PSC: Progettazione di sistemi di controllo **a.a. 2010-2011**

Lezione 21 — 31 Maggio

Docente: Luca Schenato *Stesori: Bertinato Marco, Ortolan Giulia, Zambotto Patrizio*

21.1 Analisi in media delle prestazioni di un algoritmo di *consensus*

L'analisi delle prestazioni di un algoritmo di *consensus* con matrice P tempovariante può essere fatta 'a regime', ovvero studiando il valore a cui (eventualmente) converge lo stato, oppure 'nel transitorio', ovvero guardando la velocità con cui viene raggiunto il valore di regime. A questo proposito sono stati introdotti i seguenti parametri:

- per l'analisi nel transitorio la *distanza dal* consensus

  $$d(t) = \frac{1}{N}\left|\left|x(t) - x_A(t)\mathbb{1}\right|\right|^2$$

  dove $x_A(t) = \frac{1}{N}\mathbb{1}^T x(t)$ è la media delle componenti di $x(t)$, ovvero il baricentro all'istante t;

- per l'analisi a regime lo *spostamento dal baricentro corrente* rispetto al baricentro iniziale

  $$\beta(t) = \left|x_A(t) - x_A(0)\right|^2.$$

In particolare interessa analizzare il comportamento di questi parametri per $t \to \infty$; per l'analisi del transitorio si utilizza la teoria degli esponenti di Lyapunov che consente di trovare il *rate* di convergenza per $d(t)$ a regime, ottenendo che sotto le condizioni per cui si ha *probabilistic consensus* vale

$$\lim_{t\to\infty} [d(t)]^{1/t} = \lambda \text{ quasi sicuramente,}$$

con λ costante e non aleatorio detto *secondo esponente di Lyapunov*; per l'analisi a regime si trova

$$\beta(\infty) = x(0)^T B(\omega) x(0) \text{ quasi sicuramente,}$$

con $B(\omega) = \left(\rho(\omega) - \frac{1}{N}\mathbb{1}\right)\left(\rho(\omega) - \frac{1}{N}\mathbb{1}\right)^T$ matrice aleatoria.

I problemi nascono dal fatto che per l'analisi del transitorio il secondo esponente di Lyapunov fornisce un'informazione troppo scarsa per una stima della qualità delle prestazioni (si veda l'esempio descritto nella lezione scorsa), e per l'analisi a regime è necessario studiare una matrice aleatoria. Ci si deve perciò ricondurre ad un'*analisi in media*, studiando $\mathbb{E}[\beta(\infty)]$ ed $\mathbb{E}[d(\infty)]$.

21.1.1 Analisi in media della distanza dal *consensus*

Ci si occupa ora dello studio di $\mathbb{E}[d(t)]$; è noto dalla lezione precedente che

$$\mathbb{E}[d(t)] = \mathbb{E}\left[x(t)^T \Omega x(t)\right] = x(0)^T \Delta(t) x(0) \tag{21.1}$$

dove

$$\Delta(t) \triangleq \mathbb{E}\left[P(0)^T\ P(1)^T\ \ldots\ P(t-1)^T\ \Omega\ P(t-1)\ \ldots\ P(1)\ P(0)\right]$$

per $t \geq 1$ e

$$\Delta(0) \triangleq \Omega = I - \frac{1}{N}\mathbb{1}\mathbb{1}^T.$$

É possibile provare che

$$\Delta(t+1) = \mathbb{E}\left[P(0)^T \Delta(t) P(0)\right] \tag{21.2}$$

e quindi $\Delta(t)$ segue l'evoluzione di un sistema dinamico lineare che può essere scritto nella forma

$$\Delta(t+1) = \mathcal{L}\left(\Delta(t)\right) \tag{21.3}$$

dove $\mathcal{L} : \mathbb{R}^{N\times N} \longrightarrow \mathbb{R}^{N\times N}$ è l'operatore lineare di dimensione N^2 definito in (21.2).

Risulta utile introdurre una rappresentazione matriciale alternativa per l'operatore lineare $\mathcal{L}$. Data una matrice $A \in \mathbb{R}^{N\times N}$ si definisce vect(A) il vettore colonna che ha l'elemento A_{ij} nella posizione $\big((i-1)N + j\big)$-esima. Si può dimostrare che

$$\text{vect}(ABC) = (C^T \otimes A)\text{vect}(B),$$

dove $\otimes$ indica il prodotto di Kronecker per matrici. Utilizzando questa notazione e le proprietà del prodotto di Kronecker, si trova che

$$\begin{aligned}\text{vect}\big(\Delta(t+1)\big) &= \mathbb{E}\left[P(0)^T \otimes P(0)^T\right]\text{vect}\big(\Delta(t)\big)\\ &= L \cdot \text{vect}\big(\Delta(t)\big)\end{aligned} \tag{21.4}$$

e quindi l'operatore lineare $\mathcal{L}(\cdot)$ si può rappresentare con la matrice

$$L \triangleq \mathbb{E}\left[P(0)^T \otimes P(0)^T\right] = \mathbb{E}\left[P(0) \otimes P(0)\right]^T,\ L \in \mathbb{R}^{N^2\times N^2} \tag{21.5}$$

É possibile provare che L ha tutti elementi non negativi e che L^T è una matrice stocastica per righe, ovvero ha un autovalore in 1 e tutti gli autovalori di modulo strettamente inferiore ad 1. Le condizioni per cui si ha *probabilistic consensus* garantiscono che (21.4) giunga al *consensus*; la velocità di convergenza del sistema (21.4) determina la velocità di (21.1).

21.1.2 Analisi in media dello spostamento dal baricentro corrente

Per cominciare si osserva che, poichè $\beta(t)$ converge con probabilità unitaria a $\beta(\infty) = \left[\left(\rho^T(\omega) - \frac{1}{N}\mathbb{1}^T\right)x(0)\right]^2$, allora

$$\lim_{t\to\infty} \mathbb{E}\left[\beta(t)\right] = \mathbb{E}\left[\beta(\infty)\right].$$

É stato mostrato la lezione scorsa che

$$\mathbb{E}\left[\beta(\infty)\right] = x(0)^T \overline{B} x(0),$$

dove (omettendo la dipendenza di ρ da ω)

$$\overline{B} = \mathbb{E}\left[\left(\rho - \frac{1}{N}\mathbb{1}\right)\left(\rho - \frac{1}{N}\mathbb{1}\right)^T\right] = \mathbb{E}\left[\rho\rho^T\right] - \frac{1}{N}\mathbb{E}\left[\rho\right]\mathbb{1}^T - \frac{1}{N}\mathbb{1}\mathbb{E}\left[\rho\right]^T - \frac{1}{N^2}\mathbb{1}\mathbb{1}^T. \tag{21.6}$$

In particolare, se $\overline{P}$ è doppiamente stocastica, $\mathbb{E}\left[\rho\right] = \frac{1}{N}\mathbb{1}$ e (21.6) si semplifica in

$$\overline{B} = \mathbb{E}\left[\rho\rho^T\right] - \frac{1}{N^2}\mathbb{1}\mathbb{1}^T.$$

Utile notare che $\overline{B}$ è espressa in funzione di $\mathbb{E}\left[\rho\right]$ e di $\mathbb{E}\left[\rho\rho^T\right]$, che sono gli autovettori relativi all'autovalore 1 della matrice $\overline{P} = \mathbb{E}\left[P(t)\right]$ e dell'operatore $\mathcal{L}$ rispettivamente; a questo punto, rimane da calcolare soltanto il termine $\mathbb{E}\left[\rho\rho^T\right]$. Data una generica matrice $X \in \mathbb{R}^{N\times N}$ presa come condizione iniziale, applicando iterativamente la (21.3) si trova

$$\mathbb{E}\left[x(t)^T X x(t)\right] = x(0)^T \mathcal{L}^t(X)x(0);$$

in particolare, ciò è vero con $X = \Omega$, per cui si ottiene

$$\Delta(t) = \mathcal{L}^t(\Delta(0)) = \mathcal{L}^t(\Omega) \tag{21.7}$$

e

$$\mathbb{E}\left[d(t)\right] = x(0)^T\Delta(t)x(0) = x(0)^T\mathcal{L}^t(\Omega)x(0) \tag{21.8}$$

Si osservi poi che, nelle condizioni che garantiscono il *probabilistic consensus*, si ha

$$\mathbb{E}\left[x(t)^T X x(t)\right] \xrightarrow{t\to\infty} \mathbb{E}\left[x(\infty)^T X x(\infty)\right] = \mathbb{E}\left[x(0)^T\rho\ \mathbb{1}^T X\mathbb{1}\ \rho^T x(0)\right] = \mathbb{1}^T X\mathbb{1}x(0)^T\mathbb{E}\left[\rho\rho^T\right]x(0),$$

dove $x(\infty) = \rho^T x(0)\mathbb{1}$. Ma d'altra parte deve essere

$$x(0)^T\mathcal{L}^t(X)x(0) \xrightarrow{t\to\infty} x(0)^T\mathcal{L}^\infty(X)x(0);$$

pertanto si deve avere

$$\mathbb{E}\left[\rho\rho^T\right] = \frac{1}{\mathbb{1}^T X\mathbb{1}}\ \mathcal{L}^\infty(X).$$

Questa formula è valida $\forall X\ :\ \mathbb{1}^T X\mathbb{1} \neq 0$, cioè per tutte le matrici la cui somma degli elementi è non nulla; la scelta più semplice è ovviamente quella di prendere $X = I$, mentre non si può prendere $X = \Omega$. Tuttavia è interessante notare che, quando si ha *probabilistic consensus*, $\mathbb{E}\left[d(\infty)\right] = x(0)^T\mathcal{L}^\infty(\Omega)x(0) = 0$, e ciò significa che Ω non eccita l'autovettore di $\mathcal{L}$ associato all'autovalore 1; infatti, supponendo che gli autovalori dell'operatore siano semplici e distinti, si ha

$$\mathcal{L}^t(X) = \alpha_0 1^t v_0 + \alpha_1\lambda_1^t v_1 + \ldots \xrightarrow{t\to\infty} \alpha_0 v_0$$

con

- $1, \lambda_1, \lambda_2, \ldots, \lambda_{N^2-1}$ autovalori dell'operatore, $\left|\lambda_i\right| < 1,\ i = 2, \ldots, N^2-1$;
- $v_0, v_1, v_2, \ldots, v_{N^2-1}$ autovettori associati agli autovalori;

- $\alpha_0, \ldots, \alpha_{N^2-1}$ scalari che dipendono da X;

e la convergenza a 0 richiede $\alpha_0 = 0$.

Esempio - Asymmetric gossip. Sia dato un grafo completo con N^2 archi (*edge*), e la matrice di comunicazione $P(t)$ implementi l'algoritmo di *asymmetric gossip*. Ad ogni istante t si hanno N^2 possibili matrici di comunicazione, tutte equiprobabili di probabilità $p = 1/N^2$. Il prototipo della matrice di comunicazione in cui si accende l'arco che unisce i vertici j e i è

$$R^{ij} = \begin{bmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & 1-q & q & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} = I - q \begin{bmatrix} 0 & & & & & \\ & \ddots & & & & \\ & & 0 & & & \\ & & & 1 & -1 & \\ & & & & 0 & \\ & & & & & \ddots \\ & & & & & & 0 \end{bmatrix} = I - q \underbrace{\begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}}_{\triangleq e_i} (e_i - e_j)^T .$$

Si vuole studiare innanzitutto la dinamica della distanza dal *consensus*. In generale l'operatore $\mathcal{L}(\cdot)$ è difficile da calcolare in maniera esplicita come

$$\mathcal{L}(\Delta) = \mathbb{E}\left[P(t)^* \Delta P(t)\right] = \sum_{ij} p \left(R^{ij}\right)^T \Delta \left(R^{ij}\right), \tag{21.9}$$

ma risulta fattibile effettuare il calcolo dell'operatore in corrispondenza delle due matrici I e $\frac{\mathbb{1}\mathbb{1}^T}{N}$:

$$\mathcal{L}(I) = I - \frac{2q(1-q)}{N}\,\Omega, \tag{21.10}$$

$$\mathcal{L}\left(\frac{\mathbb{1}\mathbb{1}^T}{N}\right) = \frac{\mathbb{1}\mathbb{1}^T}{N} + \frac{2q^2}{N^2}\,\Omega. \tag{21.11}$$

Sfruttando la *linearità* dell'operatore si può allora calcolare il valore di $\mathcal{L}(\Omega)$

$$\begin{aligned} \mathcal{L}(\Omega) = \mathcal{L}\left(I - \frac{\mathbb{1}\mathbb{1}^T}{N}\right) &= \mathcal{L}(I) - \mathcal{L}\left(\frac{\mathbb{1}\mathbb{1}^T}{N}\right) = I - \frac{2q(1-q)}{N}\Omega - \frac{\mathbb{1}\mathbb{1}^T}{N} - \frac{2q^2}{N^2}\,\Omega = \\ &= \Omega - \left[\frac{2q(1-q)}{N} + \frac{2q^2}{N^2}\right]\Omega = \left(1 - \frac{2q(1-q)}{N} - \frac{2q^2}{N^2}\right)\Omega, \end{aligned} \tag{21.12}$$

e quindi Ω è autovettore di $\mathcal{L}(\cdot)$.
La velocità di convergenza di $\mathcal{L}^\infty(\Omega)$ e quindi la velocità di convergenza di $\mathbb{E}\left[d(t)\right]$, in questo semplice caso, sono determinate dall'autovalore di Ω:

$$1 - \frac{2q(1-q)}{N} - \frac{2q^2}{N^2} \triangleq R < 1. \tag{21.13}$$

e quindi gli agenti tenderanno ad assumere tutti uno stesso valore con probabilità 1.

Dalle equazioni precedenti, si vede che lo spazio $V = \langle \Omega, \Pi \rangle = \alpha\Omega + \beta\Pi$, dove $\Pi = \frac{\mathbb{1}\mathbb{1}^T}{N}$, si hanno le due seguenti proprietà:

1) $\Omega \in V$, cioè la condizione iniziale appartiene a V (con $\alpha = 1$ e $\beta = 0$);

2) $\mathcal{L}(V) \subseteq V$, quindi V è invariante rispetto all'operatore $\mathcal{L}(\cdot)$;

Per quanto riguarda lo studio di $\mathbb{E}[\beta(t)]$, si vede che in questo caso $\overline{P} = \mathbb{E}[P] = \frac{1}{N^2}\sum R^{ij}$ è doppiamente stocastica, perciò

$$\overline{B} = \mathbb{E}\left[\rho\rho^T\right] - \frac{1}{N^2}\mathbb{1}\mathbb{1}^T = \frac{1}{\mathbb{1}^T X \mathbb{1}}\mathcal{L}^\infty(X) - \frac{1}{N}\Pi. \tag{21.14}$$

Per risolvere (21.14), basta prendere una qualsiasi X tale che $\mathbb{1}^T X \mathbb{1} \neq 0$: bisogna quindi calcolare $\mathcal{L}^\infty(X)$ e si nota subito che $I \in V$. In generale, in virtù del fatto che V è invariante per l'operatore $\mathcal{L}(\cdot)$, si ha che, partendo dalla condizione iniziale $\Delta(0) = \alpha(0)I + \beta(0)\Pi$, la dinamica è data da

$$\Delta(t) = \mathcal{L}^t(\Delta(0)) = \alpha(t)\Omega + \beta(t)\Pi, \tag{21.15}$$

ed è pertanto sufficiente capire come variano i coefficienti $\alpha(t)$ e $\beta(t)$ per conoscere completamente la dinamica (21.15). A questo proposito si sfruttano (21.11) e (21.12) ottenendo

$$\begin{aligned}
\mathcal{L}(\alpha(t)\Omega + \beta(t)\Pi) &= \alpha(t)\mathcal{L}(\Omega) + \beta(t)\mathcal{L}(\Pi) = R\alpha(t)\Omega + \beta(t)\Pi + \frac{2q^2}{N^2}\beta(t)\Omega \\
&= \left(R\alpha(t) + \frac{2q^2}{N^2}\beta(t)\right)\Omega + \beta(t)\Pi \\
&= \alpha(t+1)\Omega + \beta(t+1)\Pi
\end{aligned}$$

che consentono di definire completamente la mappa (lineare) $\big(\alpha(t), \beta(t)\big) \mapsto \big(\alpha(t+1), \beta(t+1)\big)$:

$$\begin{bmatrix} \alpha(t+1) \\ \beta(t+1) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 - \frac{2q(1-q)}{N} - \frac{2q^2}{N^2} & \frac{2q^2}{N^2} \\ 0 & 1 \end{bmatrix}}_{\triangleq C} \begin{bmatrix} \alpha(t) \\ \beta(t) \end{bmatrix}.$$

La matrice C ha due autovettori distinti:

1. $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$, che individua Ω, con autovalore R dato da (21.13);

2. $\begin{bmatrix} 1 \\ 1 + \frac{1-q}{q}N \end{bmatrix}$, con autovalore 1.

Gli autovettori di $\mathcal{L}$ si deducono dagli autovettori di C, e sono:

1. $1 \cdot \Omega + 0 \cdot \Pi = \Omega$;

2. $1 \cdot \Omega + \left(1 + \frac{1-q}{q}N\right) \cdot \Pi \triangleq \Lambda$

e si osserva quindi che (Ω, Λ) è base di autovettori di V. Se ora scegliamo $X = \Lambda$, si ha che

$$\mathcal{L}^{\infty}(\Lambda) = \Lambda, \quad \mathbb{1}^T \Lambda \mathbb{1} = N\left(1 + \frac{1-q}{q}N\right)$$

sostituendo in (21.14) si ricava:

$$\overline{B} = \frac{\mathcal{L}^{\infty}(\Lambda)}{\mathbb{1}^T \Lambda \mathbb{1}} - \frac{1}{N}\Pi = \frac{1}{N\left(1 + \frac{1-q}{q}N\right)}\Omega,$$

che individua un certo valore di *consensus*. Si osserva che per $N \to \infty$ la matrice $\overline{B}$ tende a 0 e quindi con probabilità 1 si raggiunge il baricentro iniziale.

Rifacendo gli stessi calcoli nel caso di algoritmo di *symmetric gossip* si troverebbe $R = 1 - \frac{1}{N} \simeq 0.94$; pertanto, almeno per valori piccoli di t, l'andamento della distanza dal *consensus* non si discosta molto da quello individuato da R^t. Questo giustifica il grafico mostrato nella scorsa lezione e mostra che il secondo esponente di Lyapunov non offre informazioni esaustive sull'analisi delle prestazioni.