PSC: Progettazione di sistemi di controllo	a.a. 2010-2011
Lezione 18 — 25 Novembre 2010	
Docente: Luca Schenato	*Stesori: Guido Albertin, Elena Toffoli, Giancarlo Baldan*

18.1 Ancora sulla progettazione di P

Vediamo due esempi che generalizzano i risultati ottenuti nell'ultimo esempio presentato la scorsa lezione.

Esempio 1. Si consideri un grafo circolante con N nodi in cui la matrice di comunicazione P sia così definita:

$$P = (1-k)I + k\Pi.$$

Per quanto visto finora l'autovalore n-esimo della matrice P risulta:

$$\lambda_n = \alpha\left(e^{i\frac{2\pi}{N}n}\right) = \sum_{j=0}^{N-1} \alpha_j\, e^{i\frac{2\pi}{N}nj} = 1-k+ke^{i\frac{2\pi}{N}n}.$$

Tali autovalori sono distribuiti, in modo discreto, sulla circonferenza del piano complesso centrata in $1-k$ ed avente raggio k. Un esempio di questa distribuzione è riportato nella figura seguente; si osservi che l'autovalore $\lambda_0 = 1$ è sempre presente e che il secondo autovalore di modulo più alto $\lambda_1 = 1-k+ke^{i\frac{2\pi}{N}}$ si avvicina sempre più ad 1 al crescere di N.

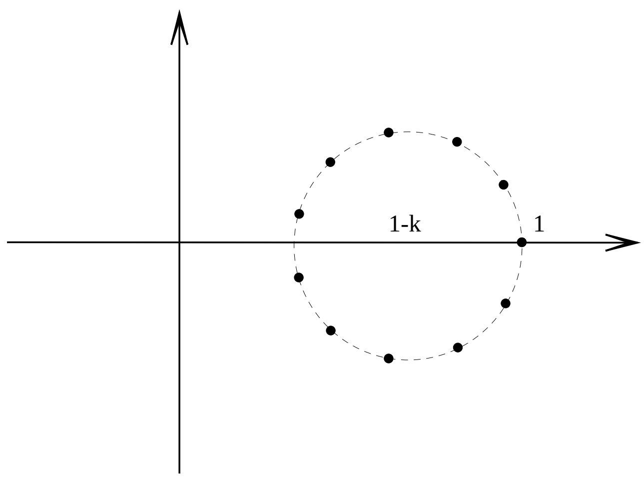


Figura 18.1. Autovalori della matrice P

È possibile fare un'analisi più precisa di quanto l'algoritmo peggiori le prestazioni di convergenza all'aumentare di N studiando il legame tra il modulo di λ_1 ed N:

$$\begin{aligned}|\lambda_1|^2 &= \left[1-k+k\cos\frac{2\pi}{N}\right]^2+k^2\sin^2\frac{2\pi}{N}=(1-k)^2+k^2+2k(1-k)\cos\frac{2\pi}{N}=\\ &= 1+2k(1-k)\left[\cos\frac{2\pi}{N}-1\right]\simeq 1-k(1-k)\frac{4\pi^2}{N^2}.\end{aligned}$$

Si evince facilmente, come già determinato la scorsa lezione, che la velocità di convergenza dell'algoritmo è inversamente proporzionale al quadrato del numero di nodi. In questo caso si può inoltre determinare il valore di k che minimizza il modulo di λ_1 ottenendo $\overline{k}=\frac{1}{2}$ cioè l'algoritmo ottimale prevede che ciascun nodo pesi in ugual misura la sua opinione e quella del vicino.

Esempio 2. Si consideri un grafo circolante con N nodi in cui però ciascun nodo comunica sia col nodo precedente che col successore. La matrice di comunicazione P assume dunque la forma:

$$P=(1-2k)I+k\Pi+k\Pi^{-1}.$$

Gli autovalori di P risultano:

$$\lambda_n=1-2k+ke^{i\frac{2\pi}{N}n}+ke^{-i\frac{2\pi}{N}n}=1+2k\left[\cos\left(\frac{2\pi}{N}n\right)-1\right],\quad n=0,\dots,N-1.$$

In questo caso l'autovalore col secondo modulo più alto non è sempre quello ottenuto con un particolare valore dell'indice. Un'analisi più accurata mostra come tale autovalore sia una funzione di N e k. Inoltre è possibile determinare la scelta ottima del parametro k:

$$\overline{k}=\frac{1}{2-\cos\left(\frac{2\pi}{N}\right)-\cos\left(\frac{2\pi}{N}\lfloor\frac{N}{2}\rfloor\right)}$$

e risulta essere anch'essa dipendente da N; il peso dato da ciascun nodo alla propria opinione ed a quelle dei vicini dipende quindi dal numero complessivo di nodi[1].

Per $k<\overline{k}$ l'autovalore col secondo modulo più grande risulta essere λ_1 mentre per $k>\overline{k}$ l'autovalore dominante è $\lambda_{\lfloor\frac{N}{2}\rfloor}$. Ovviamente per $k=\overline{k}$ i due autovalori hanno lo stesso modulo che risulta essere:

$$|\lambda_1(\overline{k})|^2=\left[1+2\frac{\cos\left(\frac{2\pi}{N}\right)-1}{2-\cos\left(\frac{2\pi}{N}\right)-\cos\left(\frac{2\pi}{N}\lfloor\frac{N}{2}\rfloor\right)}\right]^2\simeq 1-\frac{4\pi^2}{N^2}$$

in cui l'ultima approssimazione, che risulta valida per N grandi, mostra come questo algoritmo sia asintoticamente equivalente al precedente in termini di tempi di convergenza. Si può dimostrare che per rendere $|\lambda_{\max}|^2$ asintoticamente proporzionale ad $\frac{1}{N}$ è necessario avere protocolli di comunicazione che prevedano comunicazioni con agenti lontani almeno $\sqrt{N}$ come avviene, ad esempio, nel caso P assuma la forma $P=(1-k)I+k\Pi^{\sqrt{N}}$.

[1]Si osservi che, in questo caso, si ha $\lim_{N\to\infty}\overline{k}=\frac{1}{2}$ cioè, all'aumentare del numero di agenti, ciascun nodo tende a non considerare la propria opinione ma a mediare solamente quella dei due vicini

18.1.1 Simmetria toroidale

Questo tipo di simmetria si ottiene quando gli agenti possono essere logicamente distribuiti su una superficie toroidale e ciascuno di essi può comunicare con i nodi che sono gli sono più o meno vicini nella superficie. Se svolgessimo la superficie toroidale su una supeficie piana il grafo di comunicazione avrebbe non solo gli archi che collegano i nodi interni ma anche quelli che collegano un "bordo" all'altro come indicato nella seguente figura.

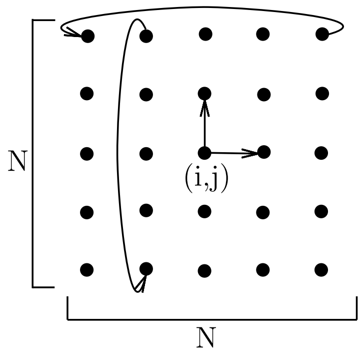


Figura 18.2. Struttura del grafo di comunicazione in presenza di simmetria toroidale

L'aggiornamento del generico agente (i, j) avviene secondo la legge:

$$x_{i,\,j}(t+1) = \sum_{h,\,k} \alpha_{h,\,k} x_{i+h,\,j+k}(t). \tag{18.1}$$

Se, ad esempio, ciascun agente (i, j), oltre al proprio stato, riceve lo stato dei soli due agenti $(i + 1, j)$ e $(i, j + 1)$[2] l'equazione di aggiornamento assume la forma:

$$x_{i,\,j}(t+1) = \alpha_{0,\,0} x_{i,\,j}(t) + \alpha_{1,\,0} x_{i+1,\,j}(t) + \alpha_{0,\,1} x_{i,\,j+1}(t).$$

Per poter analizzare questi tipi di protocolli è necessario rendere matriciale l'equazione di aggiornamento (18.1) e per prima cosa è opportuno inserire gli stati di tutti gli agenti in un unico stato vettoriale x:

$$x = [x_{0,\,0}\;\; x_{0,\,1} \cdots x_{0,\,N-1} | x_{1,\,0}\;\; x_{1,\,1} \cdots x_{1,\,N-1} | \cdots | x_{N-1,\,0}\;\; x_{N-1,\,1} \cdots x_{N-1,\,N-1}]^T$$

In secondo luogo è necessario definire una nuova operazione matriciale nota come prodotto di Kronecker:

$$\begin{array}{cccc} \otimes : & \mathbb{R}^{N\times N} \times \mathbb{R}^{N\times N} & \to & \mathbb{R}^{N^2\times N^2} \\ & (A, B) & \mapsto & \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1N}B \\ a_{21}B & a_{22}B & \cdots & a_{2N}B \\ \vdots & \vdots & \vdots & \vdots \\ a_{N1}B & a_{N2}B & \cdots & a_{NN}B \end{bmatrix} \end{array}$$

[2] Questa è la condizione minima perché il grafo sia strongly connected

e le seguenti due matrici:

$$I \otimes \Pi = \begin{bmatrix} \Pi & & & 0 \\ & \Pi & & \\ & & \ddots & \\ 0 & & & \Pi \end{bmatrix} \quad \Pi \otimes I = \left[\begin{array}{c|ccc} & I & & \\ 0 & & \ddots & \\ & & & I \\ \hline I & & 0 & \end{array}\right].$$

Le matrici appena introdotte descrivono rispettivamente le operazioni di shift per colonna e per riga degli stati come si evince facilmente osservando che:

$$(I \otimes \Pi)x = [x_{0,1} \cdots x_{0,N-1} \;\; x_{0,0} | x_{1,1} \cdots x_{1,N-1} \;\; x_{1,0} | \cdots | x_{N-1,1} \cdots x_{N-1,N-1} \;\; x_{N-1,0}]^T$$

$$(\Pi \otimes I)x = [x_{1,0} \;\; x_{1,1} \cdots x_{1,N-1} | \cdots | x_{N-1,0} \;\; x_{N-1,1} \cdots x_{N-1,N-1} | x_{0,0} \;\; x_{0,1} \cdots x_{0,N-1}]^T$$

Alla luce delle precedenti osservazioni è facile convincersi che l'equazione di aggiornamento (18.1) può riscriversi in modo matriciale come segue:

$$x(t+1) = \left[\sum_{h,k} \alpha_{h,k} \left(\Pi \otimes I\right)^h \left(I \otimes \Pi\right)^k \right] x(t) = Px(t).$$

Sfruttando ora la seguente proprietà del prodotto di Kronecker

$$(A \otimes B)(C \otimes D) = (AC) \otimes (BD),$$

possiamo riscrivere in modo più conveniente la matrice P ottenuta:

$$P = \sum_{h,k} \alpha_{h,k} \Pi^h \otimes \Pi^k.$$

Gli autovalori di P possono essere calcolati osservando che, se V è la matrice che diagonalizza Π, allora $V \otimes V$ diagonalizza P. Riportiamo di seguito la verifica di questo fatto dove viene utilizzata anche l'ulteriore proprietà del prodotto di Kronecker $(V \otimes V)^{-1} = V^{-1} \otimes V^{-1}$.

$$\begin{aligned} (V \otimes V)^{-1} P (V \otimes V) &= \sum_{h,k} \alpha_{h,k} \left(V^{-1} \otimes V^{-1}\right) \left(\Pi^h \otimes \Pi^k\right) (V \otimes V) = \\ &= \sum_{h,k} \alpha_{h,k} \left(V^{-1} \Pi^h V\right) \otimes \left(V^{-1} \Pi^k V\right) = \\ &= \sum_{h,k} \alpha_{h,k} D^h \otimes D^k. \end{aligned}$$

Ricordando come è definito il prodotto di Kronecker e che D ha la forma:

$$D = \begin{bmatrix} u^0 & & 0 \\ & \ddots & \\ 0 & & u^{N-1} \end{bmatrix},$$

si intuisce subito che tutte le matrici $D^h \otimes D^k$ sono tutte diagonali ed hanno la forma:

$$D^h \otimes D^k = \begin{bmatrix} \boxed{\begin{matrix} u^{0h}u^{0k} & & & \\ & u^{0h}u^{1k} & & \\ & & \ddots & \\ & & & u^{0h}u^{(N-1)k} \end{matrix}} & & 0 \\ & \ddots & \\ 0 & & \boxed{\begin{matrix} u^{(N-1)h}u^{0k} & & & \\ & u^{(N-1)h}u^{1k} & & \\ & & \ddots & \\ & & & u^{(N-1)h}u^{(N-1)k} \end{matrix}} \end{bmatrix}.$$

In virtù di questo risultato gli autovalori di P risultano dati semplicemente da:

$$\lambda_{ij} = \sum_{h,\,k} \alpha_{h,\,k} u^{ih} u^{jk}.$$

Esempio Si consideri un grafo di comunicazione con simmetria toroidale in cui l'algoritmo di comunicazione ha la forma:

$$x_{i,\,j}(t+1) = \frac{1}{3}x_{i,\,j}(t) + \frac{1}{3}x_{i+1,\,j}(t) + \frac{1}{3}x_{i,\,j+1}(t).$$

In base alla formula appena determinata gli autovalori che caratterizzano il protocollo sono:

$$\lambda_{nm} = \frac{1}{3} + \frac{1}{3}e^{i\frac{2\pi}{N}n} + \frac{1}{3}e^{i\frac{2\pi}{N}m}.$$

In questo caso si trova che λ_{00} è l'autovalore di modulo unitario mentre i secondi due autovalori di modulo massimo sono $\lambda_{01} = \lambda_{10} = \frac{2}{3} + \frac{1}{3}e^{i\frac{2\pi}{N}}$ ed il quadrato del loro modulo vale:

$$|\lambda_{01}|^2 = 1 - \frac{2}{9}\frac{4\pi^2}{N^2}$$

Come si può facilmente notare in questo caso l'algoritmo si comporta asintoticamente meglio degli algoritmi (locali) che sfruttano simmetrie 1-dimensionali poiché la rapidità con cui cresce $|\lambda_{\max}|^2$ è inversamente proporzionale al numero di agenti e non al suo quadrato[3]. In generale si dimostra che algoritmi locali d-dimensionali possono garantire rapidità di convergenza dell'ordine di $\frac{1}{N^{d/2}}$.

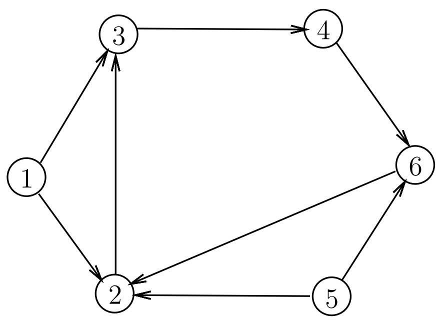
[3]Si ricorda che in questo esempio il numero complessivo di agenti è N^2

18.2 Introduzione agli algoritmi tempo varianti

Ci occuperemo d'ora in avanti di analizzare il caso in cui la matrice di aggiornamento di stato corrispondente ad un certo protocollo di comunicazione sia tempo variante e l'equazione di aggiornamento dello stato del sistema sia quindi del tipo[4]:

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

Questo tipo di protocolli può presentarsi quando gli agenti comunicano la loro informazione uno alla volta ed, a seconda del particolare agente, la matrice P può variare. Per chiarire l'origine di queste problematiche si faccia riferimento al grafo di comunicazione riportato nella figura seguente



e si supponga che ad ogni istante di tempo un solo agente[5] comunichi a tutti i suoi vicini il valore del proprio stato; chi riceve questa informazione farà la media del valore ricevuto con il proprio. Ad esempio se è l'agente 1 a comunicare la matrice P corrispondente sarà:

$$P_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ k & 1-k & 0 & 0 & 0 & 0 \\ k & 0 & 1-k & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

Allora se supponiamo che l'ordine di accensione degli agenti sia crescente la successione delle $P(t)$ sarà

$$P(1) = P_1, \quad P(2) = P_2P_1, \quad P(3) = P_3P_2P_1, \quad \ldots$$

Il problema sta nel determinare sotto quali condizioni la matrice $P(t)$ converge ad una opportuna matrice $\mathbb{1}w^T$ di rango 1. Questo problema non è affatto banale in quanto bisogna analizzare il comportamento di una produttoria di matrici stocastiche.

[4] La matrice $P(t)$ deve essere stocastica $\forall t$

[5] Il particolare ordine di scelta degli agenti può essere aleatoria o predeterminata, per la nostra trattazione non fa differenza

Poiché $P(t)$ è sempre stocastica è possibile porre il problema in un modo più conveniente. Per la stocasticità è infatti sempre possibile trovare una base[6] $S^{-1} = [\mathbb{1} \;\; V]$ tale che, se si indica con $S = [\begin{smallmatrix} q^T \\ Q \end{smallmatrix}]$, permette di fare il seguente cambio di base per $P(t)$:

$$\overline{P}(t) = SP(t)S^{-1} = \begin{bmatrix} q^T\mathbb{1} & * \\ 0 & QP(t)V \end{bmatrix} = \begin{bmatrix} 1 & * \\ 0 & \tilde{P}(t) \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & \tilde{P}(t) \end{bmatrix},$$

in cui l'ultima uguaglianza è valida perché $\overline{P}(t)$ è ancora stocastica e ogni riga deve sommare ad 1.

Poiché $\overline{P}(t)$ e $P(t)$ hanno lo stesso rango verificare la convergenza di $P(t)$ ad una matrice di rango 1 è equivalente a verificare la convergenza a 0 della matrice $\tilde{P}(t)$. Purtroppo anche questo problema non è triviale in quanto non è neppure vero che il prodotto di matrici stabili è ancora stabile[7] e, del resto, nel nostro caso non è neppure detto che i singoli fattori siano stabili.

Un'idea per poter dimostrare la convergenza a 0 della matrice $\tilde{P}(t)$ è trovare una opportuna norma di matrici che sia sub moltiplicativa cioè tale che:

$$\|AB\| \leq \|A\|\|B\|.$$

I problemi che affronteremo allora nelle prossime lezioni saranno

- Determinare che norma utilizzare in modo da riuscire a dimostrare la convergenza a 0 di $\tilde{P}(t)$.

- Capire come è fatto il vettore w^T tale che $lim_{t\to\infty}P(t) = \mathbb{1}w^T$ in modo da capire se l'algoritmo converge al valor medio dello stato iniziale.

[6] Per semplicità si è omessa la dipendenza dal tempo dei vettori della base

[7] Come controesempio si considerino le matrici $\begin{bmatrix} 0.9 & 1 \\ 0 & 0 \end{bmatrix}$ $\begin{bmatrix} 0.9 & 0 \\ 1 & 0 \end{bmatrix}$ che sono stabili mentre il loro prodotto $\begin{bmatrix} 1.81 & 0 \\ 0 & 0 \end{bmatrix}$ non lo è.