PSC: Progettazione di sistemi di controllo	III Trim. 2007
Lezione 23 — 4 Giugno	
Docente: Luca Schenato	*Stesori: Bertinato Marco, Ortolan Giulia, Zambotto Patrizio*

23.1 Problema del consensus: analisi *worst-case*

Si riprende il problema, proposto nella lezione precedente, di trovare delle condizioni necessarie e sufficienti sulla sequenza di matrici $P(t)$ che consentano di raggiungere comunque il *consensus* (*analisi del caso peggiore*). Si ricorda il modello che descrive la dinamica del sistema:

$$x(t+1) = P(t)x(t), \tag{23.1}$$

con $P(t)$ matrice stocastica tempovariante; l'*agreement* è raggiunto quando $x(t) \xrightarrow{t \to \infty} \alpha\mathbb{1}$, i.e. tutte le componenti dello stato convergono allo stesso valore.

Risulta utile definire alcuni operatori per matrici $P \in \mathbb{R}^{m \times n}$

- l'operatore $\lfloor \cdot \rfloor$, vettore riga che contiene l'elemento minore di ciascuna colonna:

$$\lfloor P \rfloor \in \mathbb{R}^{1 \times n},$$
$$(\lfloor P \rfloor)_i \triangleq \min_j P_{ji}, \ \ i = 1 \dots n;$$

- l'operatore $\lceil \cdot \rceil$, vettore riga che contiene l'elemento maggiore di ciascuna colonna

$$\lceil P \rceil \in \mathbb{R}^{1 \times n},$$
$$(\lceil P \rceil)_i \triangleq \max_j P_{ji}, \ \ i = 1 \dots n;$$

- l'operatore $\lfloor\!\lfloor \cdot \rfloor\!\rfloor$, definito come segue:

$$\lfloor\!\lfloor P \rfloor\!\rfloor \triangleq P - \mathbb{1}\lfloor P \rfloor;$$

- l'operatore $\lceil\!\lceil \cdot \rceil\!\rceil$, definito come segue:

$$\lceil\!\lceil P \rceil\!\rceil \triangleq \mathbb{1}\lceil P \rceil - P.$$

Si noti che se P ha tutti gli elementi non negativi, $\lfloor\!\lfloor P \rfloor\!\rfloor$ e $\lceil\!\lceil P \rceil\!\rceil$ sono ancora ad elementi non negativi. Inoltre si ricorda che, proprio come se $x(t)$ è soluzione di (23.1) allora $\lfloor x(t) \rfloor$ (operatore su vettori) individua una successione monotona crescente e limitata, così nel caso matriciale si ha analogamente

$$\dots \geq \lfloor P(3)P(2)P(1) \rfloor \geq \lfloor P(2)P(1) \rfloor \geq \lfloor P(1) \rfloor$$

e

$$\lceil P(1) \rceil \geq \lceil P(2)P(1) \rceil \geq \lceil P(3)P(2)P(1) \rceil \geq \dots,$$

quindi $\lfloor P(k)P(k-1)\dots P(1) \rfloor$ è successione monotona crescente e $\lceil P(k)P(k-1)\dots P(1) \rceil$ è successione monotona decrescente[1]. Inoltre, poichè $\lfloor P(k)P(k-1)\dots P(1) \rfloor \leq \lceil P(k)P(k-1)\dots P(1) \rceil$, entrambe le successioni sono monotone e limitate, pertanto ammettono limite finito

$$\begin{aligned} &\lim_{k\to+\infty} \lfloor P(k)P(k-1)\dots P(1) \rfloor \quad (\geq \lceil P(1) \rceil) \\ &\lim_{k\to+\infty} \lceil P(k)P(k-1)\dots P(1) \rceil \quad (\leq \lfloor P(1) \rfloor). \end{aligned}$$

Si presenta di seguito una condizione sufficiente per la convergenza al *consensus* della dinamica (23.1).

Teorema 23.1.

$$\begin{aligned} \left\| \llbracket P(t) \rrbracket \right\|_{i\infty} \leq \lambda < 1 \ \forall t \implies & \left\| P(k)P(k-1)\dots P(1) - \mathbb{1}\lfloor P(k)P(k-1)\dots P(1) \rfloor \right\|_{i\infty} \leq \bar{b}\lambda^k \\ \implies & \lim_{k\to+\infty} \left\| P(k)P(k-1)\dots P(1) - \mathbb{1}\lfloor P(k)P(k-1)\dots P(1) \rfloor \right\|_{i\infty} = 0 \end{aligned}$$

La condizione esposta garantisce convergenza esponenziale al consensus, *con velocità non inferiore a quella di λ^k.*

La condizione $\left\| \llbracket P(t) \rrbracket \right\|_{i\infty} \leq 1$ equivale ad avere $\forall t$ almeno una colonna di $P(t)$ con tutti gli elementi strettamente positivi; infatti vale

$$\left\| \llbracket P(t) \rrbracket \right\|_{i\infty} \iff \left\| P - \mathbb{1}\lfloor P(t) \rfloor \right\|_{i\infty} \leq 1 \iff \lfloor P(t) \rfloor \neq 0.$$

23.1.1 Effetti sul grafo di comunicazione

Riprendendo l'esempio sulla convergenza della posizione dei veicoli si supponga di avere

$$\begin{bmatrix} x_1(t+1) \\ x_2(t+1) \\ \vdots \\ x_N(t+1) \end{bmatrix} = \begin{bmatrix} * & p_{1j}(t) & * \\ * & p_{2j}(t) & * \\ * & \vdots & * \\ * & p_{Nj}(t) & * \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \\ \vdots \\ x_N(t) \end{bmatrix}, \quad p_{ij}(t) > 0 \ \forall t, \ \forall i = 1, \dots, N.$$

Il fatto che la colonna $j-$esima della matrice di aggiornamento abbia sempre elementi strettamente positivi implica che tutti gli agenti si avvicinano a x_j al passo successivo (i.e. si riduce l'area del poligono convesso che contiene gli agenti). Quando ciò accade, si dice che il grafo $\mathcal{G}_P$ è *strongly rooted*, ovvero un certo nodo j è in grado di comunicare con tutti gli altri con un percorso di lunghezza unitaria[2]. Per quanto meno restrittiva rispetto alla condizione di grafo completo, questa condizione rimane molto forte: pare quindi che il teorema 23.1 sia di scarsa utilità pratica. Si mostra

[1] Per semplicità di notazione e per attenersi a quanto mostrato a lezione, si suppone $t_0 = 1$.

[2] Quando in un grafo si ha un nodo che comunica con tutti gli altri si parla di configurazione a stella o di tipo *master-slave*.

che in realtà è possibile allentare la condizione esposta ottenendo un'altra condizione sufficiente a garantire convergenza al *consensus*.

Si considerino i prodotti di M matrici stocastiche

$$\begin{aligned}\widetilde{P}(1) &\triangleq P(M)\ldots P(1),\\ \vdots \quad &\quad \vdots\\ \widetilde{P}(k) &\triangleq P\big((k-1)M+M\big)\ldots P\big((k-1)M+1\big),\end{aligned}$$

che sono ancora matrici stocastiche, e la sequenza $\widetilde{P}(k)\widetilde{P}(k-1)\ldots\tilde{P}(1)$. Allora è sufficiente che il teorema 23.1 valga sostituendo $P(t)$ con $\widetilde{P}(t)$ per avere convergenza al *consensus*.

Per vedere l'utilità di questa considerazione si definisce la *composizione* tra grafi:

$$\mathcal{G}_{P(t+1)} \circ \mathcal{G}_{P(t)} \triangleq \mathcal{G}_{P(t+1)P(t)},$$

operatore binario che non gode di proprietà commutativa. Si vede facilmente che, se ogni nodo ha il proprio *self-loop* (i.e. gli elementi della diagonale di $P(t)$ sono non nulli $\forall t$), come è sempre nei casi in analisi, sicuramente tutti gli archi presenti al passo t vanno ad aggiungersi agli archi del passo $t+1$, ovvero

$$\big(P(t)\big)_{ij} \neq 0 \Rightarrow \big(P(t+1)P(t)\big)_{ij} \neq 0 \qquad \text{e quindi} \qquad \mathcal{G}_{P(t+1)} \circ \mathcal{G}_{P(t)} \supseteq \mathcal{G}_{P(t+1)} \cup \mathcal{G}_{P(t)}.$$

Si osservi che, in generale, $\mathcal{G}_{P(t+1)} \circ \mathcal{G}_{P(t)} \neq \mathcal{G}_{P(t+1)} \cup \mathcal{G}_{P(t)}$ dato che il prodotto tra matrici può far comparire anche altri archi; inoltre, se la proprietà di *strongly rooted* vale per l'unione dei grafi, allora essa vale anche per la composizione.

Si ricorda che, se $\mathcal{G}_{P(t)}$ è *strongly rooted* $\forall\, t$, allora c'è convergenza esponenziale al *consensus*. Il teorema seguente mostra come sia possibile ottenere grafi *strongly rooted* tramite la composizione di grafi non *strongly rooted*.

Teorema 23.2. *Se* $\forall t$ $\mathcal{G}_{P(t)}$ *è* rooted *(i.e. esiste un nodo che riesce a raggiungere tutti gli altri con cammini di lunghezza qualsiasi) allora*

$$\overline{\mathcal{G}} \triangleq \underbrace{\mathcal{G}_{P((n-1)^2)} \circ \mathcal{G}_{P((n-1)^2-1)} \circ \ldots \circ \mathcal{G}_{P(1)}}_{\textit{composizione di } (n-1)^2 \textit{ grafi}}$$

è strongly rooted.

Per applicare il teorema 23.2 non è necessario che al variare di t i grafi $\mathcal{G}_{P(t)}$ siano *rooted* rispetto allo stesso nodo; tuttavia se ciò accade allora vale il corollario seguente.

Corollario 23.1. *Se i grafi* $\mathcal{G}_{P(t)}$ *sono* rooted *rispetto allo stesso nodo* $\forall t$*, allora*

$$\overline{\mathcal{G}} \triangleq \underbrace{\mathcal{G}_{P(n-1)} \circ \mathcal{G}_{P(n-2)} \circ \ldots \circ \mathcal{G}_{P(1)}}_{(n-1)} \quad \textit{è strongly rooted.}$$

Nel caso descritto dal Corollario 23.1 ci si aspetta di avere una velocità di convergenza al *consensus* maggiore rispetto al caso del Teorema 23.2.

Banalmente, come conseguenza del Corollario 23.1, si ha che se $\mathcal{G}_{P(t)}$ è fortemente connesso per ogni t allora si ha convergenza al *consensus*: si tratta infatti di un caso particolare di sequenza di grafi *rooted* rispetto allo stesso nodo.

Le considerazioni appena fatte mostrano che non è così difficile garantire la convergenza (e.g. l'algoritmo *broadcast* consente di giungere al *consensus*): basta infatti che $P(t)$ individui $\forall t$ un grafo semplicemente *rooted* per arrivare comunque al *consensus*. È importante ribadire il fatto che le proprietà dei grafi descritte devono valere per ogni t.

Quanto finora esposto mira all'*agreement* generico; per ottenere l'*average consensus* sono invece richieste condizioni più restrittive, che non verranno discusse in questa sede.

Esempio: stima distribuita Sia dato il modello a rumori scorrelati

$$\begin{aligned} x(t+1) &= x(t) + w(t), \qquad && w(t) \sim \mathcal{N}(0,q) \\ y(t) &= x(t) + v_i(t), \qquad && v_i(t) \sim \mathcal{N}(0,r); \end{aligned}$$

si ricorda che lo stimatore tempoinvariante di Kalman è dato da

$$\hat{x}(t|t) = \hat{x}(t-1|t-1) + K^* \left(\frac{1}{N} \sum_i y_i(t) - \hat{x}(t-1|t-1) \right),$$

con K^* guadagno a regime del filtro. Se c'è la possibilità di far comunicare i nodi in modo che sia disponibile ad ogni istante e per tutti gli agenti la media aritmetica delle misure $z(t) = \frac{1}{N} \sum_i y_i(t)$, le stime locali sono date dalla seguente:

$$\hat{x}_i(t|t) = \hat{x}_i(t-1|t-1) + K^* \left(z(t) - \hat{x}_i(t-1|t-1) \right).$$

Un'idea è quella di procedere in questo modo:

- ad ogni passo T, ogni sensore inizializza $z_i(0) = y_i(t)$;
- su una scala temporale più fitta rispetto a quella usata per le misure, ogni sensore inizia a comunicare con gli altri secondo la dinamica

  $$z(m+1) = Qz(m)$$

  con Q matrice doppiamente stocastica che deve essere opportunamente progettata in modo da consentire la convergenza al *consensus*. In particolare, la doppia stocasticità porta all'*average consensus*, ovvero

  $$z(m) = Q^m z(0) = Q^m \begin{bmatrix} y_1(t) \\ \vdots \\ y_n(t) \end{bmatrix} \xrightarrow{m \to +\infty} z_i(m) \to \frac{1}{N} \sum_i y_i(t);$$

  quindi, se tra una misura e l'altra i nodi riescono a comunicare tra loro un numero sufficientemente grande di volte, ciascuno di essi riesce ad arrivare molto vicino alla media aritmetica delle misure.

In generale, non è detto che le comunicazioni tra una misura e l'altra siano così frequenti, quindi si può non arrivare nell'intorno della media aritmetica; in casi di questo tipo la scelta di K^* come guadagno statico del filtro globale può rivelarsi non ottima ai fini della convergenza all'*average consensus*.

Si chiude la trattazione elencando gli attuali problemi aperti riguardo le tematiche di stima distribuita:

- progettazione della matrice Q di comunicazione;
- calcolo del guadagno statico che ottimizza le prestazioni;
- valutazione delle prestazioni dello stimatore distribuito.