 1$, esiste cioè un percorso da j a i di lunghezza $k - 1$.

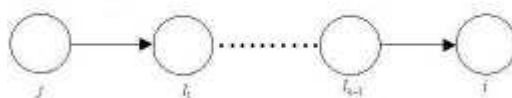


Figura 17.3.

Allora

$$(P^k)_{ij} = (P^{k-1}P)_{ij} = \sum_l (P^{k-1})_{il} P_{lj} \geq (P^{k-1})_{i,l_1} P_{l_1,j} > 0$$

L'ultima disuguaglianza vale in quanto $(P^{k-1})_{i,l_1} > 0$ per l'ipotesi di induzione, $P_{l_1,j} > 0$ per definizione di grafo: esiste un percorso da j a i , quindi anche da j a l_1 : allora $P_{l_1,j} > 0$.

La necessarietà del Lemma ($\Rightarrow$) si dimostra in modo del tutto analogo.

Applicando il Lemma appena visto per tutti gli i e j si può concludere che:

Teorema G_P è fortemente connesso $\Leftrightarrow \exists k | P^k > 0$.

Vengono ora proposti alcuni esempi.

1. $P = I \Rightarrow P^k = I \forall k$. In queste condizioni non avviene mai che $P^k > 0$ per qualche k , quindi non si arriva mai a una condizione di agreement: anche intuitivamente si capisce che questo non possa avvenire in quanto nessun agente comunica con gli altri.

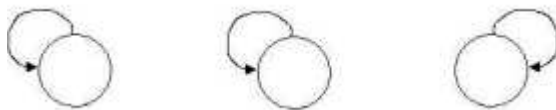


Figura 17.4.

- 2.

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Graficamente si ha:

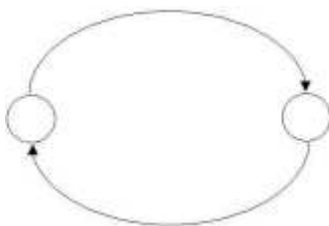


Figura 17.5.

$$P^k = \begin{cases} I & k \text{ pari} \\ P & k \text{ dispari} \end{cases}$$

Il grafo di Fig. 17.5 risulta essere fortemente connesso, ma non è presente il self loop necessario per applicare i risultati ottenuti.

3.

$$P = \begin{bmatrix} 1/2 & 1/2 \\ 0 & 1 \end{bmatrix}$$

Graficamente:

**Figura 17.6.**

Il grafo non è fortemente connesso e

$$P^k = \begin{bmatrix} 2^{-k} & 1 - 2^{-k} \\ 0 & 1 \end{bmatrix} \xrightarrow{k \rightarrow \infty} \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$$

da cui

$$\begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix} \mathbf{x}(0) = \begin{bmatrix} x_2(0) \\ x_2(0) \end{bmatrix}$$

In questo caso avviene l'agreement, pesando però solo una delle due opinioni iniziali.