Tor vergata
Laurea magistrale in Informatica
Calcolo delle Probabilità 2
A.A. 2018-2019

Lezione 20 del 29/05/2019
Equazione di Kolmogorov

Leonardo Tamiano

Contents

1 Matrice dei parametri infinitesimali

Nella lezione precedente abbiamo introdotto la matrice dei parametri infinitesimali, definita da $Q = (q_{i,j})_{i,j}$, con

$$q_{i,j} = \begin{cases} -\nu_i & , j = i \\ \nu_i p_{i,j} & , j \neq i \end{cases}$$

andiamo adesso a vedere il perchè di tale nome. Per adesso supponiamo per buona la relazione $\dot{P}_t = Q \cdot P_t$, che verrà comunque dimostrata tra poco. Ponendo $t = 0$ troviamo che $\dot{P}_0 = Q \cdot P_0$, con $P_0 = (p_{i,j}(0))_{i,j} = I$. Troviamo dunque, per $t = 0$, la seguente relazione

$$\forall i, j : \frac{d}{dt} p_{i,j}(0) = q_{i,j}$$

Utilizzando lo sviluppo di Taylor nel punto $t = 0$ abbiamo che, in generale

$$p_{i,j}(t) = p_{i,j}(0) + \frac{d}{dt} p_{i,j}(0) \cdot t + o(t)$$

Ma allora troviamo i seguenti casi

- $j = i$
$$p_{i,i}(t) = 1 + q_{i,i} \cdot t + o(t)$$
- $j \neq i$
$$p_{i,j}(t) = q_{i,j} \cdot t + o(t)$$

Dunque in generale notiamo che $q_{i,j}$ è un valore che determina $p_{i,j}(t)$ quando t è molto piccolo, ovvero quando t è *infinitesimale*. Detto altrimenti, abbiamo la seguente relazione

$$p_{i,j}(t) \approx q_{i,j} \cdot t$$

ed è per questo che $q_{i,j}$ è detto il *parametro infinitesimale* di $p_{i,j}(t)$.

Continuiamo notando che la matrice Q non è una matrice stocastica. Infatti abbiamo che $\forall i \in S$, $\nu_i > 0$ e quindi$q_{i,i} = -\nu_i < 0$. Inoltre,

$$\begin{aligned} \sum_{j \in S} q_{i,j} &= q_{i,i} + \sum_{j \in S : j \neq i} q_{i,j} \\ &= -\nu_i + \sum_{j \in S : j \neq i} \nu_i p_{i,j} \\ &= -\nu_i + \nu_i \cdot \sum_{j \in S : j \neq i} p_{i,j} \\ &= -\nu_i + \nu_i \cdot 1 \\ &= 0 \end{aligned}$$

Notiamo infine che possiamo utilizzare direttamente la matrice Q per descrivere una dinamica di markov a tempo continuo. Infatti abbiamo che

(i) Le righe nulle di Q rappresentano stati assorbenti;

(ii) Per ogni $i \in S$ non assorbente abbiamo che $\nu_i = -q_{i,i}$, e $p_{i,j} = \frac{q_{i,j}}{\nu_i}$, per ogni $j \in S$.

2 Calcolo di P_t per dinamica con due stati

Consideriamo la seguente dinamica $(X_t)_{t\geq 0}$, definita su $S = \{0, 1\}$, con matrice di transizione

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

e con $\nu_0 > 0$, $\nu_1 > 0$. Notiamo che per questa dinamica la matrice Q è la seguente

$$Q = \begin{pmatrix} -\nu_0 & \nu_0 \\ \nu_1 & -\nu_1 \end{pmatrix}$$

Andiamoci quindi a calcolare la matrice $P(t)$ per questa dinamica. Utilizzando la relazione $\dot{P}_t = Q \cdot P_t$ alla nostra dinamica troviamo il seguente sistema di Cauchy

$$\begin{cases} \frac{d}{dt} p_{0,1}(t) & = -\nu_0 \cdot p_{0,1}(t) + \nu_0 \cdot p_{1,1}(t) \\ \frac{d}{dt} p_{1,1}(t) & = \nu_1 \cdot p_{0,1}(t) + -\nu_1 \cdot p_{1,1}(t) \\ p_{0,1}(0) & = 0 \\ p_{1,1}(0) & = 1 \end{cases}$$

Ponendo $g(t) := p_{0,1}(t) - p_{1,1}(t)$ troviamo invece quest'altro problema di Cauchy

$$\begin{cases} \frac{d}{dt} g(t) & = -(\nu_0 + \nu_1) g(t) \\ g(0) = -1 \end{cases}$$

la cui soluzione, ottenuta tramite la formula generale, è $g(t) = -e^{-(\nu_0+\nu_1)t}$. Ricordando che $\frac{d}{dt} p_{0,1}(t) = -\nu_0 g(t)$ trovo quindi il seguente sistema di Cauchy

$$\begin{cases} \frac{d}{dt} p_{0,1}(t) = \nu_0 e^{-(\nu_0+\nu_1)} \\ p_{0,1}(0) = 0 \end{cases}$$

Ma allora, utilizzando il teorema fondamentale del calcolo infinitesimale, troviamo che

$$\begin{aligned} p_{0,1}(t) = p_{0,1}(t) - p_{0,1}(0) = & \int_0^t \frac{d}{dt} p_{0,1}(t) \\ & = \in_0^t \nu_0 e^{-(\nu_0+\nu_1)s} ds \\ & = \frac{\nu_0}{\nu_0 + \nu_1} [1 - e^{-(\nu_0+\nu_1)t}] \end{aligned}$$

Procedendo in modo analogo per $p_{1,1}(t)$ si è in grado di dimostrare che

$$p_{1,1,}(t) = \frac{\nu_0}{\nu_0+\nu_1} + \frac{\nu_1}{\nu_0+\nu_1} e^{-(\nu_0+\nu_1)t}$$

Ma allora, utilizzando il fatto che P_t è una matrice stocastica, possiamo calcolarci tutte le entrate di P_t, trovando la seguente matrice

$$P_t = \begin{pmatrix} \frac{\nu_1}{\nu_0+\nu_1} + \frac{\nu_0}{\nu_0+\nu_1} e^{-(\nu_0+\nu_1)t} & \frac{\nu_0}{\nu_0+\nu_1} - \frac{\nu_0}{\nu_0+\nu_1} e^{-(\nu_0+\nu_1)t} \\ \\ \frac{\nu_1}{\nu_0+\nu_1} - \frac{\nu_1}{\nu_0+\nu_1} e^{-(\nu_0+\nu_1)t} & \frac{\nu_0}{\nu_0+\nu_1} + \frac{\nu_1}{\nu_0+\nu_1} e^{-(\nu_0+\nu_1)t} \end{pmatrix}$$

Osservazione 2.1. Notiamo che, al crescre di t, la matrice P_t converge alla matrice

$$P^* = \begin{pmatrix} \frac{\nu_1}{\nu_0+\nu_1} & \frac{\nu_0}{\nu_0+\nu_1} \\ \frac{\nu_1}{\nu_0+\nu_1} & \frac{\nu_0}{\nu_0+\nu_1} \end{pmatrix}$$

Tale matrice, in modo analogo a quanto visto nel caso di catene di markov discrete, rappresenta la *misura invariante* della dinamica.

3 Dimostrazione generale: $\dot{P}_t = Q \cdot P_t$, $\forall t \geq 0$

Andiamo adesso a dimostrare la seguente relazione, nota con il nome di *equazione di Kolmogorov*: $\forall i, j \in S$,

$$\frac{d}{dt} p_{i,j}(t) = \sum_{k \in S} q_{i,k} \cdot p_{k,j}(t)$$

Proof. Fisso $i \in S$. Abbiamo due casi, a seconda della natura dello stato i.

- i **stato assorbente**.

  In questo caso $p_{i,i}(t) = 1$, e $p_{i,j}(t) = 0$, per ogni $t \geq 0$ e quindi, dato che la derivata di una costante è 0, abbiamo che $\frac{d}{dt} p_{i,j}(t) = 0$ per ogni $J \in S$. Infine, dato che i è assorbente, abbiamo che $q_{i,k} = 0$. Ma allora banalmente vale, per $j \in S$,

  $$\frac{d}{dt} p_{i,j}(t) = 0 = \sum_{k \in S} q_{i,k} \cdot p_{k,j}(t)$$

- i **stato non assorbente**.

  Sia $j \in S$ qualsiasi. In questo caso osserviamo che

  $$p_{i,j}(t) = P_i(\tau_i^P > t, X_t = j) + P_i(\tau_i^P \leq t, X_t = j)$$

  dove con τ_i^P indichiamo la v.a. il cui valore è uguale al tempo di soggiorno nello stato i. A questo punto notiamo che, se $j \neq i$, allora il primo addendo è 0. Altrimenti tale probabilità è $P_i(\tau_i^P > t, X_t = i) = P_i(\tau_i^P > t) = e^{-\nu_i t}$. Dunque trovo

$$p_{i,j}(t) = e^{-\nu_i t}\delta_{\{i\}}(j) + P_i(\tau_i^P \leq t, X_t = j)$$

Per quanto riguarda il secondo addendo, notiamo la seguente cosa

$$\begin{aligned}
P_i(\tau_i^P \leq t, X_t = j) &= \int_{-\infty}^{\infty} P_i(\tau_i^P \leq t, X_t = j | \tau_i^P = s) \cdot f_{\tau_i^P}(s) ds \\
&= \int_0^t P_i(\tau_i^P \leq t, X_t = j | \tau_i^P = s) \cdot \nu_i e^{-\nu_i s} ds \\
&= \int_0^t P_i(X_t = j | \tau_i^P = s) \cdot \nu_i e^{-\nu_i s} ds \\
&= \int_0^t \sum_{k \in S : k \neq i} P_i(X_s = k, X_t = j | \tau_i^P = s) \cdot \nu_i e^{-\nu_i s} ds \\
&= \int_0^t \sum_{k \in S : k \neq i} P_i(X_s = k | \tau_i^P = s) \cdot P_i(X_t = j | X_s = k, \tau_i^P = s) \cdot \nu_i e^{-\nu_i s} ds \\
&= \int_0^t \sum_{k \in S : k \neq i} p_{i,k} \cdot p_{k,j}(t-s) \cdot \nu_i e^{-\nu_i s} ds
\end{aligned}$$

Dunque, fissato i, trovo

$$p_{i,j}(t) = e^{-\nu_i t}\delta_{\{i\}}(j) + \int_0^t \sum_{k \in S : k \neq i} p_{i,k} \cdot p_{k,j}(t-s) \cdot \nu_i e^{-\nu_i s} ds$$

Andando ad eseguire il cambio di variaibli $u := t - s \implies \begin{cases} s = 0 & \implies u = t \\ s = t & \implies u = 0 \\ du & = -ds \end{cases}$

ottengo la seguente espressione

$$p_{i,j}(t) = e^{-\nu_i t}\delta_{\{i\}}(j) + \int_0^t \sum_{k \in S : k \neq i} p_{i,k} \cdot p_{k,j}(u) \cdot \nu_i e^{-\nu_i (t-u)} du$$

Andando a derivare notiamo come il secondo addendo sia un prodotto di funzioni. La derivata del secondo addendo è quindi

$$-\nu_i e^{-\nu_i t} \cdot [\int_0^t \sum_{k \in S : k \neq i} p_{i,k} \cdot p_{k,j}(u) \cdot \nu_i e^{-\nu_i u} du] + e^{-\nu_i t} \cdot [\sum_{k \in S : k \neq i} p_{i,k} \cdot p_{k,j}(u) \cdot \nu_i e^{-\nu_i u}]$$

Ma allora la derivata di $p_{i,j}(t)$ è pari a

$$\begin{aligned}
\frac{d}{dt}p_{i,j}(t) &= -\nu_i(e^{-\nu_i t}\delta_{\{i\}}(j) + \int_0^t \sum_{k\in S:k\neq i} p_{i,k}\cdot p_{k,j}(u)\cdot \nu_i e^{-\nu_i(t-u)}du) + \sum_{k\in S:k\neq i}\nu_i p_{i,k}\cdot p_{k,j}(t) \\
&= -\nu_i p_{i,j}(t) + \sum_{k\in S:k\neq i}\nu_i p_{i,k}\cdot p_{k,j}(t) \\
&= \sum_{k\in S} q_{i,k}\cdot p_{k,j}(t)
\end{aligned}$$

che è proprio quello che volevamo dimostrare.

□